

UNIVERSITY OF TOULOUSE CAPITOLE
TOULOUSE SCHOOL OF ECONOMICS

HABILITATION À DIRIGER DES RECHERCHES

to be presented and publicly defended by

Victor Gay

on June 18, 2026

Essays in Economic History and Political Economy

JURY

Julia Cagé, Professor, Sciences Po Paris (rapporteure)

Claude Diebolt, Research Director, CNRS and University of Strasbourg (rapporteur)

Mara Squicciarini, Associate Professor, Bocconi University (rapporteure)

Pauline Grosjean, Professor, University of New South Wales (examinatrice)

Paul Seabright, Professor, Toulouse School of Economics (professeur référent)

Contents

Introduction	1
1 Postdoctoral Research	12
1.1 State-Building and Rebellion in the Run-Up to the French Revolution	12
1.2 Revolutionary Inheritance Reforms and the Demographic Transition	95
1.3 Methodological Contributions	163
2 Ongoing Research	207
2.1 Building on the First World War	207
2.2 Measuring Economic Activity in Early Twentieth-Century France	210
2.3 Other Ongoing Research	212
3 Research Supervision	214
3.1 Master Theses	214
3.2 Doctoral Theses	215
Bibliography	216
Appendix	226
A Doctoral Research on the First World War	226
B Curriculum Vitae	307

Introduction

***Preliminary remarks** This HDR thesis is organized into an introduction, three chapters, and two appendices. The introduction provides a brief overview of my doctoral, postdoctoral, and ongoing research. Chapter 1 brings together the main articles stemming from my postdoctoral research: Albertus and Gay (2025) and Gay (2025c) on state-building and rebellion in the run-up to the French Revolution; Gay, Gobbi, and Goñi (2023a; 2023b; 2025) on revolutionary inheritance reforms and the demographic transition; and Gay (2021; 2023a; 2025d) on my methodological contributions.¹ Chapter 2 presents extended abstracts of some of my ongoing research. Chapter 3 provides an overview of my research supervision activities, including master and doctoral theses. Finally, Appendix A brings together the main articles stemming from my doctoral research (Boehnke and Gay, 2022; Gay, 2023b; Gay and Grosjean, 2023a). Appendix B provides my curriculum vitae.*

My research investigates the process of social change through the lenses of economics and history. During my doctoral studies, I examined the origins and long-run evolution of gender roles, notably by using World War I in France as an exogenous source of variation in women's labor supply to identify mechanisms of intergenerational transmission of gender norms (Section 1). My postdoctoral research shifted to a different period and thematic focus, exploring the political and demographic consequences of state-building in Ancien Régime France (Section 2). Beyond these substantive areas, a common thread in my work is a methodological focus on issues of reproducibility and data curation (Section 3). A further defining feature is its interdisciplinary orientation, drawing on approaches from fields beyond economics, including history, political science, demography, geography, sociology, and computer science (Section 4). These different strands also structure my ongoing research (Section 5).

¹Copies of my other articles, book chapters, conference proceedings, and working papers are available on my website at www.victorgay.me.

1 The Roots and Long-Run Evolution of Gender Roles

My doctoral research examined the origins and long-run evolution of gender norms, with a particular emphasis on women's labor force participation. I have pursued this research agenda along two main lines. First, I studied female labor dynamics in twentieth-century France, using World War I as a historical natural experiment (Section 1.1). Second, I investigated whether linguistic structures can shape gender gaps in economic behavior (Section 1.2). I briefly summarize these strands below and provide the main research articles related to this agenda in Appendix A (Boehnke and Gay, 2022; Gay, 2023*b*; Gay and Grosjean, 2023*a*).

1.1 World War I and Women's Labor Force Participation

My first line of doctoral research examined how the First World War reshaped women's role in French labor markets. The war generated a major demographic shock: between 1914 and 1918, 1.3 million French men died, producing imbalanced sex ratios that persisted for three decades. Although such imbalances can have long-lasting consequences on gender roles (Grosjean and Khattar, 2019; Teso, 2019), these are often difficult to identify because they typically result from factors that gradually shape labor market structures as well (Qian, 2008; Carranza, 2014). In Boehnke and Gay (2022, *Journal of Human Resources*), we address this identification challenge by exploiting exogenous variation in World War I military death rates generated by the French army's cohort-based recruitment system. Therein, we combine département-level census data on labor force participation with an original individual-level dataset of all French soldiers who died in the war (Gay and Grosjean, 2023*a*), and, using a difference-in-differences design, we show that départements with higher military death rates experienced larger increases in women's labor force participation during the interwar period. We attribute this effect primarily to shifts in labor supply: deteriorated marriage-market conditions for single women and negative income shocks among war widows drew more women into the labor force. By contrast, demand-side mechanisms—such as firms substituting female for male labor in response to male scarcity—appear to have played only a limited role. Beyond providing empirical support for a long-standing narrative in the French historiography of the Great War (e.g., Thébaud, 1986), this article also offered some of the first systematic evidence on the effects of war on women's labor outside the well-studied case of the Second World War in the United States (e.g., Acemoglu, Autor, and Lyle, 2004).

In a follow-up article (Gay, 2023*b*, *The Economic Journal*), I examined whether this shock to women's labor supply persisted over the long run using individual-level census data

extending to 2012. Beyond the specificity of the context under study, this article sought to shed light on the forces underlying the twentieth-century transformation of women's work in industrialized economies (Goldin, 2006; Olivetti and Petrongolo, 2016). Existing research on this question typically isolates a single mechanism, such as the diffusion of gender norms from mothers-in-law to daughters-in-law (Fernández, Fogli, and Olivetti, 2004), and rarely disentangles cultural transmission from changes in local labor market structures.² This article advanced on both fronts by adopting the “epidemiological” approach to culture (Fernández, 2011). For instance, I compare women born and residing in the same location but whose mothers were born in locations differentially exposed to the war, thereby varying women's cultural origins while holding their environment constant. Using this design, I show that the intergenerational transmission of gender norms following the war operated through three concurrent channels: vertical transmission from mothers to daughters, oblique transmission through marriage from mothers-in-law to daughters-in-law, and horizontal transmission through local social interactions. Beyond labor market outcomes, this article further shows that this historical shock shaped both women's and men's attitudes toward women's work over the long run. Methodologically, this article also addressed several common pitfalls in persistence studies (e.g., Abad and Maurer, 2021) by tracing outcomes continuously from the initial historical shock to the present, thereby limiting contingencies along the causal chain and strengthening the credibility of the long-run exclusion restriction.

1.2 Language Structures and Gender Roles

The second line of my doctoral research examined whether gender bias embedded in language structures can help explain gender gaps in behavior. This question relates to a long-standing debate in cognitive linguistics over whether language shapes cognition and behavior—the so-called Sapir-Whorf hypothesis. In a series of articles, conference proceedings, and book chapters (Gay et al., 2013a; 2013b; 2016a; 2016b; 2018), we show that speakers of more gender-marked languages exhibit larger gender disparities in economic and political participation, as well as a more unequal division of household labor. A central challenge in this literature, however, is to rule out spurious correlations (Roberts, Winters, and Chen, 2015). We address this issue by adopting the epidemiological approach to culture and comparing first- and second-generation immigrant women living in the same locations but with different ethnolinguistic origins. We find that the association

²A notable recent exception is Aneja, Farina, and Xu (2024), which explores several concurrent transmission mechanisms of the expansion of women's employment during World War I in the U.S. This working paper applies the same methodology as in my article to identify intergenerational transmission mechanisms.

between language structures and behavior holds in both contemporary cross-country and historical individual-level data. Although establishing causality remains difficult, the evidence runs against a strong Sapir-Whorf interpretation: our results suggest that at least two-thirds of the observed relationship between language and behavior reflects cultural norms correlated with language structures rather than a direct linguistic effect on cognitive processes. This work was among the earliest contributions in economics on this question, and the language-structures dataset we introduced has since been widely used (Gay et al., 2017).³

2 State-Building, Social Conflict, and Demographic Dynamics

My postdoctoral research has focused primarily on state-building in Ancien Régime France and its consequences for social conflict (Section 2.1) and demographic change (Section 2.2) in the run-up to the French Revolution. This body of work draws on my five-year interdisciplinary grant, ObARDI ([ANR-20-CE38-0015](#), 2021–26), which sought to improve our understanding of the process through which the Early Modern state was constructed (Gay and Poublanc, 2022). I briefly summarize this research below and provide the corresponding articles in Sections 1 (Albertus and Gay, 2025; Gay, 2025c) and 2 (Gay, Gobbi, and Goñi, 2024a; 2024b; 2025) of Chapter 1. I also briefly discuss my other postdoctoral research (Section 2.3), which has focused mainly on (historical) political economy.

2.1 State-Building and Rebellion in the Run-Up to the French Revolution

How did civil society respond to the state’s long and contested effort to impose social order (Elias, 1994 [1939]; Weber, 1968)? A central component of this process was the expansion of transportation and communication networks designed to extend the state’s reach into society and improve the administration of its territory—what Mann (1984) terms infrastructural power. Assessing the effects of such expansion is challenging, however, because infrastructures typically serve both state agents and the broader population, so existing studies often estimate bundled treatment effects (e.g., Pengl et al., 2026). In Albertus and Gay (2025, *American Political Science Review*), we address this challenge by examining the eighteenth-century expansion of the horse-post relay network in France, a one-sided, state-monopolized courier system that was distinct from both the road network

³A notable exception is an early article by Chen (2013), which studies how future marking in language explains saving, health, and retirement choices.

and the letter-post administration, each of which also served the general public. We combine original data on the rollout of the horse-post relay network with a comprehensive database of prerevolutionary rebellions (Gay, 2025c), and use a staggered difference-in-differences design at the parish level to show that the opening of a relay station increased subsequent local rebellion. We interpret this finding as evidence of backlash against the material consequences of state penetration into local society, as the monarchy sought to centralize authority and expand its informational capacity. Methodologically, our focus on parishes—the most fine-grained administrative unit of the period—also yields sharper identification than is typical in studies of Ancien Régime France, which generally rely on more aggregated levels of analysis (e.g., Waldinger, 2024).

2.2 Revolutionary Inheritance Reforms and the Demographic Transition

Can institutional change shape population behavior, and more specifically demographic dynamics? France's fertility transition in the early nineteenth century has long puzzled social scientists. Standard economic accounts emphasize the quantity-quality trade-off as a central explanation for fertility decline (Becker and Lewis, 1973; Guinnane, 2011). Yet the French transition was the first in history and occurred more than 50 years before industrialization, well before the quantity-quality trade-off had set in. In Gay, Gobbi, and Goñi (2025, *Journal of Political Economy*), we address this important puzzle by testing Le Play's (1875) long-standing hypothesis that the inheritance reforms enacted during the French Revolution—mandating equal partition among all children, including daughters—help explain this anomaly. To do so, we combine individual-level demographic data from historical surveys and online genealogies with original data on prerevolutionary inheritance rules (Gay, Gobbi, and Goñi, 2024a; 2024b), and use both difference-in-differences and regression-discontinuity designs to show that these inheritance reforms reduced completed fertility by 0.5 children, accounting for about one-third of France's overall fertility decline. The evidence suggests that parents reduced fertility to limit further land fragmentation in a society dominated by small landholders. Beyond resolving the French puzzle, this study also shows that legal institutions can durably shape fertility behavior, thereby counterbalancing the recent emphasis in the literature on cultural factors (e.g., Beach and Hanlon, 2023).

2.3 Regime Change and Elite Turnover

A third component of my postdoctoral research has examined regime change and elite turnover in both contemporary and historical settings. A central question in this area is

what drives democratization. While standard theories emphasize the role of revolutions and threats from below (e.g., Acemoglu and Robinson, 2006), Michael Albertus and I argue that such accounts leave unexplained a substantial number of recent democratic transitions—especially in Latin America—in which regime change was initiated by economic elites in the absence of a popular threat. To account for these puzzling cases, in Albertus and Gay (2017, *American Journal of Political Science*), we develop a noisy signaling model showing that, when elites and authoritarian rulers are divided, uncertainty about the type of future autocrats can lead elites to support democratization. We then illustrate the plausibility of this mechanism through a series of otherwise puzzling democratic transitions. In Albertus and Gay (2019), we extend this framework to explain patterns of private investment under dictatorship.

Another important question concerns the determinants of elite turnover in democratic regimes. In Franck and Gay (2025), we address the decline of one of the long-dominant elites of Third Republic French politics: lawyers. Exploiting exogenous changes in electoral constituency boundaries between 1928 and 1936, we show that urbanization contributed to the erosion of lawyers' dominance and to the rise of workers in parliamentary elections.

3 Methodological Contributions

Beyond these substantive research themes, a common thread in my work is a methodological focus on issues of reproducibility (Section 3.1) and data curation (Section 3.2). A research project can be conceptualized as a four-stage process: (1) a theory generation stage, (2) a data construction stage, (3) an analysis stage, and (4) a replication stage. In economics, however, the first three stages are typically bundled into a single journal article, with overwhelming emphasis placed on the analysis stage, while the replication stage is largely neglected. This disciplinary norm creates perverse incentives and often gives rise to publication biases that undermine the validity of research findings (e.g., Brodeur, Cook, and Heyes, 2020). The methodological strand of my work engages with this reproducibility crisis at several levels.⁴ I briefly summarize these aspects below and provide a replication article (Gay, 2023a) as well as a methodological book chapter (Gay, 2025d) and the data article describing my Third-Republic France Historical GIS (Gay, 2021) in Section 3 of Chapter 1.

⁴I discussed these ideas at the 2024 meeting of the European Economic Association in the session “Replication Challenges in Economics and Econometrics” (Dreber et al., 2024).

3.1 Reproducibility

Reproducibility is essential to the credibility of research findings (e.g., Maniadis, Tufano, and List, 2017), and this concern has been central to my research methodology since my doctoral studies. Beyond ensuring that all my articles are fully reproducible and making my original datasets publicly available, I have also contributed to broader replication efforts. For instance, I have published two critical replications of landmark studies (Fernández and Fogli, 2009; Schulz et al., 2019), despite substantial pushback from the original authors (le Bris and Gay, 2023; Gay, 2023a). I have also participated in three collective large-scale replication projects: (1) the SCORE project, which reproduced 146 articles across the social and behavioral sciences (Miske et al., 2026, *Nature*); (2) the Many-Economists project, a many-analysts design involving 148 researchers to assess researcher degrees of freedom in economics (Huntington-Klein et al., 2025, R&R at the *Journal of Economic Literature*); and (3) the R2E project, an ongoing effort to reproduce 30 observational studies in economics.

3.2 Data curation and data articles

The data-construction stage is central to research in economic history. Yet standards of data curation in the field remain uneven, as datasets produced by individual researchers often lack proper documentation, interoperable classifications, and consistent metadata—reflecting a weak incentive structure. For instance, in the course of my work on Third Republic France, I noticed that researchers were repeatedly reconstructing low-quality administrative shapefiles from scratch, often without adequate documentation or clear references to historical sources—a time-consuming and inefficient practice. In response, I created a carefully documented and fully interoperable annual historical geographic information system (GIS) for the Third Republic (1870–1940), covering every administrative level—a database comprising 901 nomenclatures and 830 shapefiles (Gay, 2020; 2021). This historical GIS has since become a landmark resource in the field.⁵

In a methodological book chapter (Gay, 2025d), I argue that data articles provide an incentive-compatible response to issues of data curation and reproducibility in the social sciences. Because my research routinely involves the construction of original datasets (Gay et al., 2017; 2023b; 2023a; Gay, 2020; 2025b; Gay and Grosjean, 2023b; Davoine et al., 2026c), I have written several data articles documenting their historical context, construction procedures, content, and potential biases, supported by systematic quantitative source criticism—an approach inspired by historical methods but extended through the use of quantitative tools. These contributions include data articles on a historical GIS

⁵This work was awarded the first edition of the Open Research Data Prize by the French government in 2022.

of Third Republic France (Gay, 2021), French military fatalities in the First World War (Gay and Grosjean, 2023a), institutional structures (jurisdictions, customs, and gabelles) and social conflict in Ancien Régime France (Gay, Gobbi, and Goñi, 2024a; 2024b; Gay, 2025c; Davoine et al., 2026b), and the interwar Parisian population censuses (Brée et al., 2026).

Beyond my own research, I have also worked to promote a culture of reproducible research at the institutional level. Since 2021, I have served as scientific adviser to the University Data Platform of Toulouse, the local branch of the national data infrastructure PROGEDO. And since 2024, I have co-led the Committee for Open Science at the University of Toulouse Capitole and been a member of the scientific committee of *Data & Corpus*, a new journal dedicated to data articles in the social sciences. I have also delivered dozens of talks on these issues at institutions across France.

4 Interdisciplinary Contributions

Interdisciplinarity has been central to my work since my doctoral studies, reinforced by my position at the Institute for Advanced Studies in Toulouse. For instance, I have collaborated with Margot Dazey—a qualitative sociologist of Islam in France, then an IAST postdoctoral fellow—to investigate how mosques’ visibility shaped electoral support for the National Front in the 1990s (Dazey and Gay, 2025, *Comparative Political Studies*).

More broadly, I have systematically drawn on approaches beyond economics to break disciplinary silos and push the research frontier. My interdisciplinary ANR grant ObARDI has brought together historians, geographers, digital humanists, and computer scientists to study the construction of the Early Modern state by overcoming disciplinary boundaries (Gay and Poublanc, 2022). Our aim was to move past the prevailing metanarrative of state formation (Blockmans and Genet, 1995; Blockmans, Holenstein, and Shläppi, 2009) and represent the shifting and overlapping nature of Ancien Régime institutions without resorting to anachronistic categories. To that end, we have developed ontologies—most notably the Hierarchical Historical Territory (HHT) ontology (Bourel et al., 2022)—that more faithfully capture state-building dynamics. This approach combines computer scientists’ technical expertise, historians’ domain knowledge, and geographers’ representational skills, together with digital humanists bridging these domains. Using this methodology, we have shown, for instance, that the political territory conceptualized by the monarchy for the 1788 Estates-General was heavily contested by local communities, revealing an alternative view of France’s political space “from below” (Gay and Poublanc, 2023). This interdisciplinary orientation also informs my ongoing work with political scientists, archivists, historians,

and historical demographers.

5 Ongoing Research⁶

5.1 Building on the First World War

A substantial part of my ongoing research uses the First World War as a revealing case of how critical junctures reshape societies (Collier and Munck, 2022). In revisiting my study of the consequences of the war for women's work during the interwar period (Boehnke and Gay, 2022), I identified an important limitation: our reliance on département-level labor data prevented us from examining underlying individual-level mechanisms. I address this limitation in ongoing work by assembling individual-level records of women's labor force participation before and after the war from the handwritten nominative census lists of one département (Gay and Kesztenbaum, 2026). These data make it possible to identify more precisely the mechanisms linking sex-ratio imbalances to changes in women's labor force participation. In related ongoing work, I place the French case in comparative perspective by documenting the war's effects on women's labor force participation in England and Wales, a setting that has received little attention from economic historians (Gay and Milner, 2026). We find that the war drew single women into the labor force, whereas—unlike in France—war widows withdrew from it because generous pensions available from the outset reduced the need to offset the loss of their husbands' income.

The consequences of the war, however, extended far beyond women's labor. First, trench warfare devastated land and physical capital in northeastern France. In Gay and Lévêque (2026), we assemble commune-level data on war damage and show that this destruction paradoxically accelerated long-run structural change in the affected areas. Second, sustaining industrial warfare required an unprecedented mobilization of the manufacturing sector. In Alonzo, Gay, and Tallec (2026), we compile detailed records on the hundreds of thousands of state contracts awarded to private firms for war matériel and examine whether they contributed to the country's industrialization. We find that départements that received more war contracts experienced faster postwar manufacturing growth, consistent with the argument that industrial warfare accelerated structural change. Third, the conflict intensified global demands for social rights, most notably women's suffrage (Hicks, 2013; Tripp, 2023). In Cavaillé and Gay (2026), we examine why female suffrage failed in postwar France while it succeeded elsewhere in Europe. We show that this failure resulted from the distinctive features of French political institutions and from

⁶All ongoing projects presented in this section have been presented at external seminars and conferences.

the strategic interaction between the Radicals and the Left across the lower and upper houses of Parliament during the 1920s.

5.2 Measuring Economic Activity in Early Twentieth-Century France

Drawing on my extensive work on Third Republic France, I contend that progress in the field is constrained by the scarcity of spatially localized data on economic activity. While individual-level census data provide, to some extent, credible measures of the local occupational structure (Gay, 2025a; Brée et al., 2026), they cannot recover the local sectoral structure of the economy—a prerequisite for studying structural change during the Second Industrial Revolution and the deepening of the division of labor and the rise of the factory system. To fill this gap, I propose to reconstruct this sectoral structure at the commune level throughout the Third Republic by leveraging archival firm-level records.⁷

5.3 Other Ongoing Research

My other ongoing work spans three topics. First, building on my research on Ancien Régime institutions, I have recently developed a parish-level historical GIS of gabelle (salt tax) regimes in seventeenth-century France, along with information on salt prices and quantities (Davoine et al., 2026b; 2026c). Using these data, we study tax evasion around a 1680 reform that relaxed enforcement in several gabelle districts (Davoine et al., 2026a). This design yields new quantitative evidence on the scale and consequences of tax evasion and shows how it constrained early modern state capacity.

Second, with political scientists Jan Stuckatz and Selina Hofstetter and historian Mikkel Dack, we are assembling a database of 10 thousand denazification questionnaires (Fragebögen) from the U.S. occupation zone in postwar Germany (1945–49) (Stuckatz et al., 2026). These questionnaires record detailed information on past membership in Nazi organizations, retrospective voting, education and occupations, income, assets, and other sociodemographics. This longitudinal individual-level dataset will illuminate why Germans supported and joined the Nazi Party and whether membership yielded material benefits. A key strength of this database is that it includes both NSDAP members and non-members, enabling us to construct credible control groups, a critical element that has been absent in the literature. This project is funded by the French and German research agencies (DeNazDB, [ANR-21-FRAL-0005](#), 2022–26).

Finally, with Sandra Brée, a historical demographer, I am building an end-to-end pipeline to classify occupations in the individual-level Parisian censuses of 1926, 1931,

⁷I have submitted this project to the 2025 ERC Starting Grant competition. While it passed Step 1, it received a “B” at Step 2. I will resubmit a revised application in the next call.

and 1936 (Brée and Gay, 2026), leveraging the POPP database (Brée et al., 2025; 2026). This is a substantial task as these census records contain more than one million distinct occupational strings for nearly nine million individuals. We map these occupations to standard classifications (e.g., HISCO) to produce the first high-resolution occupational and social-class maps of interwar Paris. The resulting data will enable analyses of unresolved questions, including the neighborhood-scale effects of the Great Depression, urban patterns of interwar immigration, and links between social class and the electoral dynamics of the Popular Front.

Chapter 1

Postdoctoral Research

My postdoctoral research has focused primarily on state-building in Ancien Régime France and its consequences for social conflict and demographic change in the run-up to the French Revolution. The main articles associated with this research agenda are provided in Sections 1.1 and 1.2 below. Beyond these substantive themes, a common thread in my work is a methodological focus on reproducibility and data curation. Section 1.3 brings together a replication article, a methodological book chapter, and a data article describing my Historical GIS of Third Republic France.

1.1 State-Building and Rebellion in the Run-Up to the French Revolution

This section provides the following analysis and data articles:

- Analysis article: Albertus, Michael and Victor Gay. 2025. “State-Building and Rebellion in the Run-Up to the French Revolution.” *American Political Science Review*, forthcoming.¹
- Data article: Gay, Victor. 2025c. “The Jean Nicolas Database. The French Rebellion, 1661-1789.” *Data & Corpus*, 1: 15892.²

¹The appendix to this article is available at <https://doi.org/10.1017/S0003055425101342>. The replication material for this article is available at <https://doi.org/10.7910/DVN/M1XRXT>.

²The appendix to this article is available at <https://doi.org/10.46298/dc.15892>. The original dataset associated with this article is available at <https://doi.org/10.7910/DVN/ANQXMQ>.

State-Building and Rebellion in the Run-Up to the French Revolution

MICHAEL ALBERTUS *University of Chicago, United States*

VICTOR GAY *Toulouse School of Economics, France*

Early modern European powers were beset by episodic unrest as they sought to consolidate their authority and build empires. We examine how growing state communication networks and the penetration of society impacted unrest by combining original and detailed parish-level data from pre-Revolutionary France on the expansion of the horse-post relay network with rebellion in this period. Using a staggered difference-in-differences framework, we find that new horse-post relays are associated with more local rebellion. We argue that the main mechanisms are the material consequences of state centralization. New horse-post relays are linked with more rebellion against state agents and associates—the military, police, tax collectors, and the judiciary—that conscripted civilians, enforced taxes and laws, and increasingly monopolized roads. Pre-existing state and administrative fragmentation also mediated this relationship. Our findings have implications for the scholarly understanding of the co-evolution of states and order.

Rebellion and unrest have long been central subjects of inquiry for social scientists and historians. Social conflict reflects underlying discontent as well as competition over power disparities, contested values, perceptions of fairness, and expectations for the future. Because of its critical role in social transformation and long-term development (Skocpol 1979; Thompson 1964; Tilly 1986), scholars have examined how conflict shapes phenomena as wide-ranging as state formation (Slater 2010), political consciousness (Porchnev 1963; Thompson 1971), regime durability (Levitsky and Way 2013), and identity and beliefs (Ladurie 1966). Many early students of conflict focused on episodes of violence and unrest in early modern Europe, a period marked by rapid social, cultural, and economic change alongside the emergence of centralized states (Ladurie 1966; Porchnev 1963; Thompson 1971; Tilly 1986). Because of the historical importance and prominence of such disorder, as well as rapid improvements in data availability and processing, a branch of emerging scholarship has turned to reexamining rebellion and unrest in this era.¹

We contribute to this growing literature by focusing on eighteenth-century France in the run-up to the

French Revolution. While the Revolution has played an outsized role in history and is one of the most studied political events of the past two centuries, the thousands of localized uprisings that erupted across the French territory in the preceding decades have received comparatively little attention. These rebellions often simmered and were periodically suppressed by state agents, yet they ultimately fed into the broader revolutionary upheaval. At the same time, the monarchy heavily invested in state-building throughout the eighteenth century (Burguière and Revel 1989). A central element of this effort was the expansion of the horse-post relay network, which nearly doubled in size over the period. These relays established attended lodging quarters and a well-prepared set of fresh horses for messengers carrying information for the royal administration where there had previously been no support. The horse post represented one of the primary means of consolidating the hierarchical French state's informational capacity as it sought to rule and control a fragmented territory (Marchand 2006).

Did this intensive expansion of the state's informational capacity through the horse-post relay network generate backlash among populations exposed to these state-building efforts? To investigate this question, we combine original data on the development of the horse-post relay network over the eighteenth century with a comprehensive database of rebellions in pre-Revolutionary France. Using a staggered difference-in-differences framework centered on parishes—the most granular territorial unit then—we show that the introduction of a new horse-post relay in a parish was associated with more local rebellions in subsequent decades. We attribute this finding mainly to the material consequences of state efforts at penetrating and ordering local society as it centralized and enhanced its informational capacity.

The increasingly present state that was behind and facilitated by new horse-post relays became a target of popular ire for its activities, such as heavy-handed efforts

Corresponding author: Michael Albertus , Professor, Department of Political Science, University of Chicago, United States, albertus@uchicago.edu.

Victor Gay , Assistant Professor of Economics, Department of Social and Behavioral Sciences, Toulouse School of Economics, France, victor.gay@tse-fr.eu.

Handling editor: Carles Boix.

Received: October 24, 2024; revised: February 21, 2025; accepted: November 12, 2025.

¹ See, for instance, Chambru (2019), Degraeve (2023), Chambru and Maneuvrier-Hervieu (2024), Waldinger (2024), and Davoine, Enguehard, and Kolesnikov (2025).

at forced conscription, tax enforcement, monopolization of the roads, and inflexible enforcement of royal acts and laws perceived as unfair (Aubert 2015; Nicolas 2002). Accordingly, we find that horse-post relays were strongly associated with rebellions against agents with coercive powers to maintain and enforce order: the military, the police, and the judiciary. They were also associated with rebellions against various forms of taxation as well as rebellions by notables that had private interests along postal roads. Furthermore, while these relationships broadly prevail and encompass the lion's share of rebellions, we also find a disproportionate response of rebellions to horse-post relays in places of pre-existing state fragmentation and territorial divisions. These findings demonstrate an increasingly powerful state that nonetheless struggled to control local activity across a heterogeneous territory as many of its efforts were contested.

Our empirical approach and additional data give us a unique opportunity to evaluate and challenge several alternative explanations—namely, the erosion of traditional social hierarchies, the role of war zones and war involvement, information and collective action possibly spurred by the contemporaneous letter-post system, changes in the recording of rebellions by police brigades (*maréchaussée*) charged with keeping order, and the contemporaneous transit infrastructure. These latter three alternatives are often challenging to dismiss in more commonly studied cases where communication and transportation networks are built together from scratch and can be widely used. In eighteenth-century France, however, the horse-post relay network was built atop existing transit routes and served as a state-monopolized communications network. Moreover, the development of police brigades and the letter-post system followed a different logic than the horse post.

Our findings have implications for the scholarly understanding of the co-evolution of states, order, and disorder, as well as for the vast corpus of work on the origins of the French Revolution. While greater state capacity may ultimately support political stability and order, the process of state-building itself can be disruptive to pre-existing social structures and contested, even for decades at a time. This process likely fueled the accumulation of grievances and repertoires of resistance that ultimately exploded during the Revolution.

THE LONG AND CONTESTED PROCESS OF STATE FORMATION

There is a broad, rich literature on state formation, including the emergence of states, state centralization and consolidation, and the exercise of power. This literature makes clear that building a strong state is a long, complex, and often contested process. As Elias ([1939] 1994) emphasizes, it has taken most states centuries to overcome powerful fragmented regional and religious elites that had private armies and taxation rights, and many states around the world still face fundamental weaknesses (Centeno et al. 2017; Soifer

2015). In Europe, monarchies slowly absorbed feudal capacities as they centralized their administration and bureaucracy, a process that Elias ([1939] 1994) describes as not only coercive and political but also social, as regional consciousness declined and increasing state presence in society created “civilizing” norms of state control over the use of violence.

Weber (1968) prominently described the newly powerful states that ultimately resulted from this process as defined by a monopoly over the use of violence within their territory, which entails the legitimation of authority and a centralized hierarchy of administration that embodies and enforces a legal-rational domination of society. Subsequent seminal scholarship elaborated concepts, such as state capacity and autonomy (Evans, Rueschemeyer, and Skocpol 1985) and underscored the importance of a strong state to establishing and maintaining social order (Fearon and Laitin 2003; Huntington 1968; Skocpol 1979).²

State formation in Europe advanced through both external pressures and internal incentives and innovations. Tilly (1990) emphasizes how military and fiscal demands associated with regional competition and warfighting pressured early states to accelerate their efforts at centralization and extraction in order to build the capacity to wage war and defend their territory.³ Eventually, state efforts to penetrate society and build capacity became fortified by institutional innovations that responded to disparate interests and in doing so, fostered greater compliance with extraction (Levi 1989).

However, most states spent numerous decades and even centuries attempting to penetrate society and centralize control and administration. A core component of those efforts has been the expansion of transportation and communication routes to facilitate the penetration of the state's reach into society and the administration of its territory. This is closely tied to what Mann (1984) termed infrastructural power—namely, the capacity of the central state to penetrate its territories and logistically implement its decisions. This implicates both the state's material capabilities and its territoriality (Soifer 2008). As Mann (1984) underscores, infrastructural power can vary separately from the despotic power used to impose mandates over society. It can also be very uneven across the territory (Boone 2003; Pierskalla, De Juan, and Montgomery 2019).

Historically, states and empires built paths, stone roads, postal routes, ports, and canals to reach remote populations, transmit information and correspondence, and deploy military forces (Bulliet 1990; Rogowski et al. 2022). Today, governments continue to build and improve roads in remote rural areas, expand railway networks, construct telecommunications infrastructure, and establish stations and outposts for their security

² Of course, states can deploy that same capacity to perpetrate violence against civilians or armed groups (Straus 2006).

³ Recent scholarship identifies similar pressures in Latin America (Schenoni 2025). Internal threats can also provoke state-building, as Slater (2010) demonstrates in select postcolonial Southeast Asian countries where elites coalesced behind strong states and authoritarianism to protect against systemic popular unrest.

apparatus. Alongside these efforts—and in some cases facilitated by them—states typically seek to centralize their administration, strengthen tax collection and police presence, and provide public goods as they broaden their ability to project power and penetrate more deeply into society (Centeno et al. 2017; Lindvall and Teorell 2016; Müller-Crepon, Hunziker, and Cederman 2021).⁴

Disruption and Contention During State Centralization

While stronger states can effectively monopolize violence, enforce rules, and centrally administer territory and social groups (Weber 1968), the process of state-building itself can be contested and contentious (Boone 2003; Garfias and Sellars 2022). We build on seminal theories of state formation that emphasize the development and consolidation of a centralized administrative network across the territory. Transportation and communication infrastructures that connect localities more directly to central government authorities enhance the ability of the state to penetrate society through the control and administration of local activity and rules, extraction in the form of taxation and military service, and the disruption of existing social relations and expectations. In France, as throughout Europe, this historic process entailed the extension of authority over—and even the absorption of—disparate regional nobles, lords, church authorities, and other elites (Elias [1939] 1994).

Our focus, however, highlights how this process was characterized by frictions with society more broadly, and to a lesser degree, by frictions with regional elites. Indeed, the material consequences of increased but incomplete state power can generate social contestation and popular resistance stemming from the greater local capacity of the state to enforce taxes, carry out rules of justice, and conscript civilians into military service. It can also result from the crowding out of private interests and activities in spaces that the state comes to dominate. These consequences of state expansion disrupt status-quo activities and expectations, generating grievances.⁵ At the same time, the state may lack sufficient coercive power and legitimacy to deter popular challenges, a dynamic that has been documented in other contexts of state-building (e.g., Pruett 2024).

The extent of contestation stirred by state efforts at expansion and centralization may also vary across the territory in accordance with pre-existing fragmentation in authority and administration. Centralizing states routinely encounter regional divisions linked to historical competing authorities and power centers across

their territorial expanse, which can, as Migdal (1988) highlights, condition the interactions between society and the state. That poses a challenge to the state's ability to effectively penetrate and order society in a consistent fashion across space and can raise regionalized grievances and backlash. This raises the potential for the link between state-building and unrest to be characterized at least in part by heterogeneity across space in accordance with fragmentation.

We explore these dynamics in pre-Revolutionary France, where increased connectivity and penetration of society associated with state centralization fostered disruption and social unrest. While this relationship prevailed across the entire French territory, it also varied to some extent with pre-existing administrative fragmentation: where fragmentation mapped onto historic disparities in the mode of ruling and extent of intended extraction by the state, opposition to the state's imposition of central authority through new horse-post relays varied accordingly.

STATE-BUILDING AND REBELLION IN THE RUN-UP TO THE FRENCH REVOLUTION

Eighteenth-century France remained an overwhelmingly rural society, undergoing gradual yet profound transformations with deep historical roots (Ladurie 1966). In the countryside, the state, the nobility, and the Church competed for dominance, generating persistent frictions and local conflicts. By Lefebvre's (1947) classic account, this competition gave rise to a dense web of taxes, tithes, and seigneurial dues that weighed increasingly heavily on the peasantry in run-up to the Revolution. Canonical studies have examined how these burdens spurred grievances and uprisings, and resulted in a revolution (de Tocqueville [1856] 2010; Shapiro and Markoff 1998). Recent scholarship points to additional factors behind peasant revolts during this period, such as adverse climatic shocks (Chambru 2019; Waldinger 2024) and the degree of centralization of the tax collection system (Degraeve 2023).

There is broad consensus, however, that the monarchy made significant strides during this period in centralizing and extending its authority, as its influence spread across the kingdom and into peripheral provinces (Burguière and Revel 1989). As the state expanded and assumed a more prominent role in controlling daily life—though still far from an all-encompassing one—it disrupted the delicate balance of relationships, obligations, and expectations that had long structured society. A key element of this expansion was the monarchy's exclusive communication network: the horse post (*poste aux chevaux*).

The Functioning of the Horse-Post Relay Network

From its inception under Louis XI in the late 1470s, the horse post served as a pivotal state-building instrument of the royal authority to consolidate its informational capacity across the kingdom, particularly in its outlying

⁴ Infrastructure can also enable state repression and appropriation of resources like land, as occurred in Paraguay under Stroessner (González et al. 2025). On the interrelationships between various dimensions of state capacity, see Hanson and Sigman (2021).

⁵ Transportation and communication networks could also facilitate collective action by lowering transaction costs and enabling coordination (García-Jimeno, Iglesias, and Yildirim 2022; Pengl et al. 2025). We examine these alternative mechanisms below.

FIGURE 1. Depictions of the Horse Post

(a) A Horse-Post Relay



(b) A Horse-Post Postillon

Note: This figure displays contemporary depictions of an eighteenth-century horse-post relay in Panel a and a postillon galloping two horses back to a relay in Panel b. Sources: in Panel a, lithography by Victor-Jean Adam and Louis-Philippe-Alphonse Bichebois, based on a painting of Jean-Antoine Duclaux, *La malle au relai* [sic], 1817; in Panel b, painting of the French School, *Postillon en livrée ramenant deux chevaux au galop*, mid-nineteenth century. © Musée de La Poste, La Poste, photographs by Thierry Debonnaire.

provinces (Marchand 2006, 183–221).⁶ This infrastructure consisted of a nationwide network of relays where horse-post messengers (*postillons*) could exchange their horses for fresh ones in order to ensure the rapid transportation of directives, reports, and other communications of the royal administration (Jamaux-Gohier 2001; Marchand 2006; Gazagnadou 2013, 109–16). Relays were spaced at regular intervals of 10–15 kilometers, allowing riders to switch horses after an hour or two of travel. This system enabled postillons to maintain high speeds while minimizing the risk of injury to the horse, in line with their physiological limits (Minetti 2003).⁷ The horse-post relay network was connected by a road infrastructure maintained by the Bridges and Roads (*Ponts et Chaussées*) administration, which operated independently of the horse post (Arbellot 1973).

Located at the entrances to towns or at rural crossroads, horse-post relays were relatively large buildings—comprising stables, lodging quarters, and often a tavern—that stood out prominently in the rural landscape (see Panel a of Figure 1). The postillons who galloped postal roads day and night were typically young local laborers, easily identifiable by their distinctive uniforms: a royal-blue vest, a round leather hat,

leather breeches, heavy black boots, and an arm badge bearing the name of their relay (Jamaux-Gohier 2001, 237–40; Marchand 2006, 258–64). Their horses, bred in royal stud farms, were superior to those of private horse renters and bore distinctive markings, such as a knotted or trimmed tail (Jamaux-Gohier 2001, 262–7; Marchand 2006, 270–5). Together, postillons' uniforms and horses made them instantly recognizable in the public space (see Panel b of Figure 1).

In addition to the short distance between relays—which allowed for frequent changes of horses—a series of regulations enabled the horse-post system to achieve a higher speed of information transmission than alternative postal systems available to the general public, such as the letter post or private messenger companies (Jamaux-Gohier 2001, 100–8).⁸ Indeed, postillons held a monopoly on galloping and on traveling at night on roads connecting relays, which were generally better maintained by the Bridges and Roads administration than other roads and directly connected to administrative centers (Arbellot 1973, 766). The horse post also held a monopoly on horse rentals along these roads, effectively preventing private travelers from switching horses when traveling along them.⁹ In addition, it had

⁶ The French horse-post system emulated the organization developed by the Visconti in the Duchy of Milan beginning in the 1380s, which in turn drew inspiration from earlier systems under the Mongol Empire and the Mamluk Sultanate in Egypt in the thirteenth and fourteenth centuries (Gazagnadou 2013, 95–109).

⁷ In fact, similar distances between relays were a common feature of horse-post systems throughout history—including those of the Tang Dynasty in seventh-century China, the Mongol Empire in the thirteenth century, the Mamluk Sultanate in fourteenth-century Egypt, the Duchy of Milan in fifteenth-century Italy, and the Pony Express in the nineteenth-century United States (Minetti 2003, 786).

⁸ For instance, when the *subdélégué* of Rennes sent a letter to his counterpart in Nantes at 7:15 p.m. on Monday, December 31, 1764, the postillon of the horse post arrived in Nantes by 6 a.m. the next morning—after 11 hours of travel through 11 relays (Jamaux-Gohier 2001, 39–40). By contrast, according to the 1763 edition of the *Guide des lettres* (61), a letter sent via the letter post would leave Rennes only on Wednesday at 4 p.m. and arrive in Nantes on Thursday at 7 a.m., requiring 15 hours of travel (Cauvin, Lepetit, and Reymond 1987).

⁹ The vast majority of the population traveled through alternative, regulated, modes: messenger services (*messageries*), coaches and

priority access to hay for feeding horses and the right of passage on congested roads. Postmasters (*maîtres de postes*) profited substantially from these privileges and were quick to report violations to their *intendant*—the provincial representative of the monarchy—who usually swiftly enforced these regulations (Vaillé 1953, 166–70; Jamaux-Gohier 2001, 101–4).

The establishment of horse-post relays provided the opportunity for the royal authority to form alliances with and co-opt local elites who came to serve as postmasters. Postmasters were typically recruited among wealthy local notables—bourgeois farmers and innkeepers aspiring to noble status (Jamaux-Gohier 2001, 189–94; Marchand 2006, 223–8). Each held a postal certificate (*brevet de poste*) that granted them the exclusive right to operate a given relay. The certificate imposed several obligations on postmasters: they were required to reside at the relay, to manage their relay directly (*faire-valoir direct*) and not farm out its operations, and to maintain its proper functioning—such as keeping at least five or six healthy horses available at all times. In addition, they were expected to serve as intelligence agents and auxiliary police on behalf of the central administration (Vaillé 1953, 268–71; Marchand 2006, 210–4). Their duties included monitoring passengers passing through the relay—by the late eighteenth century, this involved keeping records of names, occupations, and destinations—as well as reporting notable local political events to their *intendant*.

Overall, the horse-post relay network positioned the state as the sole authority controlling rapid communications in France, significantly expanding its ability to monitor society and exert control over its territory.

Consequences of the Horse Post for Local Order and Rebellion

The development of the horse-post relay network had several distinct consequences. At the national level, it enabled the rapid communication essential to the effective functioning of the state. For instance, it allowed central officials, tax farmers (*fermiers généraux*), and provincial *intendants* and local tax officials (*élus*) to coordinate swiftly through directives, reports, and fiscal orders on matters of tax collection and enforcement. That included efforts to suppress salt smuggling (*faux-saunage*) across the kingdom's disparate and consequential salt-tax zones (*gabelles*). This coordination required frequent and timely exchanges between central administrators and local agents, who were responsible for estimating tax yields and implementing their collection (Caillou 2005, 177–293; Deharbe 2010, 243–344). The

carriages (*coches* and *carrosses*), or stagecoaches (*diligences*) following the 1775 Turgot reform (Marchand 2006, 43). These contractors were not allowed to set up relays and had to travel with the same carriage, either walking or trotting, during the day, and on a regular schedule. They also had very local perimeters: by the early 1790s, half of the 915 *messagerie* services covered a distance under 40 kilometers (Marchand 2006, 84). The transportation of merchandise heavier than 25 kilograms was the privilege of private companies (*voituriers* and *rouleurs*).

speed of communication was especially critical in times of crisis, such as in the aftermath of natural disasters (Favier 2007). The horse-post system also facilitated the transmission and enforcement of military recruitment orders, enabling directives to travel efficiently from central authorities to *intendants* and onward to local recruitment and conscription officers (Corvisier 1964, 147–61).

Beyond these nationwide implications, the expansion of the horse-post relay network disrupted local social equilibria. A range of local actors benefited from these developments. Chief among them were postmasters. In addition to the profits they derived from their exclusive right to operate a relay, postmasters enjoyed a wide range of privileges, including substantial exemptions from taxation—most notably, from the onerous *taille*—as well as relief from military obligations, from accommodating passing soldiers, and from churchwarden responsibilities (Demangeon 1995, 44–8; Tonnaire 2005, 48–59). They were also exempt from taxes on ancillary activities tied to relay operations, such as the sale of wine in the relay's tavern. These material advantages fostered broad loyalty to the monarchy among postmasters (Vaillé 1953, 233–48; Jamaux-Gohier 2001, 189–235; Marchand 2006, 200–46).¹⁰ Other local groups also derived material benefits from the introduction of horse-post relays: innkeepers benefited from the increased traffic along roads connecting relays (Tonnaire 2005, 59–63), and families allied with postmasters often enhanced their local influence through closer ties to representatives of the royal authority (Tonnaire 2005, 48).

Local elites also likely benefited from the establishment of horse-post relays, as improved communication with provincial and central authorities may have enhanced their influence and elevated their status. For instance, historical accounts suggest that tax agents often relied on the knowledge and assistance of local notables as they traveled along the relay network in preparation for their annual round of tax collection (Caillou 2005, 177–293; Deharbe 2010, 243–344). These interactions may have fostered mutually beneficial relationships, reinforcing the authority of local elites and further integrating them into the state's administrative and logistical apparatus. Local notables may also have engaged with military officials and royal agents passing through or lodging at relays, for instance, by sharing local information, though such ties were likely kept discreet and remain difficult to corroborate in historical records.

By contrast, the various monopolies associated with the horse post crowded out competing private interests. First and foremost, horse-rental businesses were severely disrupted by the ban on renting their horses along roads connecting horse-post relays. Indeed, examples of horse renters attempting to violate this

¹⁰ Historical accounts indicate that many postmasters displayed portraits of French kings in their relays. In addition, during the Revolution, they were often accused of collaborating with the monarchy, and a significant number of them were subsequently imprisoned or executed (Marchand 2006, 216–7).

ban abound (e.g., Jamaux-Gohier 2001, 101–4).¹¹ Local farmers and innkeepers situated off postal roads also lost out, as postmasters had priority access to hay and animal fodder for their relays—a shortage potentially worsened by the increased demand brought on by rising traffic along postal roads (Tonnaire 2005, 61–4). More broadly, the monopolization of these roads by the monarchy through the horse post imposed constraints on travelers as privately owned horses were restricted to trotting during daytime and could not be exchanged at horse-rental establishments.¹²

Additional social groups, such as commoners, peasants, and merchants, may have also lost out with the establishment of horse-post relays as they brought increased surveillance and stricter enforcement of state regulations. Indeed, enhanced state communication through the horse-post relay network likely made it easier for authorities to monitor and enforce tax collection, military conscription, and judicial decisions. In addition, local residents could have been subjected to new burdens, including maintaining postal roads through the unpaid labor system of the *corvée* (Tonnaire 2005, 65–86; Conchon 2016).

We hypothesize that these disruptions to local society fueled grievances, and eventually rebellion, though the material consequences of a new horse-post relay likely took time to fully materialize and for discontent to boil over. Adverse events such as repeated food shortages, tax enforcement campaigns, labor requisitions, or conscription orders, combined with policy changes such as tax increases or stricter judicial procedures or penal codes, could have tipped simmering discontent into open rebellion. State agents who contributed to—or were perceived as contributing to—these pressures could have become targets of local hostility. Counterfactual parishes that did not host new horse-post relays were less exposed to such state intrusion and thus avoided the same stratifying pressures that divided communities into winners and losers of state centralization. Hence, residents of these parishes may not have experienced these adverse events to the same degree, and where they did experience them, they might have been less inclined to attribute them to the presence or actions of state agents.

Several rebellious events in our data exhibit just such reactions against state agents and activities. One example occurred in 1781 in the town of Mirande, in the southwestern *généralité* of Auch. A decade earlier, the establishment of a new horse-post relay marked a deeper state presence in this formerly isolated region of the kingdom, as before the 1770s, this *généralité* had

only two relays, in Pau and Orthez. Military recruitment in the area had long been among the lowest in France, but the increased needs resulting from the Anglo-French War (1778–83) brought renewed scrutiny to local enlistment.¹³ In this context, the new horse-post relay facilitated information gathering on draft-eligible men. In 1781, as the *subdélégué* (the local representative of the monarchy), military officers, and local notables assembled in Mirande to conduct a conscription draw, draft-eligible peasants and farmers emerged from the forest with sticks and knives, shouting at the authorities, “They should all be killed!” A confrontation over draft resistance followed, resulting in several arrests before a local notable managed to defuse the situation.

A rebellion in the coastal town of Blaye in Gironde in 1783, just three years after the introduction of a horse-post relay, offers another revealing example. Members of a *patache* brigade—widely despised mobile tax-farming agents charged with collecting gabelles and seizing contraband—attempted to inspect the load of a Breton ship.¹⁴ Following the establishment of a series of new relays connecting La Rochelle to Bordeaux along the eastern shore of the Gironde Estuary in the early 1780s, *patache* patrol boats began to more closely monitor maritime commerce in the region, including in Blaye. When the *patache* approached, the ship’s crew resisted inspection and joined forces with local residents to stone the *patache* agents once ashore at the encouragement of the ship’s captain and the owner of the merchandise. Members of the crew denounced the *patache* brigade as an “enemy of society” that had come to “outrage citizens” who were merely going about their usual business. In short, heightened tax enforcement facilitated by the recent expansion of the horse-post relay network sparked a wave of local resentment that quickly escalated into open resistance.

Other examples illustrate how the introduction of horse-post relays could have crowded out private interests. About a decade after a series of relays was established along the route between Dijon and Auxerre, members of the local bourgeoisie complained that inns near the new relays were monopolizing supplies of food and animal fodder (Tonnaire 2005, 63–4). In response, the *subdélégué* issued a ban preventing innkeepers from making early purchases at local markets to preserve access for other buyers—setting, for instance, an 8 a.m. cutoff in Avallon. Tensions between innkeepers and the bourgeoisie escalated, and although the restriction remained in place despite numerous complaints from innkeepers, many of the underlying grievances persisted. In another such example in 1775 in Picardy, in the context of the Flour War, a crowd of townspeople

¹¹ In fact, the horse post did not reach Brittany until 1738—following an earlier, failed attempt in 1648—partly due to the opposition of local messenger services and horse-renting companies, who feared the competition would undermine their profits (Jamaux-Gohier 2001, 11–28; Marchand 2006, 187–8).

¹² Wealthier travelers may have benefited from the expansion of the horse post, as they could afford to travel in stagecoaches (*diligences*), which could exchange horses at relays following the 1775 Turgot reform—although after 1782, this privilege depended on the postmaster’s discretion (Marchand 2006, 33–6, 88–90).

¹³ While the nationwide average recruitment rate in 1716 was 525 men per 100,000 inhabitants, it was only 327 in the *généralité* of Auch. In 1763, the national average stood at 523, compared to 301 in Auch (Corvisier 1964, 415–27). See also Figure E6 in the Supplementary Material.

¹⁴ There are numerous documented cases of violent attacks targeting *patache* brigades, as well as reports of their abusive conduct toward local populations through aggressive searches, raids, and extortion (e.g., Briais 1984).

in Chauny attacked and ransacked the local horse-post relay on rumors that the postmaster had stockpiled wheat despite the local shortage. A detachment of troops was unable to stop the crowd.

Of course, horse-post relays could have impacted rebellion through other channels beyond those illustrated in these examples. For instance, they could have generated local income shocks, disrupted traditional social hierarchies, introduced more policing, and called attention to the recording and registration of rebellious activities. We examine these and other potential alternative mechanisms below.

DATA

Our analysis dataset comprises a panel of 35 thousand parishes of the kingdom of France as of 1789—except for Corsica—spanning eight decades from the 1710s to the 1780s. Parishes were the smallest units of administration in Ancien Régime France and no prior study has analyzed this period at such a granular level.¹⁵ Further details on data sources are provided in the Data Appendix section of the Supplementary Material. For replication data, see Albertus and Gay (2025).

The Horse-Post Relay Network

To document the development of the horse-post relay network over the eighteenth century, we construct an original historical geographic information system (GIS) based on administrative sources.¹⁶ Beginning in 1708, the General Superintendent of Posts had a list of horse-post relays printed and updated annually: the *Liste générale des postes de France* (Arbellot 1980, 100–2; 1992, 110–4). This publication was designed to provide state messengers with the locations of horse-post relays together with distances between them. Each *Liste des postes* was organized by itinerary, with departure and destination relays clearly marked.¹⁷ Distances were reported in *postes*—a measure of distance equivalent to 8–12 kilometers (Bretagnolle and Verdier 2006, 68–9).¹⁸

We digitized one edition of the *Liste des postes* per decade from the 1710s through the 1790s.¹⁹ From each

edition, we extracted all the segments between relays as listed by itinerary and matched each location to the corresponding parish. We then reconstructed itineraries by drawing straight lines between successive relays and combining the resulting segments. Figure 2 displays the spatial distribution of the horse-post relay network at the beginning and end of our period of analysis.²⁰ We also report the number of horse-post relays along with the length of associated postal roads for each edition of the *Liste des postes* included in our sample in Panel a of Figure 3.

The spatial configuration of the horse-post relay network at the beginning of the eighteenth century reflected the monarchy's long-standing strategic and military concerns vis-à-vis neighboring adversaries, consistent with accounts of subnational state-building efforts that emphasize responses to prevailing geopolitical and bellicist threats (Boone 2003; Tilly 1990). In the late fifteenth century, relays were established in Bordeaux, Dijon, Lyon, Arras, and Amiens due to their strategic positions relative to the Kingdoms of Spain and England, the Duchies of Burgundy and Savoy, the County of Flanders, and the Holy Roman Empire (Gazagnadou 2013, 112). During the sixteenth century, postal routes expanded toward the kingdom's borders, with relays established in Bayonne under Louis XII and in Nantes under Henry III. This strategic expansion continued throughout the seventeenth century, following territorial conquests: northeastward to Strasbourg after the Peace of Westphalia (1648), southwestward to Narbonne after the Treaty of the Pyrenees (1659), and eastward to Besançon after the Treaties of Nijmegen (1678–79) (Marchand 2006, 186–8). These strategic and military concerns led to a center-periphery logic resulting in a star-shaped network radiating from Paris and reinforced by key nodes along the kingdom's borders.²¹

In contrast, the eighteenth century was a period of internal densification of the horse-post relay network (Bretagnolle and Verdier 2005; 2007; Verdier 2009). By then, the horse post had become a key instrument in the monarchy's strategy to assert a more uniform and centralized control over the territory and to project its power into the remotest parts of the kingdom (Marchand 2006, 189).²² The network first densified in the north following the permanent establishment of the royal court at Versailles in 1682 and the consolidation of the kingdom's territorial boundaries. From the

¹⁵ Parishes were the primary units of local organization until they gave way to *municipalités* in 1789–93 and later to *communes* (Gorry 2008).

¹⁶ The horse-post relay network data we construct is more comprehensive than that in existing literature (Bretagnolle and Verdier 2005; 2007; Verdier et al. 2025). We provide more details on this point in the Data Appendix section of the Supplementary Material.

¹⁷ Figure A1 in the Supplementary Material displays an excerpt from the 1751 edition of the *Liste des postes* that depicts the 29 horse-post relays along the postal route from Lyon to Marseille.

¹⁸ The term *poste* referred both to a relay and to the distance between two relays. For instance, Furetière's (1690, 1636) dictionary defines a *poste* as “the space between two post houses. Each post is one and a half leagues or two leagues” (quoted in Bretagnolle and Verdier 2006, 66).

¹⁹ Because all editions of the *Liste des postes* have not been preserved, we rely on the editions from 1714, 1721, 1731, 1741, 1751,

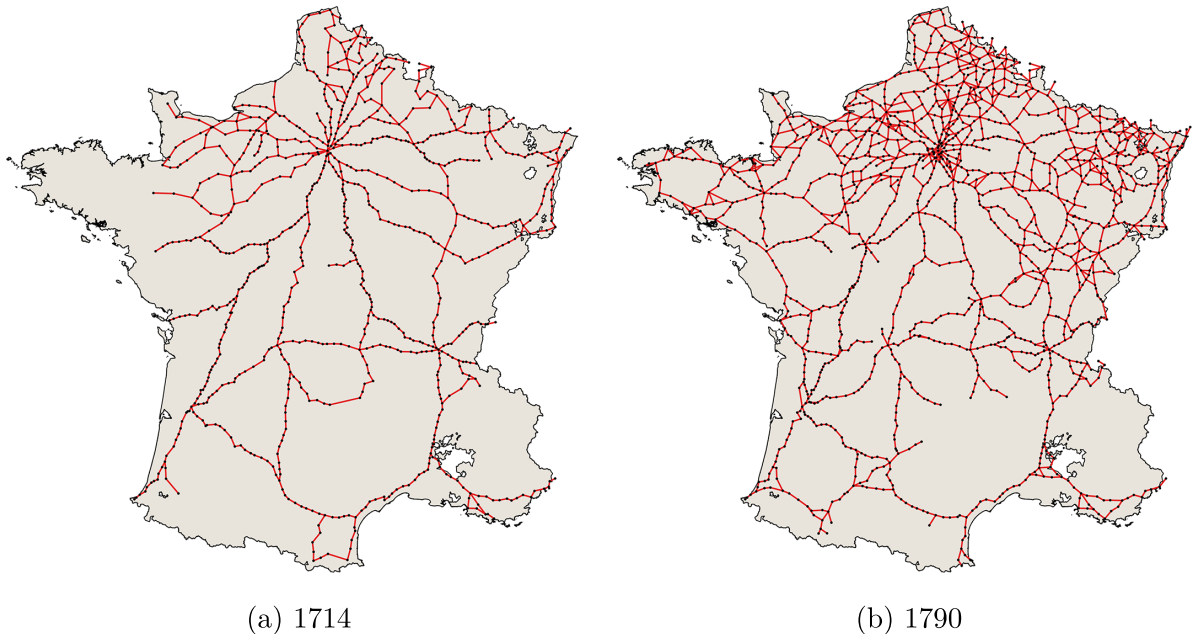
1760, 1770, 1780, and 1790. We do not use the 1708 edition—the first edition of the *Liste des postes*—because the year 1709, marked by the War of the Spanish Succession and a subsistence crisis, represents an outlier with 290 rebellions, as many as in 1789 (Aubert 2023). In addition, few horse-post relays were created between 1708 and 1714. See the Data Appendix section of the Supplementary Material for more details.

²⁰ Maps for intermediate decades are provided in Figures A3 and A4 in the Supplementary Material.

²¹ Figure A2 in the Supplementary Material reproduces a map of the horse-post relay network in 1632.

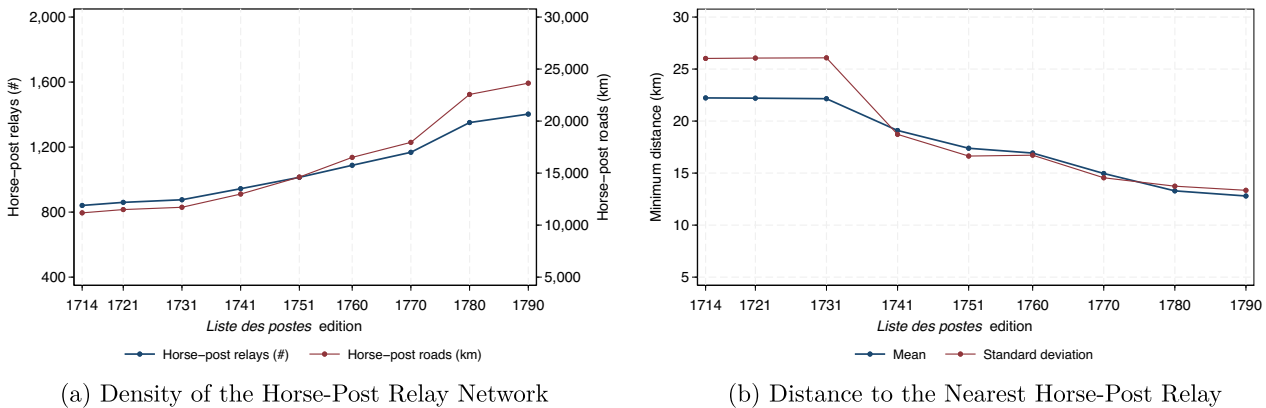
²² This strategy aligns with Pierskalla, De Juan, and Montgomery (2019, 712), who argue that early subnational state-building efforts often sought to maximize territorial coverage and fill in “blank spots” of state absence.

FIGURE 2. The Horse-Post Relay Network



Note: This figure displays the horse-post relay network based on the 1714 and 1790 editions of the *Liste des postes*. Horse-post relays correspond to black dots while horse-post roads, to red segments. The underlying shapefile of the kingdom of France as of 1789 is based on Gay, Gobbi, and Goñi's (2024a) jurisdictions database.

FIGURE 3. Evolution of the Horse-Post Relay Network



Note: Panel a displays the evolution of the number of horse-post relays and the length of postal roads in kilometers from 1714 to 1790. Panel b displays the mean and standard deviation of the distance from the 35 thousand parishes in our sample to the nearest horse-post relay in kilometers. Horse-post relay data are based on various editions of the *Liste des postes*.

mid-eighteenth century onward, the network expanded around regional centers, such as Bordeaux, Lyon, Dijon, Toulouse, and Rennes. By the end of the century, the horse post had evolved into a national network structured around the political center in the north and connected to secondary centers, creating the conditions for the emergence of an integrated city system (Bretagnolle and Franc 2017).

We quantify the expansion of the horse-post relay network over the eighteenth century in Panel a of Figure 3. The network grew steadily, essentially

doubling both in length and density over the period. Whereas in 1714, it comprised 841 relays spanning 11 thousand kilometers of postal roads, by 1790, it had expanded to 1,403 relays covering 24 thousand kilometers.²³ This expansion substantially improved local accessibility to the network, as the average distance

²³ These figures count parishes with at least one relay. Some urban parishes (e.g., Paris) had multiple, so the total number of relays was slightly higher—845 in 1714 and 1,413 in 1790.

between a parish and the nearest relay decreased from 22 kilometers in 1714 to 13 kilometers in 1790, as shown in Panel b of Figure 3. These gains were spread across all parishes, as the standard deviation of this distance also fell from 26 to 13 kilometers.²⁴ More broadly, the proportion of parishes located near a horse-post relay rose substantially, from 18% in 1714 to 31% in 1790.²⁵

However, the development of the horse-post relay network was not entirely linear as it was periodically reconfigured at the local level (Bretagnolle and Verdier 2007; Verdier 2009). While net gains amounted to 562 relays over the eighteenth century, this reflected the creation of 1,001 relays and the retirement of 439 of them.²⁶ In particular, substantial reconfigurations occurred in the interior of the country in the 1740s and 1750s—between Bordeaux and Limoges, Bordeaux and Poitiers, Poitiers and Tours, and Lyon and Clermont-Ferrand. These reconfigurations were generally driven by local practical considerations, such as the retirement of a postmaster or the development of nearby roads in better condition (Arbellot 1980, 106).²⁷ A comparison of the spatial distribution of newly established and retired relays by decade reveals that retired relays were, on average, 11 kilometers away from the nearest new relay.

While the initial configuration and expansion of the horse-post relay network were largely shaped by national strategic considerations, its local development over the eighteenth century was driven by practical constraints. Beyond the quality of the road network, the specific placement of relays was primarily determined by the need to maintain regular 10–15 kilometer intervals—reflecting the physiological limits of horses (Minetti 2003). As a result, many relays were located in small towns and villages situated along routes between administrative centers, often on riverbanks or on foot-hills.²⁸ Accordingly, about one-third of relays were located in cities with more than two thousand inhabitants—typically administrative centers—while nearly half were in towns with fewer than a thousand inhabitants, and 19% in villages with fewer than five hundred inhabitants.²⁹ Horse-post relays established after 1714 were even more likely to be located in smaller towns and villages.

²⁴ An analysis based on a Shimbel (1953) index confirms this trend (Bretagnolle, Giraud, and Verdier 2010).

²⁵ See Figure A5 in the Supplementary Material.

²⁶ See Figure A6 in the Supplementary Material.

²⁷ For instance, the relay of Vivier in Brittany—established in 1749—was transferred 7 kilometers away to Dol in 1760 to enable access to Pontorson via a better-maintained road. Likewise, a series of relay transfers between Rennes and Lorient in July 1763 shortened this route (Jamaux-Gohier 2001, 24). See Panel a of Figure A7 in the Supplementary Material.

²⁸ For instance, in Touraine, many horse-post relays were located in villages nearby river crossings in addition to administrative centers, such as Tours or Amboise. See Panel b of Figure A7 in the Supplementary Material.

²⁹ See Figure A8 in the Supplementary Material. Parishes that were administrative centers had a median population of 3.2 thousand inhabitants in 1793 and a mean of 6.5 thousand.

TABLE 1. Summary Statistics across Parish-Decades

	Mean	S.d.	Min.	Max.
A. The horse-post network				
Horse-post relay present in parish	0.029	0.167	0	1
Horse-post relay present in canton	0.224	0.417	0	1
Distance to nearest horse-post relay (km)	18.5	20.7	0	256
B. Rebellions				
All rebellions	0.020	0.225	0	34
Weak rebellions (4–10)	0.007	0.107	0	16
Moderate rebellions (11–50)	0.009	0.121	0	13
Strong rebellions (50+)	0.004	0.075	0	6
Rebellions against state authorities	0.012	0.161	0	23
Rebellions against non-state authorities	0.002	0.056	0	5
Rebellions against taxation	0.008	0.138	0	19
Rebellions against military	0.002	0.050	0	3
Rebellions against judiciary	0.002	0.057	0	10

Note: This table reports summary statistics for the main variables used in the analysis, excluding the city of Paris. The unit of observation is a parish-decade from the 1710s to the 1780s. There are 282,288 parish-decade observations in the dataset. *State authorities* include the fiscal, military, and judiciary authorities. *Non-state authorities* include the nobility, seigneurial and municipal authorities, and the Church.

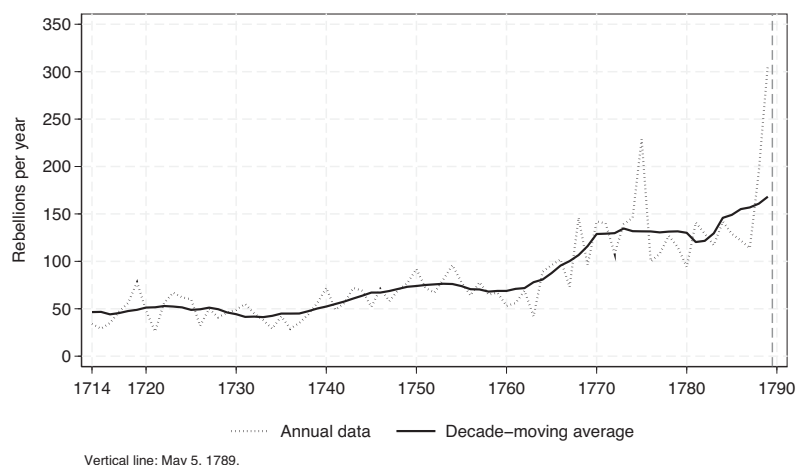
Summary statistics on the horse-post relay network at the parish-decade level are reported in Panel A of Table 1.³⁰

Rebellions

We collect data on eighteenth-century rebellions from the Jean Nicolas survey (Nicolas 1985; 2002).³¹ Compiled from extensive archival research conducted from the early 1980s to the late 1990s, this survey documents 8,516 rebellions that occurred between 1661 and 1789 (Gay 2025, 7–11). It focuses on popular uprisings aimed at challenging authorities, protesting subsistence crises,

³⁰ This table excludes the city of Paris, which is an outlier in terms of rebellious activity. An analogous table that includes Paris is provided in Table A1 in the Supplementary Material.

³¹ Our dataset on eighteenth-century rebellions in France uses one of the same sources as the HiSCoD database (Chambru and Maneuvrier-Hervieu 2024)—that is, Jean Nicolas's (2002) survey. The Data Appendix section of the Supplementary Material details how our database, in containing *all* of the variables from Nicolas's (2002) survey, extends beyond the HiSCoD database.

FIGURE 4. Annual Number of Rebellions

Note: This figure displays the annual number of rebellions from 1714 to 1789 based on Nicolas's (2002) original survey along with a decade-moving average.

or resisting state and local policies and actions—it excludes brigandage and strictly criminal activities. The survey sets an inclusive threshold, recording any uprising that lasted only a few hours or involved as few as four individuals. These criteria thus cast a wide net, making the Jean Nicolas survey an unparalleled thermometer of the rebellious fever that swept through France in the run-up to the Revolution.

Figure 4 depicts trends in rebellious activity between 1714 and 1789, our period of analysis. Rebellions began to intensify in the 1740s before accelerating sharply in the 1760s, with more than a hundred events per year from that point onward. This escalation coincided with the growing presence of the state across the territory and rising levels of taxation (Burguière and Revel 1989).³² Rebellious activities further intensified until the outbreak of the Revolution. Their territorial spread—over about three thousand locations—is shown in Figure 5.³³

We digitized all forty thousand original coding sheets of Nicolas's (2002) survey to construct a comprehensive database of the 6,000 rebellions that took place in 3,172 parishes within the boundaries of the kingdom of France between 1714 and 1789.³⁴ In addition to recording the date and location of each event, we also collected information on its motives, size, and the characteristics of its participants.³⁵ To measure

rebellious activity, we create a variable that counts the number of rebellions per parish and decade. We report summary statistics in Panel B of Table 1.³⁶

Nearly half of all rebellions (46%) that occurred from 1714 to 1789 were of moderate size, involving between 11 and 50 participants. Another 21% were relatively large, with more than 50 participants. State authorities were targeted in 60% of rebellions, while 11% were directed against non-state actors—the nobility, seigneurial and municipal authorities, and the Church. Among rebellions against the state, most were aimed at taxation (66%), followed by the military (19%) and the judiciary (23%).³⁷ Finally, 69% of rebellions in our sample include information on the social categories of their participants. Among these, 39% involved laborers; 32%, craftsmen; 23%, notables; and 22%, socially marginalized individuals.³⁸ The rebellions in our dataset thus span a broad range of the social strata of French society.

While Nicolas's (2002) survey remains the most comprehensive source of rebellious activity in France prior to the Revolution, its use requires careful attention to potential gaps and reporting biases common to historical data (Gay 2025, 29–46). These include not only the archival recording of events, but also the possibility that the establishment of a horse-post relay may have

³² The administrative divisions used for taxation—typically the *élections*—were independent of the horse post (Touzery 2024). They also operated at a much coarser scale than the units we use for identification in our analysis, that is, parishes within the same canton.

³³ Maps for individual decades are provided in Figures A3 and A4 in the Supplementary Material.

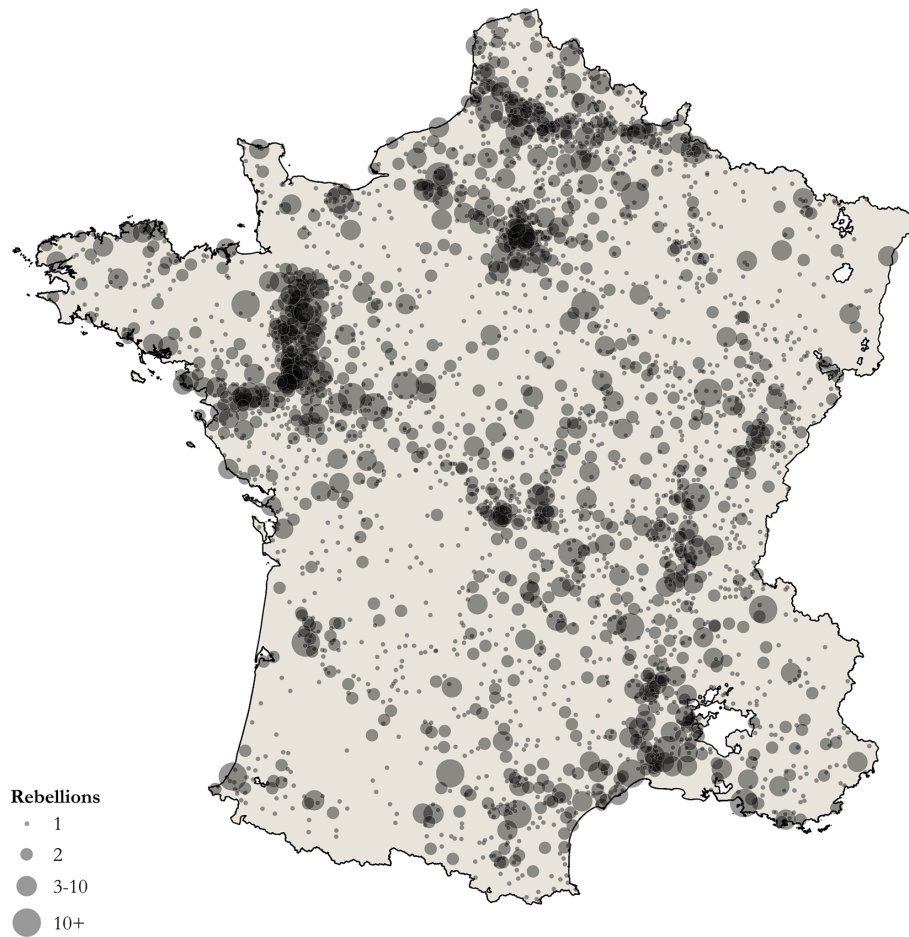
³⁴ We exclude rebellions that occurred after April 1789 and the onset of the French Revolution.

³⁵ We include both primary and secondary rebellion motives as recorded in Nicolas's (2002) survey. Of the 6,000 rebellions in our dataset, 1,398 (23%) have secondary-motive information.

³⁶ Because Paris experienced the highest number of rebellions throughout the eighteenth century—except in the 1720s—Table 1 excludes this city. An analogous table, including Paris, is provided in Table A1 in the Supplementary Material.

³⁷ Table A2 in the Supplementary Material provides a detailed breakdown of rebellion types along with the number of cases for each type.

³⁸ See Table A3 in the Supplementary Material. These categories are not mutually exclusive, as individuals from different social groups could have participated in the same rebellion. Among the 4,154 rebellions for which social categories are available (69% of all rebellions), 37% involved individuals from more than one social category.

FIGURE 5. Spatial Distribution of Rebellions (1714–89)

Note: This figure displays the spatial distribution of rebellions from 1714 to 1789 based on Nicolas's (2002) original survey. Each point corresponds to a parish that experienced at least one rebellion over the period—6,000 rebellions across 3,172 parishes in total. The underlying shapefile of the kingdom of France as of 1789 is based on Gay, Gobbi, and Goñi's (2024a) jurisdictions database.

increased the capacity of state or local authorities to record rebellions. We address these concerns below by assessing the robustness of our results to these potential sources of bias, showing that they are unlikely to account for our findings.

Administrative and Territorial Divisions

Central for our empirical strategy, parishes were embedded within a hierarchical system of administrative and territorial divisions. The largest of these were *généralités*. Created in the fourteenth century, these broad territorial divisions served as the basis for the administration of the territory under the authority of intendants (Gorry 2014). On the eve of the Revolution, there were 32 *généralités*.³⁹ *Subdélégations* were territorial divisions established in the mid-seventeenth century as subdivisions of *généralités* to extend the territorial reach of the royal authority at

the local level. The prerogatives of *subdélégués*—the royal agents who administered *subdélégations* under the authority of the intendant of their *généralité*—increased during the eighteenth century to encompass the fiscal, military, economic, social, and sanitary administration of their territory (Goubert, Arbellot, and Laclau 1989). By 1789, there were nearly 700 *subdélégations*. A more localized territorial unit was the *canton*, which held various electoral and judicial functions (Ozouf-Marignier, Graziani, and Bonin 1989). Alongside parishes, cantons represented the most granular level of administration, totaling over 4,500.⁴⁰ Although they were only formally created in 1790 along with *départements* and *districts*, their territorial scope closely reflected pre-existing patterns of local interaction and daily exchange between parishes (Tanguy, Le Bihan, and Lagadec 2009). At the median, each *généralité* had 988 parishes, 782 thousand inhabitants (as of 1793), and covered a radius of 70 kilometers; each *subdélégation* had

³⁹ The precise number of *généralités* varies depending on whether we account for *intendances* as well as territories not strictly in the kingdom of France.

⁴⁰ See Figure A9 in the Supplementary Material.

39 parishes, 31 thousand inhabitants, and a radius of 15 kilometers; and each canton had 7 parishes, 5 thousand inhabitants, and a radius of 6 kilometers.

Other Data

We supplement our parish-level data with additional sources, although data for this time period remain relatively sparse. Our research design explicitly accounts for this constraint. In particular, we construct a parish-level measure of decade-average wheat prices based on data from Ridolfi (2019), so as to control for shifting local economic conditions and for potential income shocks associated with the local development of the horse-post relay network—subsistence crises and broader changes to the price of staple foods were common triggers of collective action in eighteenth-century France (Bouton 1993; Chambru 2019; Markoff 1985).⁴¹ In addition, we measure the size of each parish by its population in the 1793 census—the first census conducted in France. We also assess the connectivity of parishes to the road network based on Cassini's maps published in 1756–90. Furthermore, we determine whether a parish served as the seat of an administrative division, and we collect information on the type of gabelle taxation regime to which it was subject. Finally, we construct a historical GIS of both the letter-post system and maréchaussée brigades, which we describe in more detail below.

RESEARCH DESIGN

The substantial variation in rebellious activity and the expansion of the horse-post relay network over the eighteenth century offer a unique opportunity to examine how state-building efforts affected unrest in the run-up to the French Revolution. Our primary estimand of interest is the causal effect of the establishment of a horse-post relay in a parish on subsequent rebellious activity in that parish over time. While rebellions ultimately stem from social and economic changes, the introduction of a relay could have triggered these changes and their associated grievances.

We employ a staggered difference-in-differences design that compares changes in rebellion in parishes that received (or lost) a horse-post relay to those that would later receive (or lose) one. This strategy accounts for fixed parish-level characteristics that might influence both the likelihood of receiving a relay and the propensity for rebellion, such as a prior history of political activism, ease of collective action, the structure of civil society, or the presence of natural resources. It relies on the identifying assumption that, absent the introduction of a relay, parishes that received a new

relay and those that did not would have followed parallel trends in rebellion. While this assumption is fundamentally untestable, we assess its plausibility through event-study regressions below.

Our treatment is staggered, binary, and non-absorbing, as parishes could both enter and exit treatment over time. Within our sample, 33,602 parishes (95.2%) were never treated—they never hosted a horse-post relay between 1714 and 1789. Another 586 (1.7%) were always treated—they maintained a relay throughout the period. The remaining 1,099 (3.1%) were “ever treated” parishes that experienced at least one change in treatment status over time. Of these, 203 (18.5%) began with a relay in 1714 but lost it at some point, while 896 (81.5%) entered treatment at least once. Among the latter, 387 (43.1%) exited treatment at least once.⁴²

To account for the specific features of this empirical design, we follow de Chaisemartin and D'Haultfœuille (2024), which provide consistent and unbiased estimators for staggered difference-in-differences models in which treatment effects may be heterogeneous as well as dynamic over time.⁴³ Specifically, de Chaisemartin and D'Haultfœuille's (2024) approach estimates period-specific average treatment effects δ_ℓ of having been treated rather than untreated for ℓ periods by comparing switchers to not-yet switchers with the same period-one treatment.⁴⁴ This setup also ensures that the coefficients are not “contaminated” by assuming continual exposure when parishes could exit treatment. We implement this approach at the parish-decade level by estimating the following regression equation:

$$\begin{aligned} \text{rebellions}_{pda} = & \alpha_p + \lambda_d + \sum_{\substack{\ell=-t \\ t \neq 0}}^{+T} \delta_\ell \times \mathbb{1}[\text{horse_post}_{pd} = 1] \\ & \times \mathbb{1}[d = \ell] + \text{prices}_{pd} + \gamma_{da} + \varepsilon_{pda}, \end{aligned} \quad (1)$$

where rebellions_{pda} denotes the number of rebellions that occurred in decade d and parish p that is part of administrative division a . Index d ranges from 1 to 8, corresponding to decades 1710 to 1780. The indicator variable horse_post_{pd} equals one if parish p hosted a horse-post relay in decade d . α_p denotes parish fixed effects; λ_d , decade fixed effects; prices_{pd} , average wheat prices in parish p and decade d ; and γ_{da} , decade-by-administrative division fixed effects, where we

⁴¹ Wheat was the most important staple, due to both its dominant role in bread consumption and its widespread cultivation (Ridolfi 2019). Because food is a normal good, local positive income shocks should have resulted in a substantial increase in the local demand for wheat and thus its price, given the inelasticity of the local supply of wheat.

⁴² Figure A10 in the Supplementary Material documents the 52 unique treatment histories among parishes that switched treatment status at least once during our period of analysis.

⁴³ Unlike other estimators, such as those proposed by Callaway and Sant'Anna (2021), Sun and Abraham (2021), or Borusyak, Jaravel, and Spiess (2024), this approach accommodates staggered binary treatments that are non-absorbing.

⁴⁴ In addition to assuming no anticipation of treatment and parallel trends, this approach requires that past treatments before the starting period do not affect outcomes during the analysis period (de Chaisemartin and D'Haultfœuille 2024, 9–12). We discuss this assumption in the Robustness section below.

alternatively consider subdélégations and cantons. This last term enables the coefficients δ_ϵ to be identified through the comparison of parishes within the same administrative division, thus accounting for administrative division-specific time trends. We cluster standard errors at the parish level.

An important practical consideration is the timing of the assignment of horse-post relays and rebellions. Our baseline approach links rebellions in a given decade to the edition of the *Liste des postes* published at the beginning of that decade. For instance, for the 1720 decade, we use the 1721 edition of the *Liste des postes* together with rebellions that occurred between 1721 and 1730.⁴⁵ A second consideration is the number of leads (t) and lags (T) to include in the estimation. While de Chaisemartin and D'Haultfoeuille's (2024) methodology allows for up to three leads and seven lags in our setting, the preferred number of leads and lags should ensure that the estimation is based on a sufficiently comparable sample of switchers. While we report estimates for the full range, our preferred specification follows de Chaisemartin and D'Haultfoeuille's (2024, 39–40) rule of thumb, which recommends including only those leads and lags for which at least half of ever-treated parishes contribute to the estimation. This corresponds to including two leads and three lags in our setting.

IDENTIFICATION

A causal interpretation of estimates resulting from Equation 1 requires that the development of the horse-post relay network be exogenous with respect to rebellion dynamics at the parish level. As discussed above, the initial spatial configuration of the network reflected long-standing strategic political and military concerns (Panel a of Figure 2). In contrast, its development over the eighteenth century was primarily driven by the state's efforts to expand and centralize its authority across the interior of the kingdom, resulting in the densification of the network around major regional nodes (Panel b of Figure 2).⁴⁶ However, the opening of new horse-post relays in growing urban centers may have been endogenous to population dynamics—a potential determinant of rebellious activity. To address this concern, our preferred specification exploits local variation through decade-by-administrative division fixed effects (γ_{da}). Indeed, while the establishment of a relay in an urban center may be endogenous, the precise location of a relay *between* regional nodes was ultimately

constrained by the need to establish rest stops every 10–15 kilometers due to the physiological limits of horses (Minetti 2003).

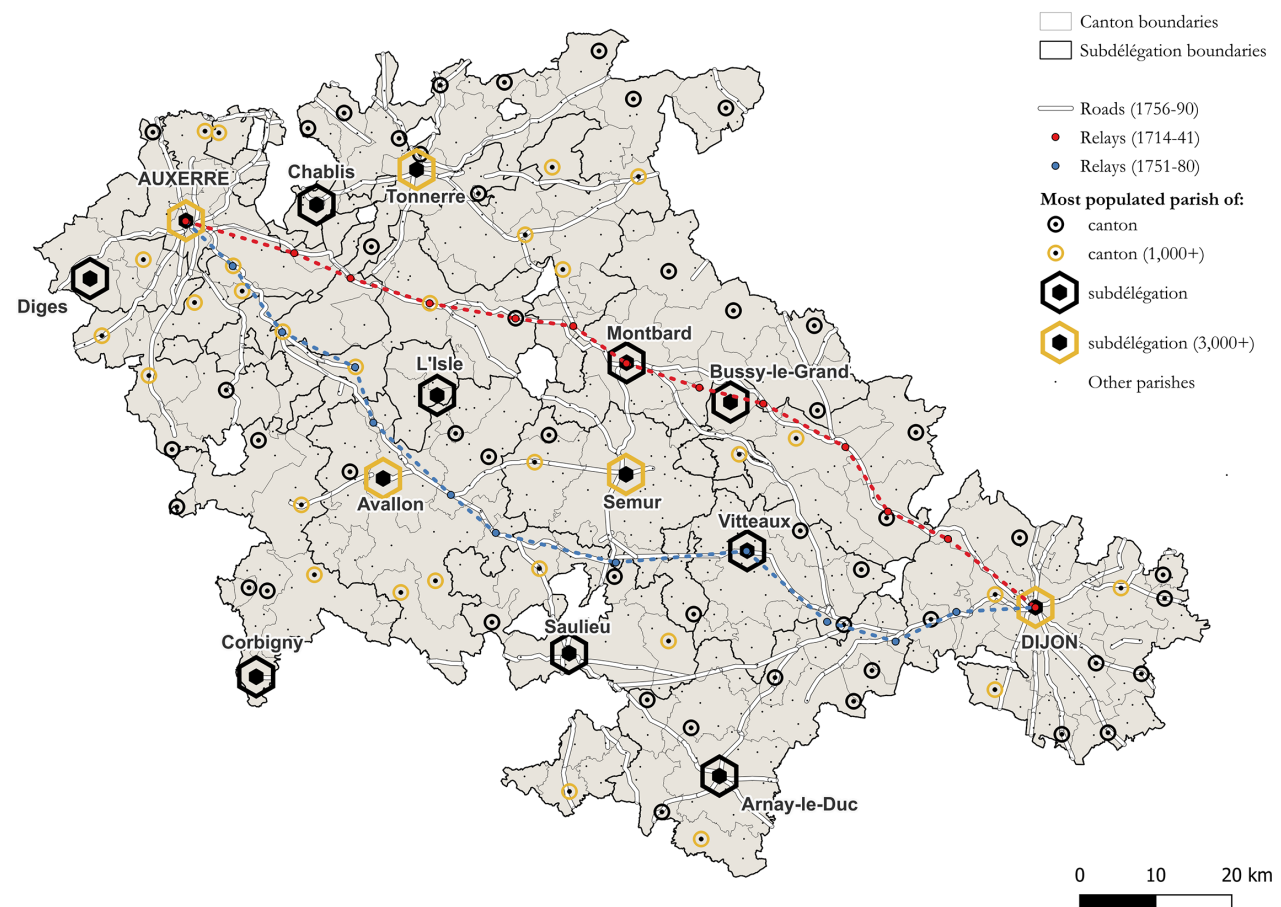
We illustrate this argument in Figure 6, which maps the postal route between Auxerre and Dijon—two regional centers and subdélégation seats in Burgundy that hosted horse-post relays throughout the eighteenth century. Since the early seventeenth century, this route had run northbound along pre-existing roads and waterways via the subdélégation seat of Montbard and eleven additional relays, spaced 11 kilometers apart on average (Nougaret 1960, 30–2). However, only two of these were located in the most populous parishes of their respective cantons; the remaining nine were located in villages with fewer than five hundred inhabitants. This relay configuration formed a relatively straight route, minimizing the travel distance between the two regional centers. And although the road passed right near the subdélégation seat of Bussy-le-Grand, no relay was ever established there. In 1742, this postal route was rerouted southward through the subdélégation seat of Vitteaux and ten other relays, following road improvements carried out in the early 1730s (Nougaret 1960, 32, 125–6; Tonnaire 2005, 98–9). As before, relays were spaced 11 kilometers apart on average, and only three of these relays were located in the most populous parishes of their respective cantons; the remaining seven were located in villages with fewer than five hundred inhabitants. This new configuration also formed a relatively straight route, minimizing travel distance between the two regional centers.

This example illustrates that while the placement of central nodes of the horse-post relay network was likely endogenous to regional population dynamics, its local configuration *between* these nodes was plausibly exogenous, as it was designed to minimize travel time along pre-existing roads, subject to the constraint that relays be spaced 10–15 kilometer apart. We therefore exploit local variation in the network through decade-by-administrative division fixed effects. Because this assumption is more credible the finer the scale of analysis, our preferred specification relies on within-canton variation—the average canton contained only two towns with more than a thousand inhabitants in 1793, compared to ten per subdélégation.

This strategy restricts the sampling variation available for identification, as it requires that a canton contains at least two parishes with the same initial treatment status, and that one of them changes treatment status over time. This restriction implies four non-mutually exclusive exclusions from the sample: (i) 201 cantons composed of a single parish; (ii) 3,694 cantons where no parish ever switched treatment—either because none ever hosted a horse-post relay or those that did never lost it; (iii) 105 cantons where only one parish switched treatment and no other parish had the same initial treatment status—typically cantons where a single parish began with a relay, then lost it, while no others ever hosted one; and (iv) 14 cantons where all parishes that ever had a relay followed identical treatment paths. Together, these restrictions remove 3,847 cantons (84% of all cantons) when decade-by-canton fixed effects are introduced.

⁴⁵ Given the availability of *Liste des postes* editions, some parish-decade observations are slightly shorter than others. For instance, for the 1710 decade, we use the 1714 edition of the *Liste des postes* together with rebellions that occurred between 1714 and 1720, since the next available edition is from 1721.

⁴⁶ We further show in Table A4 in the Supplementary Material that more populous parishes, those closer to a road, and those serving as administrative centers, were more likely to ever host a horse-post relay. However, these observable characteristics explain no more than 15% of the variation in relay assignment.

FIGURE 6. The Horse-Post Road from Auxerre to Dijon

Note: This figure displays the postal road from Auxerre to Dijon in Burgundy between 1714 and 1741 (northerly red dashed lines) and between 1751 and 1789 (southerly blue dashed lines). The extent of the map corresponds to the subdélégations (thick boundaries) and cantons (thin boundaries) through which this postal road passed. It also displays the spatial distributions of parishes in the area (marking the most populous ones by canton and subdélégation) and of 1756–90 roads (white lines). Horse-post relay data are based on various editions of the *Liste des postes*, and 1793 census data, on Cristofoli et al. (2021). The shapefile of subdélégations is based on information in Cristofoli et al. (2021), that of cantons, on information in cassini.ehess.fr (Motte and Vouloir 2007; Motte et al. 2003), and that of roads, on Perret, Gribaudo, and Barthelemy (2015).

This leaves 732 cantons (out of 1,342 that ever hosted a relay) for identification.⁴⁷

MAIN RESULTS

Table 2 reports the main results on the impact of the development of the horse-post relay network on rebellion across eighteenth-century France. Specifically, we report the average total effect of decade-specific event-study estimates δ_t , which represents the sum of

⁴⁷ We illustrate this selection in Figure A9 in the Supplementary Material. Note that while the average parish population in 1793 was 811, those located in cantons used for identification had 769. However, a regression of the log parish population on an indicator variable that equals one if a parish was in such a canton yields a small and insignificant coefficient (estimate: -0.0274 ; standard error clustered at the canton level: 0.0205), suggesting little selection into the effective sample along this dimension.

cumulated period-specific treatments (de Chaisemartin and D'Haultfœuille 2024, 24–5). This table further provides the effective number of parishes-decade and parish observations used for identification, the number of parish-decades that switched treatment, as well as the outcome mean and residual standard deviation calculated on the effective sample.⁴⁸

Column 1 reports results from a model without controls estimated using the maximum number of leads and lags allowed by the structure of the data. The establishment of a horse-post relay in a parish was associated with a subsequent increase in rebellion. The average total effect implies an increase in the

⁴⁸ We follow Mummolo and Peterson (2018) and rely on residual standard deviation as a measure of sampling variation for interpreting the magnitude of regression estimates. The raw standard deviation in rebellions in our sample is about twice as large as the residual standard deviation after accounting for parish and decade fixed effects.

TABLE 2. Average Total Effect of New Horse-Post Relays on Rebellions

Outcome	Number of rebellions					
	1	2	3	4	5	6
Horse-post relay	0.057*** [0.011]	0.029*** [0.010]	0.029*** [0.010]	0.040*** [0.009]	0.029*** [0.004]	0.029*** [0.004]
Wheat prices	No	No	Yes	No	No	Yes
8 decades × 696 subdélégations FE	No	No	No	Yes	No	No
8 decades × 4,579 cantons FE	No	No	No	No	Yes	Yes
Parish-decades	282,296	282,296	282,296	183,904	54,672	54,672
Parishes	35,287	35,287	35,287	22,988	6,834	6,834
Switchers (overall)	3,522	2,561	2,561	2,352	1,929	1,929
Switchers (at least once)	1,099	1,099	1,099	1,018	849	849
Outcome mean	0.021	0.021	0.021	0.017	0.017	0.017
Outcome residual s.d.	0.225	0.225	0.225	0.148	0.122	0.122
Periods (− <i>t</i> ; + <i>T</i>)	(−3;+7)	(−2;+3)	(−2;+3)	(−2;+3)	(−2;+3)	(−2;+3)

Note: This table reports the average total effect of non-normalized event-study effects δ_t estimated through Equation 1. All regressions include parish and decade fixed effects. *Parish-decades* and *Parishes* denote the number of observations effectively used in the estimation. *Switchers (overall)* denotes the number of parish-decade observations that switched treatment over the period, while *Switchers (at least once)* denotes the number of parishes that switched at least once. *Outcome residual s.d.* denotes the outcome standard deviation once decade, parish fixed effects, and other controls are accounted for. Summary statistics are calculated relative to the effective sample used in the estimation. *Periods (−*t*; +*T*)* denotes the number of event-study effects and placebos estimated. Estimates are calculated using de Chaisemartin et al.'s (2023) `did_multiplegt_dyn` Stata command. Standard errors are clustered at the parish level and reported in brackets. *** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.

number of rebellions by 0.057. This represents a little more than twice the average incidence of rebellion and one-quarter of its residual standard deviation. Column 2 restricts the number of periods used in the estimation to two leads and three lags. The coefficient decreases to 0.029 but remains significant at the 1% level and of sizable magnitude, corresponding to the average incidence of rebellion and one-tenth of its residual standard deviation. Column 3 introduces wheat prices as a time-varying control to assess whether local income shocks following the establishment of a relay might account for the results. The estimate remains unchanged, suggesting that the introduction of a relay was unrelated to changing local economic conditions that could have triggered rebellion. This finding aligns with evidence from Tonnaire (2005, 44–62), who concludes on the basis of *taille* and *capitation* information that parishes along the postal road between Auxerre and Dijon that received a relay in the 1740s did not become wealthier in subsequent decades. If anything, her evidence points toward distributional effects, consistent with the mechanisms we propose below. Columns 4 and 5 introduce decade-by-administrative division fixed effects to compare parishes over time within the same subdélégation or canton.⁴⁹

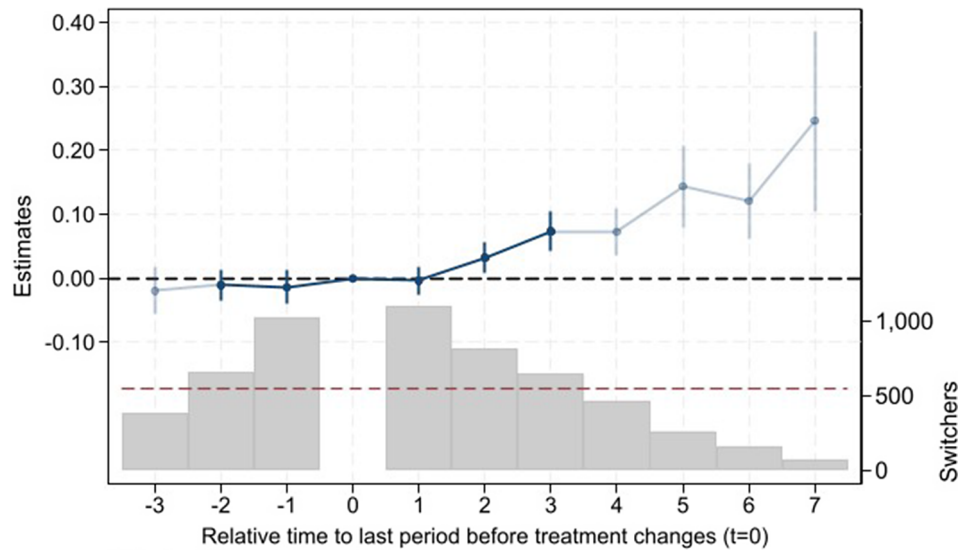
⁴⁹ The effective sample reduces to 22,988 parishes when including decade-by-subdélégation fixed effects, as only 370 of the 696 subdélégations contain at least two parishes with the same initial treatment status, with one switching treatment over the period. It further reduces to 6,834 parishes when we include decade-by-canton fixed effects, as only 732 of the 4,579 cantons meet this criterion. To ensure

Results are consistent with the baseline at 0.029, suggesting limited heterogeneity in treatment effects across administrative divisions. Finally, column 6 further controls for wheat prices. The resulting coefficient of 0.029 is again essentially identical to the baseline in column 2. However, its relative magnitude is about twice as large due to the reduced sampling variation after accounting for the additional controls used in this specification.

Figure 7 presents an event-study plot that encapsulates the estimates from columns 1 and 2. This figure serves three purposes: first, to assess the plausibility of the parallel-trends assumptions; second, to examine the temporal dynamics of the treatment effect; and third, to display event-study coefficients without restricting the number of leads and lags. Placebo lead estimates show no evidence of differential pre-trends between treated and untreated parishes, supporting the credibility of the parallel-trends assumption. In contrast, post-treatment estimates reveal substantial dynamic treatment effects: rebellions increase in the second decade after a horse-post relay is established and continue to grow thereafter. Additional analyses using sliding decadal intervals in rebellions indicate that the treatment effect begins to materialize generally within the first decade and grows thereafter.⁵⁰ Finally, the figure indicates that the third

that this sample selection does not drive the results, we re-estimated the specification of column 3 without interacted fixed effects on these restricted samples, finding broadly similar results.

⁵⁰ See Table A6 in the Supplementary Material. Our results are also robust to using alternative timing specifications such as excluding

FIGURE 7. Event-Study Effects of New Horse-Post Relays on Rebellions

Note: The top half of this figure reports event-study effects δ_c estimated through Equation 1 along with 95% confidence intervals. The specifications correspond to Columns 1 and 2 in Table 3. Dark blue estimates correspond to restricting the estimation to two leads and three lags, while light blue estimates use the full three leads and seven lags in the estimation. The bottom half of this figure reports the number of switchers used for the identification of each event-study effect. The red dashed line denotes half of first-period switchers. The full set of results is available in Table A5 in the Supplementary Material.

placebo lead and the fifth, sixth, seventh period estimates rely on a relatively small number of switchers, supporting our choice to emphasize estimates using two leads and three lags. An event-study plot relative to the specification in column 6 of Table 2 leads to similar conclusions.⁵¹

Robustness

Table 3 assesses the robustness of our main results. These tests build on the full specification reported in column 6 of Table 2—reported here in column 1—which includes controls for wheat prices and decade-by-canton fixed effects.

Periodization

To assess the credibility of the “initial conditions” assumption (de Chaisemartin and D’Haultfœuille 2024, 9–12), column 2 excludes the 1710 decade from the sample. The resulting estimate of 0.028 remains close to the baseline and is significant at the 1% level, suggesting that this assumption is reasonable.⁵²

rebellions that occurred during the year the *Liste des postes* was published, with an estimate of 0.027 (standard error: 0.004), or assigning rebellions to the following decade, with an estimate of 0.051 (standard error: 0.005).

⁵¹ See Figure A11 in the Supplementary Material.

⁵² We further report estimates when running Equation 1 on shorter panels in Table A7 in the Supplementary Material. The estimates remain stable, suggesting limited decade-specific heterogeneity.

Sample Composition

Next, we hold constant the sample of switchers used for identification to assess whether changes in the effective sample drive our results. The estimate in Column 3 of 0.035 suggests that they do not. Column 4 further restricts the sample to parishes that switch *into* treatment—that is, parishes that receive a new horse-post relay—as opposed to using both switchers in and switchers out—that is, parishes that also lose a relay. The estimate declines slightly to 0.022 but remains significant at the 1% level and consistent with the baseline.

Heterogeneity

In column 5, we estimate Equation 1 with 1793 population weights to gauge the heterogeneity of treatment effects across parish sizes. We find an effect of 0.075, suggesting that larger parishes may respond more strongly to the introduction of a new relay. However, given the higher weighted outcome mean and residual standard deviation, the magnitude remains comparable to the baseline. Similarly, in column 6, we further include decade-by-rural status fixed effects, where rural status is defined as an indicator variable that equals one if the parish had fewer than two thousand inhabitants in 1793—the threshold used to classify rural communes in nineteenth-century censuses (Le Mée 1972). We find an estimate of 0.019, again comparable to the baseline. We further investigate the nature of heterogeneous treatment effects along various characteristics in Appendix B of the Supplementary Material, showing that baseline differences in population, administrative status, gabelle taxation regime, historical opposition to

TABLE 3. Robustness										
Outcome	Number of rebellions									1[rebellion]
	1	2	3	4	5	6	7	8	9	10
Horse-post relay	0.029*** [0.004]	0.028*** [0.004]	0.035*** [0.004]	0.022*** [0.004]	0.075*** [0.008]	0.019*** [0.003]	0.017*** [0.003]	0.037*** [0.005]	−0.005 [0.006]	0.023*** [0.003]
Specification	Baseline	No 1710	Same switchers	Switchers in	Pop. weights	Rural- decade FE	Moderate/ strong	New relays	Reconf. relays	Indicator
Parish-decades	54,672	45,934	54,672	54,672	54,672	43,088	54,672	42,984	16,544	54,672
Parishes	6,834	6,562	6,834	6,834	6,834	5,386	6,834	5,373	2,068	6,834
Switchers (overall)	1,929	1,826	1,404	1,840	1,896	1,565	1,929	1,424	573	1,929
Switchers (at least once)	849	817	468	809	835	689	849	643	234	849
Outcome mean	0.017	0.018	0.017	0.017	0.073	0.014	0.011	0.017	0.019	0.014
Outcome residual s.d.	0.122	0.130	0.122	0.122	0.235	0.108	0.099	0.122	0.153	0.098
<i>Note:</i> This table reports the average total effect of non-normalized event-study effects δ_{ℓ} estimated through Equation 1. All regressions include parish and decade fixed effects, as well as wheat prices and decade-by-canton fixed effects. The estimation is restricted to two leads and three lags. Column 2 removes parish-decades of the 1710s. Column 6 includes an additional set of decade-by-rural status fixed effects. Column 8 restricts horse-post relays to those that are at least 15 kilometers from a retired relay in the same decade, and column 9, to those that are within 15 kilometers, where <i>Reconf.</i> denotes reconfigured relays. The outcome is the number of rebellions in columns 1–9, and an indicator variable for the occurrence of at least one rebellion in column 10. <i>Parish-decades</i> and <i>Parishes</i> denote the number of observations effectively used in the estimation. <i>Switchers (overall)</i> denotes the number of parish-decade observations that switched treatment over the period, while <i>Switchers (at least once)</i> denotes the number of parishes that switched at least once. <i>Outcome residual s.d.</i> denotes the outcome standard deviation once decade, parish, and canton fixed effects, as well as wheat prices are accounted for. Estimates are calculated using de Chaisemartin et al.'s (2023) <code>did_multiplegt_dyn</code> Stata command. Standard errors are clustered at the parish level and reported in brackets. *** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.										

the central state, and mode of ruling do not account for our results. We also show that the treatment effect is generally homogeneous across regions and that the results are also robust to excluding one region at a time.⁵³

Rebellion Intensity

In column 7, we show that the effect we identify is not driven by small rebellions, as results are similar when focusing on moderate to strong rebellions.⁵⁴

Type of Horse-Post Relays

Columns 8 and 9 explore the nonlinear expansion dynamics of the horse-post relay network by contrasting entirely new relays with those arising from local reconfigurations.⁵⁵ We find that new relays are driving the results, while those stemming from local reconfigurations have no discernible impact on rebellion—likely because they did not meaningfully alter the local presence of the state.

Rebellions Indicator

Column 10 confirms that the results are robust when the outcome is instead specified as an indicator variable that equals one if there was at least one rebellion in a parish-decade.

Archival Reporting Biases

In addition to the robustness tests reported in Table 3, we assess whether our results could be driven by archival gaps or recording biases Nicolas's (2002) rebellion data. For instance, information for some départements is mostly drawn from national archives, which are generally less comprehensive than departmental archives. This could introduce biases in the representativeness of rebellious events across départements. Reassuringly, our empirical strategy mitigates this potential source of bias by leveraging variation in rebellious activity within the reach of departmental archives through the use of decade-by-canton fixed effects.

⁵³ See Tables A8 and A9 in the Supplementary Material, where we divide the territory into eight regions (see Figure A12 in the Supplementary Material). Only one region—Normandy—exhibits a negative coefficient, driven by four outlier subdélégations (out of 58 in Normandy). Removing these subdélégations yields an estimate of 0.017 (standard error: 0.005) for Normandy (see Table A10 in the Supplementary Material). Excluding the city of Paris from the sample—an outlier in terms of rebellious activity—yields an estimate of 0.029 (standard error: 0.004).

⁵⁴ Moderate rebellions are defined as involving 10–50 participants, and strong rebellions, more than 50. We report the average total treatment effects by rebellion intensity in Table A11 in the Supplementary Material. We also provide additional information on the measurement and reliability of rebellion intensity in the Data Appendix section of the Supplementary Material.

⁵⁵ A typical example of a local reconfiguration is shown in Panel a of Figure A7 in the Supplementary Material. We define entirely new relays as those located at least 15 kilometers away from a retired relay in the same decade.

But other forms of selection bias could affect our results. One concern is that the expansion of the horse-post relay network may have enhanced the state's ability to record rebellions. However, monitoring rebellions was not part of the horse post's surveillance function, which was primarily to control individuals passing through relays (Marchand 2006, 210–4). Moreover, spatial patterns presented below show no evidence of over-recording in areas surrounding relays beyond host parishes themselves. Furthermore, archives of the horse post are not among the sources of the Jean Nicolas survey, as a major flood in 1910 and destruction during the Second World War damaged much of the archives of Ancien Régime postal institutions (Krakovitch et al. 2000, 2). Surviving records were deposited in the National Archives in 1998, and were therefore largely inaccessible at the time the survey was conducted.⁵⁶ Instead, the main actors monitoring and punishing rebellious activity were the *maréchaussée* and the judiciary, much as the police and the courts today. The main sources of information on rebellion thus come from the archives of these institutions, which are well preserved and were extensively surveyed by Jean Nicolas and his collaborators (Gay 2025, 27–8).

Alternatively, the introduction of a horse-post relay may have increased the recording of rebellions by other local actors. To assess this possibility, we reproduce our analysis using only rebellions recorded by state-level authorities and preserved in national archives, finding that the results remain robust.⁵⁷ We further sequentially exclude rebellions recorded by the most frequent reporting institutions: the central government, the central financial administration, the *maréchaussée*, local courts and justices, and provincial administrations. Results are again consistent with the baseline, suggesting little selection bias arising from Nicolas's (2002) archival sources.

Additional Robustness Tests

In addition to the tests discussed above, we conduct further analyses to address the high incidence of zeroes in the outcome variable and to probe robustness to clustering standard errors at higher levels of aggregation. We also examine the potential for post-treatment bias arising from the use of canton fixed effects, given that canton boundaries were established after the Revolution (Homola, Pereira, and Tavits 2024). To that end, we re-estimate our models using fixed effects

⁵⁶ Postal administrations and their agents are also absent from Nicolas's (2002) typology. This omission reflects the thematic focus of the typology, which emphasized actors and institutions highlighted by the historiography when Jean Nicolas developed it in the early 1980s (Nicolas 1985, 15). In particular, research at that time had primarily focused on resistance to the fiscal, judicial, and military actions of the central state, the nobility, the Church, and municipal authorities, but not to the postal administrations (see, e.g., Bercé 1980; Mandrou 1969).

⁵⁷ We report the distribution of rebellions in our sample across the type of sources used by Nicolas's (2002) survey in Table A12 in the Supplementary Material. The results of this robustness test are reported in Table A13 in the Supplementary Material.

TABLE 4. Mechanisms

Outcome	Number of rebellions					
	1	2	3	4	5	6
Horse-post relay	0.018*** [0.002]	−0.001 [0.001]	0.007*** [0.002]	0.006*** [0.002]	0.007*** [0.001]	0.005*** [0.002]
Authorities Protagonists	State	Non-state	Taxation	Military	Judiciary	Notables
Parish-decades	54,672	54,672	54,672	54,672	54,672	54,672
Parishes	6,834	6,834	6,834	6,834	6,834	6,834
Switchers (overall)	1,929	1,929	1,929	1,929	1,929	1,929
Switchers (at least once)	849	849	849	849	849	849
Outcome mean	0.009	0.002	0.006	0.002	0.002	0.003
Outcome residual s.d.	0.089	0.040	0.073	0.043	0.039	0.046

Note: This table reports the average total effect of non-normalized event-study effects δ_c estimated through Equation 1. All regressions include parish and decade fixed effects, as well as wheat prices and decade-by-canton fixed effects. The estimation is restricted to two leads and three lags. The outcomes are rebellions against state authorities in Column 1, non-state public authorities (seigneurial, nobility, municipal, and Church) in Column 2, taxation authorities in Column 3, military authorities in Column 4, and judiciary authorities in Column 5. Column 6 includes rebellions involving local notables as protagonists. *Parish-decades* and *Parishes* denote the number of observations effectively used in the estimation. *Switchers (overall)* denotes the number of parish-decade observations that switched treatment over the period, while *Switchers (at least once)* denotes the number of parishes that switched at least once. *Outcome residual s. d.* denotes the outcome standard deviation once decade, parish fixed effects, and other controls are accounted for. Estimates are calculated using de Chaisemartin et al.'s (2023) `did_multiplegt_dyn` Stata command. Standard errors are clustered at the parish level and reported in brackets. *** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.

based on arbitrary square grids of varying sizes. In all cases, the results remain robust. Further details and results are provided in Appendix C of the Supplementary Material.

MECHANISMS

In this section, we provide evidence that suggests that the mechanisms linking the introduction of new horse-post relays to increased rebellion are rooted in the material consequences of the network's expansion. By facilitating state centralization and penetration into local society, the horse post enabled greater tax enforcement and conscription, stricter administration of laws and regulations, and the disruption of private interests.

Table 4 probes these mechanisms in several ways. Column 1 restricts the dependent variable to rebellions targeting state authorities, including tax, military, and judicial authorities.⁵⁸ These agents represented some of the most consequential manifestations of central state power, charged with transmitting royal communications, maintaining order, and enforcing rules and directives handed down from the monarchy—reaching deep into local society. Our rebellion data capture a wide range of actions against these agents, including

resistance against court rulings, conscription efforts, tax collection, and interventions by the military and *maréchaussée*. Results reveal a robust link between the establishment of new horse-post relays and rebellions against these authorities. These account for over three-fifths of the total effect of horse-post relays on rebellions. In contrast, column 2 indicates no comparable relationship between the expansion of the horse post and rebellions directed against non-state authorities—the nobility, the Church, as well as seigneurial and municipal authorities.

Columns 3–5 unpack the relationship between the expansion of the horse-post relay network and rebellions targeting state authorities. Column 3 focuses on rebellions specifically directed against taxation. As the reach of the central state expanded, obligations to local nobles diminished, but were replaced by new and increasingly burdensome taxes owed to the state, which became a major source of popular anger (Touzery 2024). Indeed, drawing on an analysis of the Revolution-era *Cahiers de doléances*, Shapiro and Markoff (1998) find evidence that peasants harbored grievances over taxation that were shaped by concerns over utility and fairness. Although the establishment of new horse-post relays did not directly alter local tax regimes, the improved speed and reliability of communication likely facilitated stricter tax enforcement, thereby crystallizing ire against state tax obligations. Column 3 confirms that this mechanism translated into greater local rebellion, as the estimate suggests that rebellions against taxation account for about one-fourth of the total effect of the introduction of new horse-post relays on rebellions.

⁵⁸ See Table A2 in the Supplementary Material for a breakdown of rebellion types by target—state versus non-state authorities—along with the number of cases in each category.

Column 4 examines rebellions targeting the military, including its conscription activities. This category further encompasses rank-and-file mutinies as well as uprisings targeting the *maréchaussée*, the state-sponsored security force policing the countryside. The results show a positive and statistically significant effect. Consistent with this result, we find that rebellions are more responsive to the introduction of horse-post relays in areas that were larger contributors of troops over the eighteenth century.⁵⁹ The coefficient in column 4 accounts for about one-fifth of the total effect of the introduction of new horse-post relays on rebellions.

In column 5, we turn to rebellions targeting the judiciary, including uprisings triggered by judicial actions, such as court-ordered imprisonments and the repossession of property. Indeed, the stricter enforcement of order and rules set by the monarchy—likely facilitated by greater information flow through the horse-post relay network—may have disrupted local social balance and generated grievances. Once again, we find a positive and statistically significant link between new horse-post relays and this subset of rebellions. The coefficient in column 5 comprises another one-fourth of the total effect of the introduction of new horse-post relays on rebellions.

Lastly, column 6 adopts a different approach by focusing on the protagonists of rebellions rather than their targets. Specifically, it examines whether rebellions led by notables increased following the introduction of new horse-post relays. As discussed above, the private interests of notables—such as the provisioning of goods along postal roads—may have been disrupted as the state asserted its presence by monopolizing roads through the new horse-post relays. Consistent with this mechanism, we find a positive link between new relays and rebellions led by notables, with this subset accounting for about one-fifth of the total effect of the introduction of new horse-post relays on rebellion.⁶⁰ This finding is also in line with De Juan (2016) and Garfias and Sellars (2022), who argue that local elites may resist infrastructure and associated state expansion when such efforts threaten their authority and local dominance.

The Role of Pre-Existing State and Administrative Fragmentation

The causal relationship we have identified between horse-post relays and rebellions thus far was part of a broader process of state-building—one in which persistent state fragmentation and enduring subnational

divisions continued to undergird local unrest. For instance, boundaries between different tax regimes, particularly those associated with the *gabelles*, became hotspots for tax evasion and smuggling that disproportionately resisted the state's centralization efforts (Davoine, Enguehard, and Kolesnikov 2025). More broadly, responses to the establishment of horse-post relays may have varied with the historical relationship to the central state, as territories incorporated earlier into the kingdom (*pays d'Élection*) were under direct monarchic authority—especially in fiscal matters—whereas peripheral territories incorporated later (*pays d'États* and *pays d'Imposition*) retained some degree of autonomous ruling. Indeed, these latter regions had historically offered stronger opposition to the central state and were more likely to oppose efforts at state centralization and its penetration into local society (Degraeve 2023).

We provide evidence for the conditional role of pre-existing state and administrative fragmentation along these dimensions. To test for the role of fiscal fragmentation, we split the sample between parishes located near *gabelles* frontiers—defined as those in cantons directly adjacent to these fiscal boundaries—and those farther away. We then estimate the treatment effect of the introduction of horse-post relays on rebellions specifically targeting taxation authorities.⁶¹ The results, reported in Table 5, indicate a markedly stronger response to state expansion in fiscally fragmented areas. Indeed, the estimate in column 1 represents a little more than the average incidence of rebellions against taxation, while the column 2 estimate represents only one fifth. Next, to test for the role of the mode of ruling more broadly, we split the sample between parishes in *pays d'États* and *pays d'Imposition* relative to the more historically centralized *pays d'Élection*.⁶² We then estimate the treatment effect of the introduction of horse-post relays on all rebellions. Consistent with our hypothesis, we find a stronger response to state expansion in areas that were historically more remote from the state's direct authority, with the estimate in column 3 representing about three times the average incidence of rebellions relative to a little less than this average in column 4.⁶³

⁵⁹ See Table A14 in the Supplementary Material, where we divide the kingdom of France into low and high military-recruitment areas based on data from Corvisier (1964, 410–48) for 1716 and Komlos (2003) for 1716–84. See the Data Appendix section of the Supplementary Material for more details.

⁶⁰ Unfortunately, Nicolas's (2002) survey does not document whether other non-state elites—such as members of the clergy or the nobility—participated in rebellions.

⁶¹ See Figures A14 and A15 in the Supplementary Material. Rebellions targeting taxation account for 66% of all rebellions in areas near *gabelles* frontiers (677 out of 1,025), compared to just 34% in areas farther away (1,671 out of 4,975).

⁶² See Figure A12 in the Supplementary Material and Appendix B of the Supplementary Material.

⁶³ Another dimension of administrative fragmentation stemmed from the legal system, which varied substantially across customary regions. At the border of these jurisdictions, litigants frequently resisted the state's attempts to impose uniform legal structures, oftentimes pitting jurisdictions against each other in the hope of securing more favorable legal outcomes (Gay, Gobbi and Goñi 2024a; 2024b). When splitting the sample between parishes located near customary frontiers—defined as those in cantons directly adjacent to these judicial boundaries—and those farther away (see Figure A16 in the Supplementary Material), we find a stronger response to state expansion in judicially fragmented areas.

TABLE 5. Heterogeneity Across State and Administrative Fragmentation

Outcome	Number of rebellions			
	Gabelles		Pays	
	1	2	3	4
Horse-post relay	0.036*** [0.010]	0.003* [0.002]	0.044*** [0.004]	0.017*** [0.006]
Authorities Group	Taxation Close to gabelles front.	Taxation Far from gabelles front.	All États or Imposition	All Élections
Parish-decades	7,768	46,904	22,280	31,864
Parishes	971	5,863	2,785	3,983
Switchers (overall)	255	1,674	834	1,089
Switchers (at least once)	119	730	340	507
Outcome mean	0.021	0.016	0.014	0.019
Outcome residual s.d.	0.139	0.119	0.110	0.130

Note: This table reports the average total effect of non-normalized event-study effects δ_{ℓ} estimated through Equation 1. All regressions include parish and decade fixed effects, as well as wheat prices and decade-by-canton fixed effects. The estimation is restricted to two leads and three lags. Columns 1 and 2 distinguish between parishes that are in cantons adjacent to gabelles frontiers versus not, and focus on rebellions targeting taxation authorities, and Columns 3 and 4 distinguish between parishes that are in *pays d'États* or *pays d'Imposition* versus *pays d'Élection*, and focus on all rebellions. *Parish-decades* and *Parishes* denote the number of observations effectively used in the estimation. *Switchers (overall)* denotes the number of parish-decade observations that switched treatment over the period, while *Switchers (at least once)* denotes the number of parishes that switched at least once. *Outcome residual s.d.* denotes the outcome standard deviation once decade, parish, and canton fixed effects, as well as wheat prices are accounted for. Estimates are calculated using de Chaisemartin et al.'s (2023) `did_multiple_dyn` Stata command. Standard errors are clustered at the parish level and reported in brackets.

*** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.

Gathering Effects

Next we investigate whether the establishment of new horse-post relays impacted rebellious activity in nearby parishes. If new relays brought material changes that extended beyond their host parish—as state agents sought to reach into surrounding areas as well—they could have “gathered” rebellions from nearby, drawing people to express grievances at the locales of this new representation of state authority. The built environment and increased traffic associated with new horse-post relays may have opened new spaces for contention, taking a step in the direction of Tilly’s (1986) theory of evolving repertoires of contention from dispersed, parochial revolts to more clustered and cosmopolitan forms of collective action. Contributing to this shift, new horse-post relays and postillons clearly identified with the markings of increased state visibility and may have represented a disruptive threat to the existing social order. This dynamic constitutes a complementary potential mechanism that extends beyond the parishes where new relays were established.

To test this hypothesis, we estimate three versions of Equation 1, redefining the treatment as an indicator variable that equals one for the presence of a horse-post relay in neighboring parishes within 5 kilometers, 5–10 kilometers, and 10–15 kilometers.⁶⁴ Results in

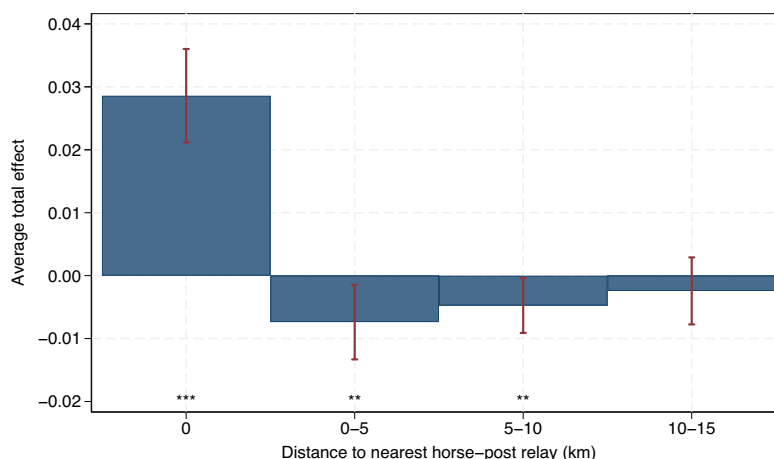
Figure 8 imply that the number of rebellions *decreased* by 0.007 in parishes located within 5 kilometers of a new horse-post relay and by 0.005 in those 5–10 kilometers away, with both estimates significant at the 5% level. Rebellions further decreased by 0.002 in parishes located 10–15 kilometers away from a new relay, although this estimate is not statistically significant.

We interpret these results as suggestive evidence that part of the main effect we identify may reflect a local reconfiguration of rebellious activity from surrounding parishes toward those hosting new horse-post relays. Given the small size of parishes and their close proximity, it would have been relatively easy for people to coordinate resistance locally and direct it toward nearby symbols of state authority. Importantly, however, even after accounting for this “gathering effect” of rebellion, the overall net effect of a new relay on rebellion remains positive and twice as large as the cumulated spillover effects.

ALTERNATIVE EXPLANATIONS

We now explore five main alternative explanations that could account for the observed link between the expansion of the horse-post relay network and rebellion, independent of our primary hypothesized mechanisms: (i) disruption of local social hierarchies; (ii) the role of wartime dynamics and the proximity to war zones; increased capacity of the populace for

⁶⁴ Recall that we are comparing parishes over time within cantons, and the average radius of a canton was about 6 kilometers.

FIGURE 8. Average Total Effect of New Horse-Post Relays on Rebellions across Distance to Nearest Horse-Post Relay

Note: This figure reports the average total effect of a new horse-post relay on rebellion when various treatment variables are input into Equation 1, along with 95% confidence intervals: a horse-post in a given parish and decade, up to 5 kilometers away from a new horse-post (excluding parishes with a relay), 5–10 kilometers away, and 10–15 kilometers away. All regressions include parish and decade fixed effects as well as wheat prices, decade-by-canton fixed effects, and distance to the nearest horse-post relay. Standard errors are clustered at the parish level. The full set of results is available in Table A15 in the Supplementary Material. *** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.

collective action thanks to the concurrent expansion of alternative information channels, such as (iii) the letter-post system and (iv) the underlying road infrastructure; and (v) shifts in the presence of the *maréchaussée*, the public security force policing the countryside.

Disruption of Local Social Hierarchies

Research in other contexts suggests that the erosion of traditional social hierarchies and networks that once structured local society can lead to disorder and to reprisals against declining authorities. For instance, Scott (1976) and Lecce, Ogliari, and Orlando (2022) show that when the disruption of traditional social relations weakens the monopolistic grip of local elites and social networks, it can leave populations vulnerable to severe risks and enable everyday grievances to metastasize rather than be absorbed or resolved locally. These grievances may be directed at both declining elites and the state. Hence, if the expansion of the state near horse-post relays undermined traditional authorities—such as the Church, the nobility, and seigneurial authorities—without eliminating the rents and tithes they extracted, resulting discontent against these declining figures could generate the observed link between new horse-post relays and rebellion. Grievances may also have been targeted at the state itself, especially if it failed to replace the functions previously carried out by these traditional institutions with equivalent state-provided services.

Several pieces of evidence cast doubt on this competing explanation. The establishment of new horse-post relays typically involved forging alliances with

select local elites rather than their weakening, suggesting little disruption to traditional hierarchies (Jamaux-Gohier 2001; Marchand 2006). Accordingly, we find little evidence that new horse-post relays triggered rebellions against traditional non-state authorities. Unpacking the null coefficient in column 2 of Table 4, we find no significant impact on rebellions against the Church or municipal authorities, and if anything, a slightly negative effect on rebellions against seigneurial authorities and nobles. We also find no effect on the broader set of local civilian elites, such as notables and labor authorities.⁶⁵

War Zones and Wartime

War considerations may also have impinged on the horse-post relay network and rebellion by fueling both the network's expansion and grievances against the military in border areas prone to or exposed to war. While most changes to the network during the eighteenth century occurred in the interior of the country, substantial extensions took place in the 1740s and 1750s along the northeastern border with the Holy Roman Empire and in the 1760s along the Swiss border. However, the treatment effect is broadly homogeneous across regions and not driven by a single one. Moreover, these border areas were not necessarily those

⁶⁵ Specifically, the estimate for rebellions against the Church is 0.0002 (standard error: 0.0009); against municipal authorities, 0.0002 (standard error: 0.0002); against seigneurial authorities and nobles, -0.0016 (standard error: 0.0009); against notables, -0.0001 (standard error: 0.0002); and against labor authorities, 0.0009 (standard error: 0.0007).

with the highest levels of conscription, nor were rebellions against the military concentrated there.⁶⁶ Furthermore, the supply routes and stations for troops were organized through the *routes d'étape* system, which was distinct from the horse post and followed a different spatial logic (Kroener 1980).⁶⁷ Nonetheless, to ensure that the military geography does not confound our results, we exclude subdélégations that experienced battles or were along the borders of the kingdom.⁶⁸ The results remain in line with our baseline estimates. We also assess the role of wartime involvement directly by excluding all rebellions that occurred during periods of war, where we define wartime as any year in which France was engaged in a military conflict in Europe.⁶⁹ The results are again similar.

Information and Collective Action Through the Letter-Post System

Another potential alternative explanation is the simultaneous development of the letter-post system, which may have increased the likelihood of rebellion by lowering barriers to collective action through the wider dissemination of information.⁷⁰ Indeed, such dynamics linking improved communication infrastructure to collective mobilization have been observed with the expansion of railways and telegraphs in nineteenth- and twentieth-century United States (García-Jimeno, Iglesias, and Yildirim 2022) and Europe (Melander 2020; Pengl et al. 2025). In eighteenth-century France, the letter post operated as a private service for collecting, forwarding, and delivering letters via an infrastructure of post offices (Marchand 2006, 95–180). If the establishment of letter-post offices coincided with that of horse-post relays—and if those offices facilitated collective action—this correlation could confound our results.

To address this potential concern, we collected information on the locations of letter-post offices for the same years as our data on the horse-post relay network.⁷¹ Critically, the spatial configuration of

letter-post offices followed a different logic from that of horse-post relays, as their establishment followed economic rather than strategic and logistical considerations (Cauvin, Lepetit, and Reymond 1987; Verdier and Chalonge 2018). Letter-post offices were as numerous as horse-post relays and expanded at a comparable pace throughout the eighteenth century, so that they covered most of the territory by the early eighteenth century.⁷² However, unlike the horse post, the letter post was not designed as a network requiring offices at regular intervals. Instead, letter-post offices were generally located in populous towns, in sharp contrast with horse-post relays.⁷³ Their buildings were also much smaller, as they handled only letter distribution and did not serve as rest stops or provide fresh horses. Mail delivery in the letter-post system relied primarily on private couriers (*entrepreneurs de dépêches*), who rode their own horses, and on mail coaches (*malles-poste*), which consisted of two-wheeled carts operated by employees of the letter-post administration.⁷⁴

To empirically assess whether the development of letter-post offices confounds our results, we apply de Chaisemartin and D'Haultfœuille's (2023) estimator, which accommodates staggered designs with binary, non-absorbing *multiple* treatments that may exhibit heterogeneous and dynamic effects. Specifically, we augment Equation 1 by including a set of event-study coefficients on an indicator variable that equals one if a letter-post office was present in a given parish and decade. The results closely mirror our baseline estimates: the average total effect of the horse post on rebellion, accounting for the concurrent development of the letter post, is 0.027 (standard error: 0.011), suggesting that this alternative mechanism does not drive our findings.⁷⁵

Transit Infrastructure and the Road Network

Another potential confounder is the underlying road network on which the horse-post relay system depended, although this infrastructure was managed separately by the Bridges and Roads administration

⁶⁶ See Figures E6 and A13 in the Supplementary Material.

⁶⁷ See, for example, the 1787 map of the routes d'étape at <https://gallica.bnf.fr/ark:/12148/btv1b7100446h>.

⁶⁸ See Figure A13 in the Supplementary Material. Battle locations are drawn from Miller and Bakar's (2023) Historical Conflict Event Dataset (HCED). See the Data Appendix section of the Supplementary Material for more details. The results of this analysis are reported in Table A16 in the Supplementary Material.

⁶⁹ This includes the War of the Austrian Succession (1744–47), the Seven Years' War (1756–62), and the Anglo-French phase of the American Revolutionary War (1778–83). The results of this analysis are reported in Table A17 in the Supplementary Material.

⁷⁰ Although the horse-post relay network was initially reserved for official state communication, by the mid-eighteenth century, it had opened to private correspondence and passenger transport through stagecoaches. However, its private use remained prohibitively expensive compared to the letter post and private-messenger services, effectively limiting its use to the wealthy (Arbellot 1980, 115; Jamaux-Gohier 2001, 35–80).

⁷¹ We draw on the *Ordre général des courriers* published in the annual editions of the *Almanach royal*, the contents of which we cross-validated with the *Cartes des bureaux de postes* of 1741 and with the 1754 and 1782 editions of the *Dictionnaires des postes*. See the

Data Appendix section of the Supplementary Material for more details.

⁷² See Figures A17 and A18 in the Supplementary Material.

⁷³ See Figure A19 in the Supplementary Material.

⁷⁴ Although the mail could also be transported by agents of the letter-post administration (*courriers*) using horses from the horse-post relay network (but escorted by postillions), their numbers remained modest throughout the eighteenth century: there were only six *courriers* in 1694; 18 in 1741; 29 in 1771; 79 in 1785; and 173 in 1792 (Marchand 2006, 162–4). By contrast, the number of postillions was far larger—four thousand in 1763 and as many as seven thousand in 1790 (Marchand 2006, 259).

⁷⁵ See Panel a of Figure A20 in the Supplementary Material. We also report event-study estimates of the treatment effect of letter-post offices on rebellion in Panel b of Figure A20 in the Supplementary Material. We find a slightly positive effect, but pre-trends are such that one may not have a causal interpretation of these results: letter-post offices developed in locations that were already experiencing upward trends in rebellion, consistent with the economically-oriented development logic of the letter post.

(Arbellot 1973). If improvements to the road network coincided with the establishment of new horse-post relays, it is possible that the effects we observe were instead driven by changes in the transit infrastructure itself—whether by fostering economic development, disrupting non-state local authorities, or facilitating popular coordination.

In the early eighteenth century, French roads were in a deplorable state (Arbellot 1973, 766). These conditions had been a longstanding concern and lamented since the Colbert administration as a major obstacle to the expansion of internal trade (Blond 2013, 65). However, it was not until the 1730s—by which point the situation had become critical—that the Controller General of Finance Philibert Orry launched an ambitious program of road repair and construction (Massounie 2013).⁷⁶ Spearheaded by Daniel-Charles Trudaine, the program ultimately resulted in the construction of nearly 26 thousand kilometers of road by the 1780s—although only two-thirds of this network remained in passable condition by that time (Arbellot 1973, 772–3).

The order of priority for new road construction, however, was not directly linked to the location of horse-post relays, as roads connecting relays were considered only after those linking Paris to ports, provincial capitals, and the kingdom's borders (Blond 2007, 69–70). While the upkeep of existing roads may have occasionally favored those used by the horse post (Arbellot 1973, 766), it was exceedingly rare for the creation of a horse-post relay to prompt a request for road construction or maintenance from the Bridges and Roads administration, or that such requests would be granted (Demangeon 1995, 66–71; Verdier 2009, 13). Overall, historical accounts suggest that the horse post was not tied to the road network and followed its own logic, and that the actual routes taken by postillons between relays often varied seasonally depending on weather and road conditions (Jamaux-Gohier 2001, 125–45).

A comprehensive empirical test of this alternative mechanism is however not feasible due to the severe limitations of available quantitative information on road construction and maintenance at the local level in the eighteenth century.⁷⁷ Nevertheless, we show in Tables B1 and B2 in the Supplementary Material that there is little heterogeneity in the treatment effect of horse-post relays across parishes located at different distances from a road. This suggests that our results are unlikely to be driven by connectivity to the road network.

⁷⁶ This initiative relied on the implementation of the *corvée des chemins*, a system of forced labor requiring able-bodied men living within 16 kilometers of a road to contribute up to one month of labor per year. Though widely unpopular and a source of local resentment (Shapiro and Markoff 1998, 397; Tonnaire 2005, 78–80), only ten rebellions were directed against this labor institution in our data. Indeed, opposition to this system rather took the form of passive resistance—workers simply failed to appear to work or slacked on the job (Conchon 2016, 120–6).

⁷⁷ This is in part because the Bridges and Roads administration primarily relied on local topographic maps rather than statistical tables (Lepetit 1984, 17–44; Blond 2013).

Policing by the Maréchaussée Brigades

A final potential alternative mechanism is the presence of *maréchaussée* brigades, the state-sponsored security force that policed the countryside. These men-at-arms formed part of the security apparatus of the royal authority and maintained public order along the kingdom's roads (Fressin 2022). Following their consolidation in 1720, brigades consisting of four to five men were stationed in barracks located in the center of towns and urban areas (Fressin and Georges 2022). Their spatial distribution closely resembled that of letter-post offices: they covered the entire territory and were generally located in populous towns.⁷⁸ However, the expansion of the *maréchaussée* was relatively limited: the number of brigades remained stable at five hundred units during the first half of the eighteenth century, increased to about eight hundred in the 1750s and 1760s, and then plateaued until the Revolution.⁷⁹ While these brigades embodied the symbolic presence of the state, they never mustered more than three thousand men and primarily operated in the vicinity of towns and major roads.⁸⁰

Maréchaussée brigades could potentially account for our findings through two distinct channels. First, if brigades systematically accompanied the establishment of new horse-post relays, the effect we identify might reflect the presence of the *maréchaussée* rather than that of the horse post itself. Second, if the presence of brigades made rebellions more likely to be recorded by these state actors, our results could be driven by selection bias in the archival sources used to document rebellious activity.

We address these concerns in a number of ways. First, we show that although parishes that received a horse-post relay were more likely to eventually host a *maréchaussée* brigade, pretreatment trends indicate that they were also more likely to have received a brigade *prior* to the establishment of the relay.⁸¹ In other words, while the spatial development of the two institutions was correlated, there is no evidence of a causal link between them. Second, we extend Equation 1 to control for the potential incidence of *maréchaussée* brigades using de Chaisemartin and D'Haultfœuille's (2023) estimator for multiple treatments. The results closely resemble our baseline

⁷⁸ See Figures A18 and A19 in the Supplementary Material. Data on the spatial distribution of *maréchaussée* brigades are drawn from Fressin's (2021) *Atlas historique de la gendarmerie*. See the Data Appendix section of the Supplementary Material for more details.

⁷⁹ See Figure A17 in the Supplementary Material.

⁸⁰ Urban security remained under the jurisdiction of municipal authorities. More broadly, policing in Ancien Régime France was highly fragmented, with each jurisdiction maintaining its own corps. Rebellions specifically targeting municipal security forces were rare, with only 39 cases recorded in our data.

⁸¹ See Figure A21 in the Supplementary Material. Our data contain 300 rebellions explicitly targeting the *maréchaussée* (5% of all rebellions). Consistent with the response of rebellions against military authorities in column 4 of Table 4, we find that the horse-post relay network fostered rebellions against the *maréchaussée* (estimate: 0.0037; standard error: 0.0012).

estimates, suggesting that the presence of *maréchaussée* brigades does not account for our findings.⁸² Third, we restrict the analysis to parishes that never hosted either a letter-post office or a *maréchaussée* brigade. Results in Table A18 in the Supplementary Material suggest that these “uncontaminated” parishes still exhibit a positive causal relationship between the establishment of horse-post relays and rebellion. Finally, we test the robustness of our findings to removing rebellions recorded in the archives of the *maréchaussée*, finding similar results (see column 6 of Table A13 in the Supplementary Material).

CONCLUSION

Alongside ongoing international rivalries and wars across early modern Europe, major powers—especially France and England—also faced episodic but significant outbreaks of rebellion and revolt. These episodes of domestic unrest unfolded against the backdrop of ambitious state-building efforts. France was among the first states to pursue a systematic program of state consolidation, encompassing a wide range of infrastructural, communication, and extractive initiatives. Yet state-building initiatives can themselves become a site of contestation, even when their ultimate goal is to establish centralized control and social order. We find that improvements in communication infrastructure through the expansion of France’s horse-post relay network sparked episodic rebellions across the territory in the run-up to the French Revolution. We attribute this relationship to pushback against the material consequences of state penetration, such as greater tax and law enforcement, conscription, control over the roads, and the crowding out of private interests. This pushback also varied with prior levels of state fragmentation and underlying territorial heterogeneity.

The dynamics we identify may also have had important implications for the French Revolution itself. Indeed, exploratory analyses suggest that, while the horse-post relay network in itself did not facilitate information diffusion among citizens or reduce their collective action costs, parishes that received relays—and experienced an associated uptick in rebellion—were more likely to later host political societies during the Revolution.⁸³ These societies played a central role

in coordinating revolutionary action and asserting local control as the Ancien Régime collapsed. Prior grievances fostered by the expansion of the horse-post relay network thus appear to have resurfaced when subsequent shocks created an opportunity for mobilization, amplifying more immediate dynamics such as those linked to the 1788 drought or to the diffusion of democratic ideals by French soldiers returning from the American Revolutionary War (Jha and Wilkinson 2023; Waldinger 2024). This could have operated through long-standing frustrations over issues such as taxation or conscription, or even *corvée* labor—grievances that may not have triggered major rebellions in earlier decades but were prominently expressed in the *Cahiers de doléances* (Shapiro and Markoff 1998). This suggests that efforts to build state capacity can generate long-lasting consequences for social order that may unfold in ways that state planners neither anticipate nor intend.

Our findings underscore the importance of conceptually distinguishing the process of state-building from state strength itself. State-building can affect social order through multiple channels, from material consequences, such as taxation and conscription to information diffusion, ease of communication, and popular coordination. Scholars might fruitfully explore these channels in other canonical cases of early contested state-building in order to deepen our understanding of historical as well as ongoing state-building efforts.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0003055425101342>.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/M1XRXT>.

ACKNOWLEDGEMENTS

The authors thank Carles Boix, Charlotte Cavaillé, Anne Degraeve, Bogdan G. Popescu, Louis Rouanet, Megan Stewart, Ronan Tallec, and Dean Yang for insightful feedback. They also thank Loris Cuenot, Annie Dago, Mario Fernandez, Magdelaine Fikri, Wyatt Leaf, Navid Mazidabadifarhani, Gabriel Perez, Maximilien Pointurier, Carly Taylor, Diego Villamizar, and Louis Vitrand for excellent research assistance.

FUNDING STATEMENT

The authors gratefully acknowledge funding from the Pearson Institute at the University of Chicago for this

⁸² See Panel a of Figure A22 in the Supplementary Material. We also report event-study estimates of the treatment effects of *maréchaussée* brigades on rebellion in Panel b of Figure A22 in the Supplementary Material. While we find a positive effect, pre-trends are such that one cannot have a causal interpretation of these results. Indeed, *maréchaussée* brigades developed in locations that were already experiencing upward trends in rebellion, suggesting that their placement was endogenous.

⁸³ The relationship between the eighteenth-century horse-post relay network and political societies during the Revolution holds both in an OLS and a 2SLS framework, using horse-post relays as an instrument for rebellion in predicting the presence of political societies, along with controls and canton fixed effects. See Appendix D of the Supplementary Material for details.

project. V.G. also acknowledges funding from the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program) and ANR-20-CE38-0015 (ObARDI).

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

ETHICAL STANDARDS

The authors affirm this research did not involve human participants.

REFERENCES

- Albertus, Michael, and Victor Gay. 2025. "Replication Data for: State-Building and Rebellion in the Run-Up to the French Revolution." Harvard Dataverse. Dataset. <https://doi.org/10.7910/DVN/MIXRXT>.
- Arbellot, Guy. 1973. "La grande mutation des routes de France au XVIII^e siècle." *Annales. Économies, Sociétés, Civilisations* 28 (3): 765–91.
- Arbellot, Guy. 1980. "Le réseau des routes de poste, objet des premières cartes thématiques de la France Moderne." In *Actes du 104^e Congrès national des Sociétés savantes, Bordeaux 1979, Section d'histoire moderne et contemporaine*, ed. Ministère des universités, 97–115. Paris: Bibliothèque nationale.
- Arbellot, Guy. 1992. *Autour des routes de poste. Les premières cartes routières de la France*. Paris: Bibliothèque nationale, Musée de la Poste.
- Aubert, Gauthier. 2015. *Révoltes et répressions dans la France Moderne*. Paris: Armand Colin.
- Aubert, Gauthier. 2023. *1709. L'année où la révolution n'a pas éclaté*. Paris: Calype.
- Bercé, Yves-Marie. 1980. *Révoltes et révolutions dans l'Europe moderne (XVI^e–XVIII^e siècles)*. Paris: Presses universitaires de France.
- Blond, Stéphane. 2007. "L'atlas des routes royales de Trudaine. La carte, instrument de politique routière des lumières." *Siècles* 25: 67–82.
- Blond, Stéphane. 2013. "The Trudaine Atlas: Government Road Mapping in Eighteenth-Century France." *Imago Mundi* 65 (1): 64–79.
- Boone, Catherine. 2003. *Political Topographies of the African State*. Cambridge: Cambridge University Press.
- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *The Review of Economic Studies* 91 (6): 3253–85.
- Bouton, Cynthia. 1993. *The Flour War*. University Park, PA: Penn State Press.
- Bretagnolle, Anne, and Alain Franc. 2017. "Emergence of an Integrated City-System in France (XVIIth–XIXth Centuries)." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 50 (1): 49–65.
- Bretagnolle, Anne, and Nicolas Verdier. 2005. "Images d'un réseau en évolution. Les routes de poste dans la France préindustrielle." *M@ppemonde* 79 (3): 1–8.
- Bretagnolle, Anne, and Nicolas Verdier. 2006. "La mesure postale, reflet des évolutions dans le rapport au temps et à la distance au siècle des Lumières." *Cahiers de métrologie* 22–3: 63–74.
- Bretagnolle, Anne, and Nicolas Verdier. 2007. "L'extension du réseau des routes de poste en France, de 1708 à 1833." In *Postes d'Europe, XVIII^e–XXI^e siècles*, ed. Muriel Le Roux, 155–71. Paris: Comité pour l'histoire de la Poste.
- Bretagnolle, Anne, Timothée Giraud, and Nicolas Verdier. 2010. "Modéliser l'efficacité d'un réseau. Le cas de la poste aux chevaux dans la France pré-industrielle (1632–1833)." *L'espace géographique* 39 (2): 117–31.
- Briaïs, Bernard. 1984. *Contrebandiers du sel. La vie des faux sauniers au temps de la gabelle*. Paris: Aubier.
- Bulliet, Richard W. 1990. *The Camel and the Wheel*. New York: Columbia University Press.
- Burguière, André, and Jacques Revel, eds. 1989. *Histoire de la France. L'État et les pouvoirs*. Paris: Éditions du Seuil.
- Caillou, François. 2005. *Une administration royale d'Ancien Régime. Le bureau des finances de Tours*. Tours, France: Presses universitaires François-Rabelais.
- Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–30.
- Cauvin, Colette, Bernard Lepetit, and Henry Reymond. 1987. "Cartes postales. Un espace de relation dans la France pré-industrielle." *Histoire & mesure* 2 (3–4): 89–113.
- Centeno, Miguel, Atul Kohli, Deborah Yashar, and Dinsha Mistree. 2017. *States in the Developing World*. Cambridge: Cambridge University Press.
- Chambru, Cédric. 2019. "Do the Right Thing! Leaders, Weather Shocks and Social Conflicts in Pre-Industrial France." EHES Working Paper 161.
- Chambru, Cédric, and Paul Maneuvrier-Hervieu. 2024. "Introducing HiSCoD: A New Gateway for the Study of Historical Social Conflict." *American Political Science Review* 118 (2): 1084–91.
- Conchon, Anne. 2016. *La corvée des grands chemins au XVIII^e siècle*. Rennes, France: Presses universitaires de Rennes.
- Corvisier, André. 1964. *L'armée française de la fin du XVII^e siècle au ministère de Choiseul. Le soldat*. Paris: Presses universitaires de France.
- Cristofoli, Pascal, Marie-Christine Vouloir, Bertrand Duménieu, and Claude Motte. 2021. "Des chefs-lieux de Cassini aux communes de France (1756–1999)." Dataset, V1. EHESS. https://didomena.ehess.fr/concern/data_sets/6395wb092.
- Davoine, Eva, Joseph Enguehard, and Igor Kolesnikov. 2025. "The Political Costs of Taxation." EU Tax Observatory Working Paper 27.
- de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2023. "Two-Way Fixed Effects and Difference-in-Differences Estimators with Several Treatments." *Journal of Econometrics* 236 (2): 105480.
- de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2024. "Difference-in-Differences Estimators of Intertemporal Treatment Effects." *The Review of Economics and Statistics* 1–45. https://doi.org/10.1162/rest_a_01414.
- de Chaisemartin, Clément, Diego Ciccia, Xavier D'Haultfœuille, Felix Knaut, Méline Malézieux, and Doulo Sow. 2023. "DID_MULTIPLEGT_DYN: Stata Module to Estimate Event-Study Difference-in-Difference (DID) Estimators in Designs with Multiple Groups and Periods." Statistical Software Components S459222. Boston College. <https://ideas.repec.org/c/boc/bocode/s459222.html>.
- Degrave, Anne. 2023. "Local Rule, Elites, and Popular Grievances: Evidence from Ancien Régime France." *Journal of Historical Political Economy* 3 (1): 1–29.
- Deharbe, Karine. 2010. *Le Bureau des finances de la généralité de Lyon. XVI^e–XVIII^e siècles*. Paris: Institut de la gestion publique et du développement économique.
- De Juan, Alexander. 2016. "Extraction and Violent Resistance in the Early Phases of State Building." *Comparative Political Studies* 49 (3): 291–323.
- Demangeon, Pierre. 1995. *La poste en Basse-Normandie*. Condé-sur-Noireau, France: Éditions Charles Corlet.
- Elias, Norbert. [1939] 1994. *The Civilizing Process*. Oxford: Basil Blackwell.
- Evans, Peter, Dietrich Rueschemeyer, and Theda Skocpol. 1985. *Bringing the State Back In*. Cambridge: Cambridge University Press.
- Favier, René. 2007. "Une aide instrumentalisée. L'impôt et l'indemnisation des catastrophes dans la France du XVIII^e siècle." In *Les modalités de paiement de l'État moderne*, ed. Marie-Laure Legay, 107–31. Paris: Institut de la gestion publique et du développement économique.
- Fearon, James, and David Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1): 75–90.

- Fressin, Thomas. 2021. "Atlas historique de la gendarmerie." Centre de recherche de l'École des officiers de la gendarmerie nationale. <https://atlas-gendarmerie.fr/>.
- Fressin, Thomas. 2022. "La maréchaussée et les grands chemins, 1720–1791." *Inflexions* 49 (1): 31–9.
- Fressin, Thomas, and Jonathan Georges. 2022. "Maillage territorial et état de la maréchaussée à la fin du XVIIIe siècle." *Les notes du CREOGN* 65: 1–4.
- Furetière, Antoine. 1690. *Dictionnaire universel*. The Hague, Netherlands: Arnout and Reinier Leers.
- García-Jimeno, Camilo, Angel Iglesias, and Pinar Yildirim. 2022. "Information Networks and Collective Action: Evidence from the Women's Temperance Crusade." *American Economic Review* 112 (1): 41–80.
- Garfias, Francisco, and Emily Sellars. 2022. "When State Building Backfires: Elite Coordination and Popular Grievance in Rebellion." *American Journal of Political Science* 66 (4): 977–92.
- Gay, Victor. 2025. "The Jean Nicolas Database. The French Rebellion, 1661–1789." *Data & Corpus* 1 (1): 15892.
- Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2024a. "The Atlas of Local Jurisdictions of Ancien Régime France." *Journal of Historical Geography* 84: 49–60.
- Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2024b. "The Customary Atlas of Ancien Régime France." *Explorations in Economic History* 93: 101588.
- Gazagnadou, Didier. 2013. *La poste à relais en Eurasie*. Paris: Éditions Kimé.
- González, Felipe, Josepa Miquel-Florensa, Mounu Prem, and Stéphane Straub. 2025. "The Dark Side of Infrastructure: Roads, Repression and Land in Authoritarian Paraguay." *The Economic Journal* 135 (666): 653–69.
- Gorry, Jean-Michel. 2008. "Des communes de la Révolution aux communes actuelles." In *Des paroisses de Touraine aux communes d'Indre-et-Loire*, ed. Elisabeth Zadora-Rio, 121–83. Tours, France: FERACF.
- Gorry, Jean-Michel. 2014. "L'administration générale et fiscale d'Ancien Régime." In *Atlas archéologique de Touraine*, ed. Elisabeth Zadora-Rio, 27. Tours, France: FERACF.
- Goubert, Jean-Pierre, Guy Arbellot, and Alexandra Laclau. 1989. "Les subgélédations à la veille de la Révolution de 1789." In *Atlas de la Révolution française. Le territoire (2): Les limites administratives*, eds. Daniel Nordman, Marie-Vic Ozouf-Marignier, and Alexandra Laclau, 47–52. Paris: Éditions de l'EHESS.
- Hanson, Jonathan, and Rachel Sigman. 2021. "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research." *The Journal of Politics* 83 (4): 1495–510.
- Homola, Jonathan, Miguel Pereira, and Margit Tavits. 2024. "Fixed Effects and Post-Treatment Bias in Legacy Studies." *American Political Science Review* 118 (1): 537–44.
- Huntington, Samuel. 1968. *Political Order in Changing Societies*. New Haven, CT: Yale University Press.
- Jamaux-Gohier, Théotiste. 2001. *La poste aux chevaux en Bretagne (1738–1873)*. Mayenne, France: Éditions régionales de l'Ouest.
- Jha, Saumitra, and Steven Wilkinson. 2023. "Revolutionary Contagion." Stanford GSB Research Paper 4084.
- Komlos, John. 2003. "An Anthropometric History of Early-Modern France." *European Review of Economic History* 7 (2): 159–89.
- Krakovitch, Odile, Marie-Thérèse Lalgüë-Guilhemsans, Thierry Roquincourt, and Eugène Vaille. 2000. *Postes. F 90 19001–19104. Répertoire numérique détaillé*. Paris: Centre historique des archives nationales.
- Kroener, Bernhard R. 1980. *Les Routes et les Étapes. Die Versorgung der französischen Armeen in Nordostfrankreich (1635–1661)*. Münster, Germany: Aschendorff.
- Lecce, Giampaolo, Laura Oglia, and Tommaso Orlando. 2022. "State Formation, Social Unrest and Cultural Distance. Brigandage in Post-Unification Italy." *Journal of Economic Growth* 27: 453–83.
- Lefebvre, Georges. 1947. *The Coming of the French Revolution*. Princeton, NJ: Princeton University Press.
- Le Mée, René. 1972. "Population agglomérée, population éparse au début du XIXe siècle." *Annales de démographie historique* 1971 (1): 455–510.
- Lepetit, Bernard. 1984. *Chemins de terre et voies d'eau*. Paris: Éditions de l'EHESS.
- Ladurie, Emmanuel Le Roy. 1966. *Les paysans du Languedoc*. Paris: S.E.V.P.E.N.
- Levi, Margaret. 1989. *Of Rule and Revenue*. Berkeley: University of California Press.
- Levitsky, Steven, and Lucan Way. 2013. "The Durability of Revolutionary Regimes." *Journal of Democracy* 24 (3): 5–17.
- Lindvall, Johannes, and Jan Teorell. 2016. "State Capacity as Power: A Conceptual Framework." STANCE Working Paper.
- Mandrou, Robert. 1969. "Vingt ans après, ou une direction de recherches fécondes. Les révoltes populaires en France au XVIIIe siècle." *Revue historique* 242 (1): 29–40.
- Mann, Michael. 1984. "The Autonomous Power of the State." *European Journal of Sociology* 25 (2): 185–213.
- Marchand, Patrick. 2006. *Le maître de poste et le messenger. Les transports publics en France au temps des chevaux*. Paris: Belin.
- Markoff, John. 1985. "The Social Geography of Rural Revolt at the Beginning of the French Revolution." *American Sociological Review* 50 (6): 761–81.
- Massounie, Dominique. 2013. "Philibert Orry et l'embellissement du territoire autour de l'instruction de 1738. Genèse d'un paysage routier et urbain." *Livraisons de l'histoire de l'architecture* 26: 89–104.
- Melander, Eric. 2020. "Transportation Technology, Individual Mobility and Social Mobilisation." CAGE Working Paper 471.
- Migdal, Joel S. 1988. *Strong Societies and Weak States*. Princeton, NJ: Princeton University Press.
- Miller, Charles, and K. Shuvo Bakar. 2023. "Conflict Events Worldwide since 1468BC: Introducing the Historical Conflict Event Dataset." *Journal of Conflict Resolution* 67 (2–3): 522–54.
- Minetti, Alberto. 2003. "Efficiency of Equine Express Postal System." *Nature* 426: 785–6.
- Motte, Claude, and Marie-Christine Vouloir. 2007. "Le site cassini. ehess.fr. Un instrument d'observation pour une analyse du peuplement." *Bulletin du Comité français de cartographie* 191: 68–84.
- Motte, Claude, Isabelle Séguy, Christine Théré, and Dominique Tixier-Basse. 2003. *Communes d'hier, communes d'aujourd'hui. Les communes de la France métropolitaine, 1801–2001. Dictionnaire d'histoire administrative*. Paris: INED.
- Müller-Crepon, Carl, Philipp Hunziker, and Lars-Erik Cederman. 2021. "Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa." *Journal of Conflict Resolution* 65 (2–3): 563–90.
- Mummolo, Jonathan, and Erik Peterson. 2018. "Improving the Interpretation of Fixed Effects Regression Results." *Political Science Research and Methods* 6 (4): 829–35.
- Nicolas, Jean. 1985. "Un chantier toujours neuf." In *Mouvements populaires et conscience sociale: XVIe–XIXe siècles*, ed. Jean Nicolas, 13–20. Paris: Maloine S. A. Éditeur.
- Nicolas, Jean. 2002. *La rébellion française. Mouvements populaires et conscience sociale (1661–1789)*. Paris: Éditions du Seuil.
- Nougaret, Pierre. 1960. "Histoire de la poste en Bourgogne des origines à 1793." *Mémoires de l'Académie des Sciences, Arts et Belles-Lettres de Dijon* 113: 7–173.
- Ozouf-Marignier, Marie-Vic, Antoine-Marie Graziani, and Serge Bonin. 1989. "Les départements, districts et cantons en 1790." In *Atlas de la Révolution française. Le territoire (2): Les limites administratives*, eds. Daniel Nordman, Marie-Vic Ozouf-Marignier, and Alexandra Laclau, 53–62. Paris: Éditions de l'EHESS.
- Peng, Yannick I., Carl Müller-Crepon, Roberto Valli, Lars-Erik Cederman, and Luc Girardin. 2025. "The Train Wrecks of Modernization: Railway Construction and Separatist Mobilization in Europe." *American Political Science Review* 1–18. <https://doi.org/10.1017/S0003055425000048>.
- Perret, Julien, Maurizio Gribaudi, and Marc Barthelemy. 2015. "Roads and Cities in 18th Century France." *Scientific Data* 2: 150048.
- Pierskalla, Jan, Alexander De Juan, and Max Montgomery. 2019. "The Territorial Expansion of the Colonial State: Evidence from German East Africa 1890–1909." *British Journal of Political Science* 49 (2): 711–37.

- Porchnev, Boris. 1963. *Les soulèvements populaires en France de 1623 à 1648*. Paris: S.E.V.P.E.N.
- Pruett, Lindsey. 2024. "Resisting the Blood Tax: Coercive Capacity, Railroads, and Draft Evasion in Colonial West Africa." *The Journal of Politics* 86 (4): 1207–20.
- Ridolfi, Leonardo. 2019. "Six Centuries of Real Wages in France from Louis IX to Napoleon III: 1250–1860." *The Journal of Economic History* 79 (3): 589–627.
- Rogowski, Jon, John Gerring, Matthew Maguire, and Lee Cojocar. 2022. "Public Infrastructure and Economic Development: Evidence from Postal Systems." *American Journal of Political Science* 66 (4): 885–901.
- Schenoni, Luis. 2025. *Bringing War Back In: Victory, Defeat, and the State in Nineteenth-Century Latin America*. Cambridge: Cambridge University Press.
- Scott, James. 1976. *The Moral Economy of the Peasant*. New Haven, CT: Yale University Press.
- Shapiro, Gilbert, and John Markoff. 1998. *Revolutionary Demands: A Content Analysis of the Cahiers de Doléances of 1789*. Stanford, CA: Stanford University Press.
- Shimbel, Alfonso. 1953. "Structural Parameters of Communication Networks." *Bulletin of Mathematical Biology* 15 (4): 501–7.
- Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.
- Slater, Dan. 2010. *Ordering Power*. Cambridge: Cambridge University Press.
- Soifer, Hillel. 2008. "State Infrastructural Power: Approaches to Conceptualization and Measurement." *Studies in Comparative International Development* 43: 231–51.
- Soifer, Hillel. 2015. *State Building in Latin America*. Cambridge: Cambridge University Press.
- Straus, Scott. 2006. *The Order of Genocide*. Ithaca, NY: Cornell University Press.
- Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–99.
- Tanguy, Jean-François, Jean Le Bihan, and Yann Lagadec, eds. 2009. *Le canton, un territoire du quotidien?* Rennes, France: Presses universitaires de Rennes.
- Thompson, Edward. 1964. *The Making of the English Working Class*. New York: Pantheon Books.
- Thompson, Edward. 1971. "The Moral Economy of the English Crowd in the Eighteenth Century." *Past & Present* 50 (1): 76–136.
- Tilly, Charles. 1986. *The Contentious French*. Cambridge, MA: Harvard University Press.
- Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990–1990*. Oxford: Basil Blackwell.
- de Tocqueville, Alexis. [1856] 2010. *The Old Regime and the French Revolution*. New York: Anchor Books.
- Tonnaire, Noémie. 2005. "Monter une route en poste au XVIII^e siècle. L'itinéraire de Dijon à Auxerre, par Vitteaux." Master's thesis. Université de Bourgogne.
- Touzery, Mireille. 2024. *Payer pour le roi. La fiscalité monarchique. France, 1302–1792*. Paris: Champ Vallon.
- Vaillé, Eugène. 1953. *Histoire générale des postes françaises. La ferme générale et le groupe Grimod-Thiroux (1738–1789)*. Paris: Presses universitaires de France.
- Verdier, Nicolas. 2009. "La liabilité du réseau aux XVIII^e et XIX^e siècles. Le cas de la poste aux chevaux dans le quart nord-ouest de la France." *Les nouvelles de l'archéologie* 115: 13–7.
- Verdier, Nicolas, and Ludovic Chalonge. 2018. "La question des échelles en géohistoire. Les bureaux de la poste aux lettres du XVIII^e siècle à nos jours." *Cybergeo: European Journal of Geography* 851. <https://doi.org/10.4000/cybergeo.29116>.
- Verdier, Nicolas, Timothée Giraud, Christophe Mimeur, and Anne Bretagnolle. 2025. "Postal Horse Relays and Roads in France, from the 17th to the 19th Centuries." *Cybergeo: European Journal of Geography* 1088. <https://doi.org/10.4000/13gxr>.
- Waldinger, Maria. 2024. "'Let Them Eat Cake': Drought, Peasant Uprisings, and Demand for Institutional Change in the French Revolution." *Journal of Economic Growth* 29 (1): 41–77.
- Weber, Max. 1968. *Economy and Society*. New York: Bedminster Press.



Article de données

The Jean Nicolas Database

The French Rebellion, 1661–1789

GAY, Victor *  ¹

¹ Department of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole

* Mail : victor.gay@tse-fr.eu

Domaines

Sciences sociales — Histoire

Date de soumission : 18 juin 2025

Date d'acceptation : 7 novembre 2025

Date de publication : 11 décembre 2025

© The Authors

Licence : [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Publié par l'Université de Lorraine

DOI : <https://doi.org/10.46298/dc.15892>

Résumé

Cet article présente la base de données Jean Nicolas, une ressource exhaustive recensant environ 8 516 rébellions survenues en France entre 1661 et 1789. Fondée sur une enquête menée par Jean Nicolas du début des années 1980 à la fin des années 1990, la base de données consigne pour chaque événement sa typologie, sa chronologie, son lieu, les caractéristiques des participants, les formes de confrontation et de violence, les conséquences judiciaires, les sources et l’auteur de la notice. En plus de détailler la méthodologie de construction de la base, l’article en évalue de manière critique la fiabilité, en analysant les biais introduits lors de la collecte initiale des données. Il fournit également des conseils méthodologiques aux utilisateurs finaux afin d’en atténuer les limites.

Abstract

This article introduces the Jean Nicolas database, a comprehensive resource documenting 8,516 rebellions in France between 1661 and 1789. Based on a survey conducted by Jean Nicolas from the early 1980s to the late 1990s, the database records each event’s typology, chronology, location, participant characteristics, forms of confrontation and violence, legal consequences, sources, and authorship. In addition to detailing the construction methodology of the database, the article critically evaluates its reliability by analyzing biases introduced during the original data collection process. It also provides methodological guidance to end users to mitigate the database’s limitations.

Mots-clés [conflit social](#), [troubles publics](#), [histoire sociale](#), [histoire européenne](#), [données](#)

Keywords [social conflict](#), [civil disturbances](#), [social history](#), [European history](#), [data](#)

Informations sur le jeu de données

Title	The Jean Nicolas Database
Author	Gay, Victor
Publication date	2025-09-22 (Version 2.0) 2025-06-17 (Version 1.0)
Version	2.0

Link	https://doi.org/10.7910/DVN/ANQXMQ
Identifier	doi:10.7910/DVN/ANQXMQ
Availability	Open access
License	CC-BY 4.0
Formats and metrics	Database: 6 folders and 9,019 files (2.03 GB) ▪ README file (TXT 5 KB) ▪ Image file <code>nicolas_logbook</code> (PDF 22 MB) ▪ Folder <code>nicolas_records</code>: 8,977 image files (PDF 1.51 GB) ▪ Folder <code>nicolas_events</code>: 8 data files (CSV, TXT, DTA 223 MB) ▪ Folder <code>nicolas_events_all</code>: 8 data files (CSV, TXT, DTA 234 MB) ▪ Folder <code>nicolas_sources</code>: 8 data files (CSV, TXT, DTA 32 MB) ▪ Folder <code>nicolas_authors</code>: 8 data files (CSV, TXT, DTA 123 KB) ▪ Folder <code>nicolas_hiscod</code>: 8 data files (CSV, TXT, DTA 20 MB)
Language	en, fr
Bibliographic reference	Gay, Victor. (2025, 22 septembre). The Jean Nicolas Database (version 2.0) [Dataset]. Harvard Dataverse. https://doi.org/10.7910/DVN/ANQXMQ

1. INTRODUCTION

Rebellions have long been a central concern for social scientists and historians. Far from being mere disruptions, episodes of social conflict expose deeper tensions over power, inequality, values, and competing visions of society. Because of their transformative potential, scholars have examined how these conflicts shape a wide range of societal outcomes such as state formation (Slater, 2010), political consciousness (Porchnev, 1963; Thompson, 1971), regime durability (Levitsky & Way, 2013), or collective identity (Le Roy Ladurie, 1966). Early works in this tradition focus on early modern Europe—a period of rapid social and economic change accompanied by the consolidation of state power (Le Roy Ladurie, 1966; Porchnev, 1963; Thompson, 1971; Tilly, 1986). Among these conflicts, the French Revolution stands out as a landmark event that has inspired a vast scholarly literature (e.g., Andress, 2013). Yet the thousands of localized uprisings that erupted across the French countryside throughout the eighteenth century have received comparatively less attention, despite their significance as precursors to the Revolution (Aubert, 2015, 2023; Nicolas, 2002).

In recent years, however, renewed interest in historical political economy and improved access to historical data have spurred a new wave of empirical research on rebellion over this period (Cirone, 2023; Jenkins & Rubin, 2024). A key driver of this resurgence has been the availability of the Historical Social Conflict Database (HiSCoD; Chambru and Maneuvrier-Hervieu, 2024), which compiles about 21 thousand episodes of social conflict in Europe from the High Middle Ages to the late nineteenth century. This resource has proved instrumental in enabling new empirical studies on the roots of rebellion, particularly in early modern France (e.g., Chambru, 2019; Davoine et al., 2025; Degrave, 2023; Giommoni et al., 2025; Jha & Wilkinson, 2023; Ottinger & Rosenberger, 2023). A significant part of the HiSCoD database draws from the Jean Nicolas survey, which documents more than 8,500 rebellious events that occurred in France between 1661 and 1789 (Nicolas 1973, 1974, 1985a, 2002).¹ However, the database includes only a limited subset of the information originally recorded in the Jean Nicolas survey.² Moreover, the companion article to the database offers little detail about the context in which the Jean Nicolas survey was developed or on the potential biases embedded in its content (Chambru & Maneuvrier-Hervieu, 2024). Yet the meaningful reuse of data from past surveys requires a clear

¹ These events represent about 40 percent of the HiSCoD database and 54 percent of its entries concerning France.

² The HiSCoD database provides information on the typology, date, location, number and gender of participants of each event, together with the sources and authors of each record.

understanding of the conditions under which the data were initially produced (Girault, 2019).

To address these two limitations, I offer a detailed account of the development of the Jean Nicolas survey, tracing its origins in the early 1970s through to the publication of Jean Nicolas's (2002) seminal book, *The French Rebellion (La rébellion française)*. In particular, I highlight the *de facto* regional stratification of the survey. I further examine Jean Nicolas's definition of the concept of *rebellion*, emphasizing how its broad scope may blur the line between political unrest and more ambiguous phenomena, such as ordinary criminal acts (Section 2). I then introduce the Jean Nicolas database, a comprehensive resource that captures *all* the information originally recorded in the Jean Nicolas survey. For each event, the database documents its typology, chronology, location, participant characteristics, forms of confrontation and violence, and legal consequences—a total of 284 variables. In addition, the Jean Nicolas database includes auxiliary datasets offering detailed information about the sources and authors of each record. It also provides the scanned images of all the original records of the Jean Nicolas survey (Section 3). Next, I provide a critical assessment of the reliability of the Jean Nicolas database. Specifically, I first compare the content of the database with the corresponding records in HiSCoD, identifying some discrepancies in Chambru and Maneuvrier-Hervieu's (2024) database. I then analyze *quantitatively* the biases introduced by the survey methodology due to source selection, authorship patterns, and regional archival coverage. I also propose a grading system that documents the quality of each record (Section 4). On the basis of this quantitative source criticism exercise, I offer practical recommendations for researchers using the database in empirical work to help ensure that their analyses yield robust and credible results (Section 5). Finally, Section 6 describes the availability of the Jean Nicolas database, and Section 7, its potential uses.

The construction of the Jean Nicolas database is part of a threefold movement. First, it reflects the recent revival of quantitative history in both the French and Anglo-American traditions (Cirone, 2023; Karila-Cohen et al., 2018; Lemerrier & Zalc, 2019; Ruggles, 2021). Second, it contributes to an ongoing cross-disciplinary movement to upcycle historical data collections, whose objective is to make these collections interoperable and reusable through reproducible workflows (Scheltjens, 2023). Finally, it adopts an interdisciplinary perspective and aims to foster a dialogue between history and the quantitative social sciences by providing users with the conceptual tools needed to avoid common pitfalls in quantitative historical research (Armatte, 2020; Dennison & Gehlbach, 2024; Karila-Cohen et al., 2018; Lemerrier et al., 2013; Lemerrier & Zalc, 2019).

2. THE JEAN NICOLAS SURVEY OF REBELLIONS

This section traces the development of the Jean Nicolas survey, from its origins in the early 1970s to its eventual digitization in the form of the Jean Nicolas database in 2025 (Section 2.1). It also examines Jean Nicolas's definition of the concept of *rebellion*, which formed the basis for the survey's data collection (Section 2.2).³

2.1. The Making of the Jean Nicolas Survey

2.1.1. Origins of the Survey (1973–74)

The Jean Nicolas survey of rebellions originated as the initiative of a single historian. While studying the nobility and bourgeoisie of eighteenth-century Savoy (Nicolas, 1978), Jean Nicolas became increasingly interested in the dynamics of popular uprisings. Challenging the prevailing view in the 1970s that popular revolts had largely disappeared by the mid-seventeenth century, he argued that social conflict remained a vital force in shaping historical change well into the eighteenth century. Guided by his Marxist convictions, he sought to restore a sense of “conflictual and creative coherence” to popular action, emphasizing that the people acted rationally and that their participation in conflict had meaningful political implications. This line of inquiry resulted in the programmatic article “Chronicle of Refusal. Toward an Inquiry into Popular Emotions in the Eighteenth Century, the Case of Savoy” (“Éphémérides du refus. Pour une enquête sur les émotions populaires au XVIII^e siècle, le cas de la Savoie”), in which Jean Nicolas compiled 234 rebellions that occurred in Savoy between 1650 and 1792 (Nicolas, 1973, 1974). Although regional in scope—reflecting the broader orientation of French historiography on early modern revolts in the 1970s (Bercé, 1974; Foisil, 1970; Pillorget, 1975)—this study laid the conceptual foundations for what would evolve into a nationwide investigation of popular uprisings.

2.1.2. Design and Launch of the Survey (1982–85)

While Jean Nicolas was the driving force behind the survey from the outset, its concrete design took shape through a collaborative process in 1982–83 as he consulted a number of early modern historians—including Yves-Marie Bercé, Denis Richet, and Hugues Neveux—to define the project's scope and methodology. The survey aimed to systematically and cumulatively document all traces of popular uprisings in France between the beginning of Louis XIV's personal reign in 1661 and the outbreak of the French Revolution in 1789. Each event was to be recorded on an individual record through an elaborate coding grid and categorized within a detailed typology. The

³ This section draws extensively on Hamon (2021, 2022) as well as informal conversations with Philippe Hamon, who interviewed Jean Nicolas in 2017.

resulting (paper-based) database was then to be processed by computer for statistical analysis.⁴

Following this preparatory phase, the survey was officially launched in 1983. Based in Paris, Jean Nicolas reached out to modern history professors across France, inviting them to participate directly in the project or to involve their graduate students.⁵ To achieve comprehensive national coverage, he tried to articulate three spatial frameworks: the provincial geography of Ancien Régime France, the spatial distribution of university centers, and the location of départemental archives—the most promising archival repositories. This approach resulted in a *de facto* regional stratification of the survey—Section 4.2 examines the biases introduced by this strategy. A key moment in the survey’s development came in May 1984 with the conference “Popular Movements and Social Consciousness” (“Mouvements populaires et conscience sociale”) organized by Jean Nicolas at Paris–VII. While its thematic scope extended from the fifteenth to the nineteenth century and beyond France, the event included a dedicated workshop on the survey. The proceedings of the conference featured the survey’s typology of protest movements and coding grid (Nicolas, 1985*b*, pp. 761–767).⁶

From that point on, Jean Nicolas assumed a central role as both coordinator and data processor. He personally managed the entire workflow: mailing out survey forms, collecting completed records, numbering them, correcting inconsistencies, and supplementing missing information when necessary. As uneven regional participation revealed gaps in the survey’s coverage, Jean Nicolas increasingly stepped in as an investigator, conducting archival research himself to fill these gaps. Although the data collection was initially scheduled to conclude in 1985, it ultimately continued until 1999.

2.1.3. Results of the Survey (1986–2002)

The first results of the survey were presented by Jean Nicolas during a roundtable held at Paris–VII in October 1986 on “Popular Uprisings in France in the Seventeenth and Eighteenth Centuries.” His unpublished presentation, titled “Emotions in the Computer. Preliminary Results of a Collective Survey” (“Les émotions dans l’ordinateur. Premiers résultats d’une enquête collective,” cited in Markoff, 1990, p. 413), reported 2,878 recorded

⁴ On the transformative impact of computer-assisted database construction on historical research, see Genet (1986), and more recently, Dedieu (2024).

⁵ The survey operated with minimal funding. Jean Nicolas occasionally secured financial support from the mission of the Bicentennial of the Revolution and from the 1983 CNRS program “Genesis of the Modern State” (“Genèse de l’État moderne,” Genet, 1997). However, the project generally lacked institutional support and relied largely on the unpaid contributions of university faculty.

⁶ Appendix Figures A1–A3 reproduce the original typology of the survey, while Appendix Figures A4–A7 reproduce its original coding grid.

rebellions. By 1988, the number of documented events had grown to 5,125 (Burguière, 1991).⁷ Although Jean Nicolas retired in 1994, he continued to develop the survey until 1999. In its final version, the survey comprised 8,516 distinct rebellion records. The culmination of this vast undertaking was the publication of Nicolas's (2002) seminal book, *The French Rebellion. Popular Movements and Social Consciousness (1661–1789)* (*La rébellion française. Mouvements populaires et conscience sociale (1661–1789)*), which extensively drew on his survey.

2.1.4. From the Survey to the Database (2015–25)

The original records of the Jean Nicolas survey resurfaced a few years after the publication of *The French Rebellion* thanks to the serendipitous workings of academic networks. Shortly after retiring, Nicolas sought to deposit the forty boxes of survey materials at his home institution, the University of Paris-VII, but the administration declined—as did several other Parisian universities. In 2015, Nicolas raised the issue with Jean Aubert, a historian who had worked in several museums in Savoie, Nicolas's native region. Jean Aubert then relayed the matter to his son, Gauthier Aubert, who had recently completed his *mémoire* for the Habilitation to Supervise Research on early modern rebellions in Brittany (Aubert, 2014). Seizing the opportunity, Gauthier Aubert personally transported the boxes to his home institution, the University of Rennes-2, where he deposited them at the François Lebrun Library.

In 2017, I contacted Philippe Hamon and Gauthier Aubert to inquire about the whereabouts of Jean Nicolas's original survey records. I then traveled to the François Lebrun Library, where I digitized the entire collection using a portable scanner.⁸ The 8,977 PDF files of the Jean Nicolas survey—totaling 33,490 pages—are included in the Jean Nicolas database under the `nicolas_records` folder.⁹ Over the following years, and with the assistance of three research assistants, I transcribed the full content of these

⁷ Appendix Figure A8 reproduces Burguière's (1991) figure of rebellions from 1750 to 1789 drawn based on the Jean Nicolas survey as of 1988.

⁸ This digitization effort was partially carried out by Estefania Santacreu-Vasut. The resulting scans were distributed in 2021 via the En colère platform (<https://www.projetcolere.org/>), developed by Pascal Bastien (Université du Québec à Montréal) and Jamie Folsom (Performant Software).

⁹ This total page count does not correspond to 35,944 (= 8,986 × 4) because some records include supplementary description pages, while others contain two grids per page. Jean Nicolas's logbook tracking his correspondence with collaborators was also deposited at the François Lebrun library and scanned (`nicolas_logbook.pdf`).

records into a Microsoft Access database, which I then processed and formatted into tabular datasets using the Stata software.¹⁰

2.2. The Notion of Rebellion in the Jean Nicolas Survey

2.2.1. The Semantics of Revolt

Before Nicolas's (2002) book, the term *rebellion* was rarely used in the French historiography to describe popular uprisings in early modern France. For instance, in his seminal *History of the Croquants. A Study of Popular Uprisings in 17th-Century Southwestern France* (*Histoire des Croquants. Étude des soulèvements populaires au XVII^e siècle dans le Sud-Ouest de la France*), Bercé (1974, p. 173) uses *rebellions* in a narrow sense, referring specifically to acts led by royal officers. He explicitly distinguishes them from *revolts*, noting that *rebellions* refer to "violent but isolated act[s] with no lasting consequences." This reflects an implicit hierarchy of intensity by which Bercé (1974) considered *rebellions* as less important than *revolts*.¹¹

This preference for an alternative terminology is consistent across the seminal historiography: Porchnev (1963) refers to *popular uprisings*; Mousnier (1967) and Foisil (1970), to *revolts*; and Pillorget (1975), to *insurrectionary movements*. Likewise in historical dictionaries: Cabourdin and Viard's (1978) entries are *peasant revolts* and *emotions*; Pillorget and Pillorget's (1996), *popular uprisings*; and Bély's (1996), *riots* and *revolts*.

Even Jean Nicolas initially employed alternative terminology. His pioneering 1973–74 article referred to *popular emotions*; the 1984 conference, to *popular movements*; and the 1986 roundtable, to *popular uprisings*, with Jean Nicolas's own presentation referring to *emotions*. It was ultimately with Nicolas's (2002) *The French Rebellion* that the term *rebellion* gained significant visibility, for instance by quickly making its way into Grenier et al.'s (2003) and Bourquin et al.'s (2005) historical dictionaries.

2.2.2. What Counts as Rebellion?

In the first chapter of his book, "The Archives of Disorder," Nicolas (2002, pp. 19–25) revisits the concept of rebellion that guided the construction of his survey. While he draws on categories used by historical actors—such as *emotions*, *riots*, or *sedition*, as found in archival sources—he ultimately constructs his own analytical framework. He defines a rebellion as an event that involves violence, engages at least four individuals from more than

¹⁰ The research assistants who contributed to the transcription were Loris Cuenot, Annie Dago, and Louis Vitrand. This process was supervised in collaboration with Michael Albertus.

¹¹ Bercé (1974, p. 681) reinforces this distinction in his conclusion, where he describes the topographical unit of *rebellions* as the local community, further suggesting that he associates the term with smaller-scale, localized events.

one family, may last for only a few hours, and expresses a group's need for survival or collective assertion—his ambition was to “reach the threshold of the barely perceptible, where the most rudimentary forms of awareness emerge” (p. 27).¹² Although most of these events implicitly involve a confrontation with authorities, this criterion is notably absent from Nicolas's formal definition. In fact, he explicitly includes “violent forms of collective release, neighborhood or parish rivalries, student riots, [and] festive disturbances” (p. 27), which often unfold among social equals and do not necessarily involve any confrontation with authorities. In the specific case of strikes, he even sets aside violence as a necessary criterion (Nicolas, 2002, Footnote 41, p. 27).

Taken together, these criteria allow the Jean Nicolas survey to cast a wide net, capturing the “sporadic” forms of social contestation that might otherwise escape the historian's scope (Nicolas, 1973, p. 594). However, this expansive scope also presents analytical challenges, as it blurs the boundaries between political unrest and more ambiguous phenomena, such as ordinary criminal acts (e.g., prison escapes) or certain forms of intercommunal violence. As a result, users of the Jean Nicolas database should proceed with caution and pay close attention to the typology of events they include in their analyses.

3. THE JEAN NICOLAS DATABASE

The Jean Nicolas database comprises three main datasets: an events dataset (Section 3.1), a sources dataset (Section 3.2), and an authors dataset (Section 3.3).

3.1. The Events Dataset

The events dataset (`nicolas_events`) is the central component of the Jean Nicolas database. It provides detailed information on 8,516 distinct rebellions, organized across ten sets of variables. For each rebellion, the dataset records its typology, chronology, location, participant characteristics, forms of confrontation and violence, legal consequences, sources, authorship, and a description of the event—the 284 variables included in the events dataset are listed in Panel A of Appendix Table A1. Each event is assigned a unique four-digit `nicolas` identifier, ranging from 0001 to 8843

¹² Jean Nicolas selected a threshold of four individuals to reflect contemporary regulations that prohibited gatherings exceeding that number when public order was perceived to be under threat (Nicolas, 1985*b*, p. 75).

and matching the number found on the top-right corner of each original record.¹³

Records *not* included in the events dataset. The Jean Nicolas survey originally comprises 8,977 records, of which 461 are not included in the final events dataset. These excluded records fall into three categories: deleted, duplicate, and incomplete records. First, Jean Nicolas marked 19 records for deletion, either because they occurred before 1661, did not describe a rebellious event, or were inaccurate.¹⁴ Second, he identified 238 records as duplicates, typically when multiple authors created distinct records describing the same event—I refer to these as “formal” duplicate records.¹⁵ In addition, I identified 202 further “effective” duplicates by systematically searching for events that occurred concurrently in the same location and manually inspecting them.¹⁶ Altogether, the events dataset includes 381 records for which duplicates exist in the original survey. Finally, I removed 2 incomplete records that had not been filled out using the original survey grid.¹⁷ I provide access to the entire set of records of the Jean Nicolas survey in the `nicolas_events_all` dataset, which includes the same variables as the `nicolas_events` dataset. In addition, this dataset links each duplicate record to the identifier of its corresponding reference record.

3.1.1. Typology

The first set of variables concerns the typology of rebellions. Originally developed to encompass 68 types grouped into 13 distinct categories (Nicolas, 1985a, pp. 15–16, 761–763), the typology was later extended to 72 types (Nicolas, 2002, pp. 26–27, 548–550).¹⁸ Each event is characterized

¹³ Each record is also associated with a unique `hiscod` identifier. See Section 4.1 for more details. Some rebellions are further linked to related events—up to three—by a note on the top-right corner of their original record—this is the case for 315 rebellion records.

¹⁴ See, e.g., Appendix Figure A9 for an example based on record 1048.

¹⁵ See, e.g., Appendix Figure A10 for an example based on record 0412.

¹⁶ In selecting which record to retain as the reference, I prioritized the one kept in the HiSCoD database. Note that reference records were not complemented with the content of their duplicate records, as they may draw from different—and potentially conflicting—sources.

¹⁷ See, e.g., Appendix Figure A11 for an example based on record 8780. See Appendix Table A2 for the distribution of these types of records.

¹⁸ Specifically, the types `Opposition to the actions of officers of the Eaux et Forêts jurisdiction (0312)` and `Hiring dispute (1108)` were added to the typology, while the type `Strike (1106)` was subdivided into three types: `Industrial sector strike (1106)`, `Agricultural sector strike`

by a primary general type and primary detailed type. In addition, 24 percent of events are further characterized by a secondary general type and a secondary detailed type. The general typology of rebellions is reported in Table 1.¹⁹ The most common type of rebellion is Resistance to state taxation, which accounts for 37 percent of events, followed by Resistance to the state judiciary, military, or police (18 percent), and Subsistence (17 percent).

Table 1 – General typology of rebellions.

Type of rebellion	Primary		Secondary		Valid
	Freq.	Percent	Freq.	Percent	
01 Rejection of state reform initiatives	52	0.61	3	0.04	0.15
02 Resistance to state taxation or incidental taxation	3,122	36.66	320	3.76	15.52
03 Resistance to the state judiciary, military, or police	1,497	17.58	796	9.35	38.60
04 Acts of hostility toward a seigneurial authority	428	5.03	113	1.33	5.48
05 Acts of hostility toward nobility or nobiliary privilege	13	0.15	12	0.14	0.58
06 Acts of hostility toward the Church	103	1.21	159	1.87	7.71
07 Accusation of a notable	157	1.84	70	0.82	3.39
08 Accusation of a municipal authority	148	1.74	111	1.30	5.38
09 Subsistence	1,483	17.41	186	2.18	9.02
010 Religion, beliefs	265	3.11	34	0.40	1.65
011 Labor dispute	434	5.10	61	0.72	2.96
012 Regional idiosyncrasy	92	1.08	46	0.54	2.23
013 Miscellaneous	722	8.48	151	1.77	7.32
Total	8,516	100.00	2,062	24.21	100.00
.c Type missing	0	0.00	6,454	75.79	
Total	8,516	100.00	8,516	100.00	

Notes. This table reports the distribution of the general primary and secondary types of rebellions in the `nicolas_events` dataset. It corresponds to the `type_prim` and `type_sec` variables.

3.1.2. Chronology

Date. The second set of variables concerns the chronology of rebellions, beginning with the date of each event. Dates are disaggregated into year, month, calendar day, and weekday variables. While the year of each rebellion is systematically available, complete dates are missing in 17 percent of cases.²⁰ Figure 1 displays the temporal distribution of recorded rebellions from 1661 to 1789.²¹ It reveals a relatively calm period following

(1109), and Commercial sector strike (1110). The translation of the typology from French to English was carried out by Nicole Charley.

¹⁹ The detailed typology is reported in Appendix Table A3.

²⁰ See Appendix Table A4 for the distribution of date variables.

²¹ Although the Jean Nicolas survey focuses on rebellions that occurred *before* the French Revolution—which began with the opening of the Estates General on May 5,

the upheavals of the 1620s–40s, from the *Cascadeu* in Provence (Pillorget, 1975) to the *Croquants* in Quercy (Bercé, 1974) and the *Nu-Pieds* in Normandy (Foisil, 1970). Nevertheless, a few spikes emerge in the mid-1670s, with the revolt of the Roure in Vivarais in 1670 (Ribon, 2001) and the revolts of the *Papier Timbré* in Brittany and Guyenne in 1675 (Aubert, 2014), followed by tumults due to the great famine of 1693–94. Other notable events include the “peasant rage” of the *Tard-Avisés* in Quercy in 1707 (Gossare, 1997) and the revolutionary outbreak of 1709, fueled by a subsistence crisis, the War of the Spanish Succession, and renewed Protestant opposition in the Massif Central (Aubert, 2023). From the 1730s onward, rebellious activity rises steadily, accelerating sharply during the 1760s and exceeding one hundred annual events following the Flour War of 1775 (Bouton, 1993), persisting through the end of the period and the run-up to the French Revolution.

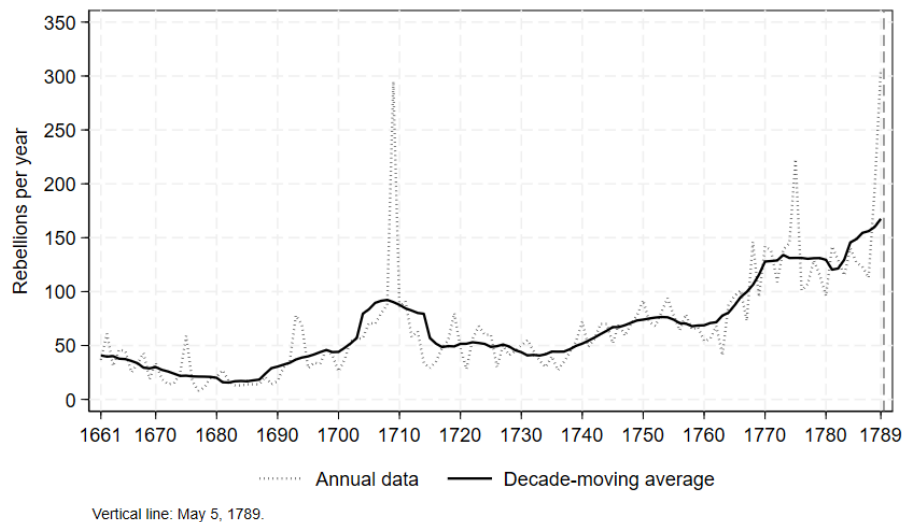


Figure 1 – Temporal Distribution of Rebellions.

Notes. This figure displays the temporal distribution of the 8,516 rebellions in the events dataset based on the `date_year` variable.

An examination of the monthly and daily distribution of rebellions indicates that neither the month of January nor the first day of the month is overrepresented, suggesting that recorded dates are reliable.²²

1789—two recorded events occurred *during* the Revolution on May 15 and June 14, 1789.

²² See Appendix Figure A12. In 262 cases, records specify a date interval, indicating that the event spanned multiple days. For these events, the date variables record both the start and end dates of the rebellion. An additional variable further measures

Period of the Day. The starting hour of the rebellion is available for 11 percent of events.²³ However, a broad period of the day—morning (6–12), afternoon (12–18), evening (18–22), and/or night (22–6)—is recorded for 48 percent of them.²⁴

In addition, information on the overall duration of the rebellion is available for 52 percent of cases across 5 categories: one hour, one to two hours, half a day, one day, or more than a day. Among the events for which duration information is available, 55 percent lasted less than two hours, suggesting that the Jean Nicolas database captures brief rebellious outbursts with notable precision.²⁵

3.1.3. Location

The third set of variables concerns the location of rebellions. It is organized into three subsets of variables: an Ancien Régime geography subset, which further includes population information, a contemporary geography subset, and a subset of miscellaneous spatial variables.

Ancien Régime geography. I manually match the location of each rebellion to a spatial reference system of 1789 parishes based on the *Histoire Administrative des Communes* (HAC) database. This database—which also serves as a reference framework for the TRF-GIS (Gay, 2021) and COMMUNE (Litvine et al., 2023) databases—is distributed via the website [cassini.ehess](https://cassini.ehess.fr/) (LaDéHiS, 2021; Motte & Vouloir, 2007) and is (partially) accessible in tabular format (Cristofoli et al., 2021). It contains individual entries for the forty thousand Ancien Régime parishes that appear on Cassini's *Carte générale de la France*—surveyed between 1756 and 1789 (Pelletier, 1990)—and that became municipalities in 1790, then communes in 1793 (see, e.g., Gorry, 2008).

Using the HAC database, I assign a location identifier (*cassini*) to each rebellion.²⁶ 96 percent of rebellions can be assigned to a specific parish

the duration of the rebellion in days. Conditional on lasting more than a single day, these events had an average duration of three days.

²³ See Appendix Table A5 for the distribution of hour variables. In 59 cases, an hour interval is recorded. For these events, the hour variable is duplicated to indicate the start and end hour of the rebellion. An additional variable further records the duration of the rebellion in hours. Conditional on an interval being indicated, these rebellions lasted an average of two hours. In 13 cases, the recorded hour contradicts the indicated period of the day. In such instances, I prioritize the period of the day.

²⁴ See Appendix Table A6 for the distribution of the period of the day variables.

²⁵ See Appendix Table A7 for the distribution of the duration variable.

²⁶ Eight rebellions occurred in the modern territories of Switzerland and Belgium and therefore fall outside of the coverage of the HAC database, which is limited to the territory of France as of 1999. These cases include rebellions that occurred in the

without ambiguity. For 31 cases, however, the original records do not provide a precise location. In such instances, I assign the rebellion to the *chef-lieu* of the relevant *canton*, *subdélégation*, or province cited in the original record.²⁷ An additional 267 rebellions cannot be precisely located because their records refer to *lieux-dits* or hamlets that did not become municipalities or communes, and are thus absent from the HAC database. In these cases, I assign the rebellion to the relevant village, town, or city mentioned alongside the *lieu-dit* or hamlet.²⁸ Finally, for 54 rebellions, the records refer to locations created after 1789. In these cases, I identify and assign the corresponding parent parish that existed in 1789.²⁹ A flag variable documents each of these ambiguous cases.³⁰

Notwithstanding the considerations discussed above, each rebellion is precisely situated in Ancien Régime geography through its `cassini` identifier and associated spatial coordinates.³¹ The 8,508 rebellions included in the events dataset with a `cassini` identifier span 4,264 distinct parishes (Figure 2). Three cities—Paris, Nantes, Bordeaux—each hosted over

Republic of Geneva (Collonge-Bellerive, Compesières, Corsier, Meinier, and Vézenaz) and in Mariembourg, which was incorporated into the Kingdom of the Netherlands under the Second Treaty of Paris in 1815, and into the Kingdom of Belgium following the Belgian Revolution of 1830.

²⁷ For instance, rebellion 5754 is recorded as having occurred in the canton of Toulouse without further precision. Accordingly, I assign this event to the *chef-lieu* of the canton, i.e., the parish of Toulouse (37818). See Appendix Figure A13. 9 uncertain cases further involved subjective choices—these are documented in Appendix Table A8.

²⁸ For instance, rebellion 0174 is recorded as having occurred in the hamlet of Les Fontêtes ([Fonteste]) near the village of Saint-Ours. I therefore assign it to the parish of Saint-Ours (33885). See Appendix Figures A14 and A15.

²⁹ For instance, rebellion 0177 is recorded as having occurred in La Sauvetat (35591), a commune established in 1872. I therefore assign it to its parent parish, i.e., Authezat (01964).

³⁰ See Appendix Table A9 for the distribution of this flag variable. Another ambiguity arises when rebellions are reported to have occurred in multiple locations: in seven cases, across multiple parishes (see, e.g., Appendix Figure A16), and in six cases, across an entire region, such as a *bailliage*, an *élection*, or a *généralité* (see, e.g., Appendix Figure A17). In the case of multiple parishes, I assign the most populous one. In the case of a regional rebellion, I assign the *chef-lieu* of the corresponding region. A flag variable documents these cases (see Appendix Table A10).

³¹ The spatial coordinates for parish locations are provided in both RGF93 and WGS84 projections. These coordinates are derived from the `position_1999` variable in Cristofoli et al. (2021)—which is itself based on IGN's BD-TOPO 1999—supplemented by the `position_cassini` variable for locations that were not communes in 1999.

100 events during the period, while 18 locations saw more than 20.³² However, the distribution of locations that ever hosted a rebellion is heavily left-skewed, as the median location hosted only a single event.³³

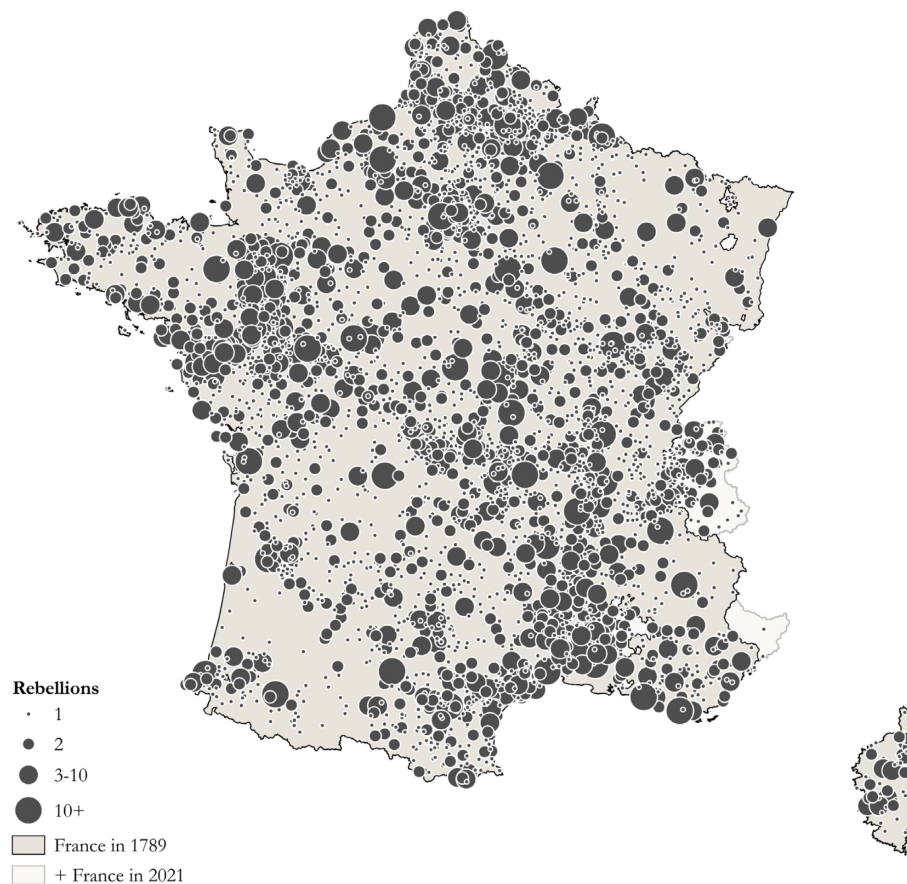


Figure 2 – Spatial Distribution of Rebellions.

Notes. This figure displays the spatial distribution of the 8,508 rebellions in the `nicolas_events` dataset that have a cassini identifier in RGF93 projection. It is based on the `lat_cassini_rgf` and `lon_cassini_rgf` variables. The underlying shapefile of the Kingdom of France as of 1789 is based on Gay et al.'s (2023) jurisdictions shapefile (Gay et al., 2024). The underlying shapefile of France as of 2021 is based on IGN's (2021) ADMIN-EXPRESS shapefile.

³² See Appendix Table A11 for the distribution of the most frequent locations.

³³ See Appendix Figure A18.

Each location is further characterized by a set of historical toponyms: the name of the parish as it appears on Cassini's map, the name of the corresponding municipality as listed on the "Year III copy" of the 1793 population census, and the name of the corresponding commune as listed on the 1801 *Arrêtés de réduction des justices de paix*, which constitute the first official nomenclature of administrative constituencies.³⁴

To characterize the nature of rebellion locations within the broader landscape, I include a variable indicating their representation on Cassini's map: as a church tower (parish), a hamlet, a town, a city, an abbey, or a castle.³⁵ According to this measure, 43 percent of rebellions occurred in parishes, 10 percent in towns, and 45 percent in cities.³⁶

I also include information on the administrative setting of each location as of 1789 based on Brette's (1904) historical atlas and Gay et al.'s (2024) corresponding dataset: its sovereignty, *généralité*, and *bailliage*. 97 percent of rebellion locations were situated within the Kingdom of France, with the remainder largely located in the Duchy of Savoy.³⁷

1793 and 1800 Population. To further characterize the type of locations that experienced rebellions, I include population information. I draw on the 1793 census for the corresponding municipality and on the 1800 census for the corresponding commune.³⁸ Note that population information for 1793 is available for 98 percent of rebellion locations, leaving about 200 cases without population data. There are three primary reasons for these gaps. The first is related to the operations of the 1793 census (Aberdam, 2004; Meuriot, 1918). This census recorded populations at the level of municipalities, which sometimes encompassed multiple parishes that would later be divided into independent communes. As a result, some parishes did not have distinct

³⁴ While parish names from Cassini's map and commune names from the 1801 *Arrêtés* are provided by the `nom_cassini` and the `nom_1801` variables in Cristofoli et al. (2021), 1793 municipality names are only accessible through the web version of the HAC database. The source of the `nom_an3` variable in Cristofoli et al. (2021) remains unclear.

³⁵ This classification is derived from the `graphie` variable in Cristofoli et al. (2021). Appendix Figure A19 illustrates the corresponding map symbols.

³⁶ See Appendix Table A12 for the distribution of this representation variable.

³⁷ See Appendix Table A13. I further include information on each location's *intendance* and *subdélégation* as of 1789 as well as *district* and *département* as of 1793 based on Cristofoli et al.'s (2021) dataset. This information is included as it appears in this dataset.

³⁸ These values are based on the `pop_an3_info`, `pop_an3_val`, `pop_an8_info`, and `pop_an8_val` variables in Cristofoli et al. (2021).

population counts.³⁹ Moreover, some municipalities were omitted from the census due to the disruptions caused by external wars and civil unrest during its operations. In some cases, original census returns were never recovered.⁴⁰ Finally, Corsica was not included in the 1793 census. Second, some parishes had been absorbed into neighboring ones that became municipalities between 1789 and 1793, and thus were not counted separately in 1793.⁴¹ Third, a number of parishes were only incorporated into France after 1793.⁴² For these ca. 200 cases, I assign the corresponding population from the 1800 census. Each of these cases is documented with a flag variable. A similar flag variable tracks the ca. 50 missing population values for the 1800 census.⁴³

Figure 3 displays the distribution of rebellions across location by 1793 population size. Rebellions were relatively evenly distributed among small, medium, and large municipalities, with a moderate concentration in urban municipalities with more than 5,000 inhabitants.⁴⁴ The median municipality that ever hosted a rebellion had approximately 1,700 inhabitants.⁴⁵

Contemporary geography. I also match each parish to its corresponding commune in 2021 geography based on INSEE's (2021) Official geographic code (COG) and IGN's (2021) ADMIN-EXPRESS, so that each location is associated with both an `insee` and an `ign` identifier.⁴⁶ I further include each commune's toponymy, spatial coordinates, and administrative setting—region, *département*, *arrondissement*, and *canton*. Because France's communal landscape consolidated over time (Bideau & Verdier, 2024),

³⁹ For instance, rebellion 0230 occurred in the parish of Saint-Cirgues (30943), which was grouped in with the parish of La Voutte (19040) in the municipality of La Voutte et Saint Cirgues during the 1793 census.

⁴⁰ This concerned the municipalities of five districts: La Roche-sur-Yon and Montaigu in the département of Vendée, La Rochefoucauld in the département of Charente, Machecoul in the département of Loire-Inférieure, Sarrelouis in the département of Moselle, and Valenciennes in the département of Nord.

⁴¹ For instance, rebellion 0397 occurred in the parish of Parilly (61165), which had been absorbed by Chinon (9378) prior to the 1793 census.

⁴² For instance, rebellion 0397 took place in Mulhouse (24395), which was reunited with France in 1798.

⁴³ See Appendix Tables A14 and A15 for the distribution of these flag variables.

⁴⁴ Appendix Figure A20 shows the corresponding distribution using 1800 population figures.

⁴⁵ See Appendix Table A43 for summary statistics of quantitative variables.

⁴⁶ The correspondence between `cassini` and `insee` identifiers is based on the `commune_mars_2021` variable in Cristofoli et al. (2021).

rebellions are distributed across fewer locations in contemporary than in Ancien Régime geography: 4,149 communes compared to 4,264 parishes.

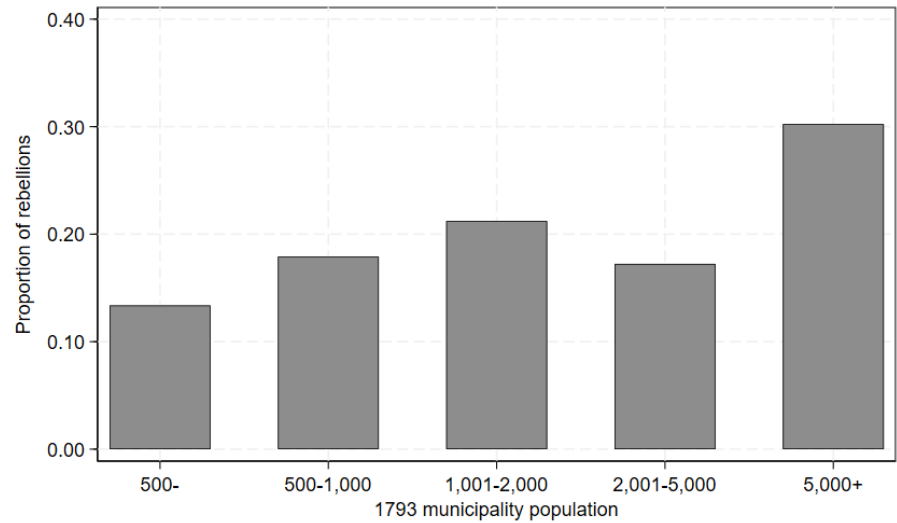


Figure 3 – Distribution of Rebellions by 1793-Municipality Population.

Notes. This figure displays the distribution of rebellions by 1793-municipality population for the 8,508 rebellions in the `nicolas_events` dataset that have a `cassini` identifier. It is based on the `pop_1793` variable.

Miscellaneous spatial variables. Besides their locations, original rebellion records further contain three spatial indicators: whether the location where the rebellion occurred was higher than 500 meters in altitude, closer than 30 kilometers from a border, and whether there was a strong Protestant presence in the area. I report the original altitude measure alongside the minimum, maximum, and *chef-lieu* altitude of the rebellion location based on contemporary geography.⁴⁷ I also report the original border-distance variable, along with a measure of the distance to the nearest sea or external

⁴⁷ See Appendix Tables A16 and A43. I match each 2021 commune in the data with France's Digital elevation model (MNT) provided by Sauvion (2024). The original altitude variable is relatively accurate as 99 percent of rebellions recorded as having occurred below 500 meters in the original record are confirmed by the MNT data, while 84 percent of those recorded above 500 meters are likewise confirmed (see Appendix Table A17).

land frontier, and the type of that frontier.⁴⁸ Finally, I report the original Protestant-presence variable.⁴⁹

3.1.4. Participant Characteristics

The fourth set of variables relates to the characteristics of rebellion participants. The events dataset records the number of participants in each event, a broad measure of its “intensity,” and the gender, age, and social categories of those involved. However, since this information is not available for all records, users of the Jean Nicolas database should assess the spatial and temporal representativeness of their sample when including these variables in their analysis—the same caution applies to the variables discussed in the following subsections.

Number of participants. The number of participants in a rebellion is recorded in various ways, depending on the level of detail provided by source materials: a precise number, an interval, a lower bound, or a relatively vague term. Still, information on rebellion size is entirely missing in 56 percent of cases. To make this information usable in empirical analyses when available, I apply the following transformations: when a participant count is provided as an interval (3 percent of cases), I use the midpoint of this interval; when a lower bound is provided (2 percent), I use this lower bound; and when an approximate number is provided (1 percent), I use the reported figure as-is. I document the nature of the original string provided through a flag variable.⁵⁰ I also retain the original string from rebellion records. For the 3,769 rebellions for which this information is available, the average number of participants is 164, and the median is 21. This distribution is heavily left-skewed.⁵¹ The Jean Nicolas database thus captures both small and large rebellious events.

Intensity. The intensity of a rebellion is an element written by Jean Nicolas in the lower left corner of the first page of each record. It is available for nearly all records (99.5 percent). Ranging from 1 to 3, it indicates the broad importance of the rebellion. As defined in his book (Nicolas, 2002, pp. 27–28), “weak” rebellions (*intensity*=1) are those with fewer than 10 individuals, “moderate” rebellions (*intensity*=2), those with 10–

⁴⁸ See Appendix Tables A18 and A43. These distance measures are calculated based on Gay et al.’s (2023) shapefile of the Kingdom of France as of 1789 (Gay et al., 2024).

⁴⁹ See Appendix Table A19 for the distribution of the Protestant-presence variable. Users wishing to validate this measure may consult Mours (1958) or Robert (1961).

⁵⁰ See Appendix Table A20 for the distribution of this flag variable. When the information provided is a vague term (31 cases), such as “foule,” I input the number of participants based on the intensity variable: 7 participants for weak rebellions, 30 for moderate ones, and 100 for strong ones. See below for more details.

⁵¹ See Appendix Figure A21 for the distribution of the number of participants.

50 individuals, and “strong” rebellions (*intensity* = 3), those with more than 50 individuals.⁵² However, as discussed above, information on the number of participants is missing in more than half of cases. Jean Nicolas was well aware of this issue.⁵³ His solution was to also rely on qualitative information available in the sources.⁵⁴ Per his assessment, 31 percent of rebellions had a weak intensity, 45 percent, a moderate intensity, and 24 percent, a strong intensity.⁵⁵

How accurate is Jean Nicolas’s assessment of the intensity of a rebellion? While we cannot assess his accuracy for all rebellions, it is possible to do so for the subset for which we have precise information on the number of participants, i.e., 38 percent of cases. Appendix Table A22 provides a tabulation of the intensity measure by participant count bin. Although the overlap between Jean Nicolas’s assessment and the actual number of participants is not perfect, it is reasonably accurate as 84 percent of cases are “correctly” classified $((792 + 979 + 958) / 3,254)$.

Gender and age composition. Rebellion records document the gender composition of rebellion participants in 76 percent of cases, across 8 categories: men and women; only men; mixed, majority men; only women; mixed, majority women; women and children; men, women, and children; and only children.⁵⁶ From these measures, I further create an indicator variable for whether women participated in the rebellion.⁵⁷ This was the case

⁵² “Nous avons affecté chaque cas d’un coefficient d’intensité relative en distinguant trois échelons: affaires de faible, de moyenne et de considérable importance, de 4 à 10 participants, de 11 à 50 et plus de 50” (“We have assigned a relative intensity coefficient to each case, distinguishing three levels: small, medium and large cases, from 4 to 10 participants, from 11 to 50 and above 50,” Nicolas, 2002, p. 27).

⁵³ “On sait que les sources restent souvent imprécises ou peu fiables pour les données chiffrées [sur les participants]” (“One knows that sources are often imprecise or unreliable when it comes to statistics [on participants],” Nicolas, 2002, p. 28).

⁵⁴ “Nous avons donc tenu compte de toutes les estimations qualitatives fournies par les autorités ou par des témoins” (“We have therefore taken into account all qualitative estimates provided by the authorities or by witnesses,” Nicolas, 2002, p. 28).

⁵⁵ See Appendix Table A21 for the distribution of the intensity variable. In the 21 cases for which the intensity variable is missing from the original record but information on the number of participants is available, I input the intensity based on this participant count.

⁵⁶ See Appendix Table A23 for the distribution of the participants variable.

⁵⁷ Specifically, this indicator variable takes the value of 1 if the gender composition of the rebellion is men and women; mixed, majority men; only women; mixed, majority women; women and children; or men, women, and children. Conversely, it takes the value of 0 if the gender composition of the rebellion is only men or only

for about half of the rebellions for which the gender composition of participants is documented.

Rebellion records also document the age composition of participants in 94 percent of cases, distinguishing between adults and children. 95 percent of rebellions had a majority of adults.⁵⁸

Social categories. The social categories of rioters are documented in 70 percent of cases, across 17 categories: farmers, market gardeners, winemakers, agricultural workers, lumberjacks, artisans, journeymen, industrial workers, domestic servants, soldiers, mariners, inmates, beggars, bohemians, migrants, outcasts, and notables.⁵⁹ Because a rebellion can involve rioters from multiple social categories, I create an indicator variable for each category that takes the value of 1 if that category is mentioned in the rebellion record. The most frequently mentioned social categories are farmers (28 percent of cases), notables (18 percent), artisans (13 percent), and outcasts (12 percent).⁶⁰

Involvement of notables alongside rioters. The involvement of notables alongside rioters is documented in 22 percent of cases, across 9 categories of notables: lords, nobles, civil officers or magistrates, lawyers or clerks, priests or members of a religious order, municipal representatives, wealthy commoners, tavern keepers, occult figures, and other categories.⁶¹ Because different types of notables can be involved alongside rioters, I create an indicator variable for each category that takes the value of 1 if that category is mentioned in the rebellion record. The most frequently mentioned categories of notables are municipal representatives (4 percent of cases), members of a religious order (4 percent), and wealthy commoners (4 percent).⁶²

children. It takes a missing value if the gender composition of the rebellion is missing. See Appendix Table A24 for the distribution of the gender variable.

⁵⁸ See Appendix Table A25 for the distribution of the age variable.

⁵⁹ The original terminology of social categories is *paysans*, *marais*, *vignerons*, *ouvriers agricoles*, *bûcherons*, *artisans*, *compagnons*, *ouvriers industriels*, *domestiques*, *soldats*, *marins*, *détenus*, *mendiants*, *bohémiens*, *migrants*, *marginiaux*, and *notables*.

⁶⁰ See Appendix Table A26 for the distribution of the social categories variable.

⁶¹ The original terminology of notables is *seigneur*, *noble*, *officier civil* or *magistrat*, *gens de loi* or *avocat*, *curé* or *vicar* or *religieux*, *représentant du pouvoir municipal*, *roturier fortuné*, *cabaretier*, *personnage occulte*, and *autres*.

⁶² See Appendix Table A27 for the distribution of the notables involvement variable.

3.1.5. Forms of Confrontation

The fifth set of variables concerns the forms of confrontation during rebellions: the modes of expression of rioters and whether they used weapons or uttered insults. Because these elements are composed of several not mutually exclusive categories, I create an indicator variable for each category that takes the value of 1 if that category is mentioned in the rebellion record.

Modes of expression. Rebellion records provide (limited) information on the modes of expression of rioters, across 4 categories: whether they used a tocsin, masks, cross-dressing or disguises, or musical instruments. Among the 7 percent of records that mention at least one mode of expression, the most frequently mentioned is the use of a tocsin (4 percent of cases).⁶³

Weapons. The types of weapons used by rioters are documented in 60 percent of cases, across 7 categories: rocks, sticks or bars, tools, knives, firearms, other weapons, and no weapons. The most frequently mentioned types of weapons are sticks or bars (29 percent), rocks (21 percent), and knives (17 percent).⁶⁴

Insults. The types of insults uttered by rioters are documented in 31 percent of cases, across 4 categories: sexual slurs, social slurs, slurs with double reference, and threats of bloodshed. The most frequently mentioned type of insult is threats of bloodshed (23 percent of cases).⁶⁵

3.1.6. Forms of Violence

The sixth set of variables concerns the forms of violence in rebellions: against rioters, their adversaries, goods, and buildings. As before, these elements are composed of several not mutually exclusive categories, so I create an indicator variable for each category that takes the value of 1 if that category is mentioned in the rebellion record.

Violence against rioters. The forms of violence against rioters are documented in 35 percent of cases, across 6 categories: whether some rioters were wounded, killed, captured, captured then released, arrested after the confrontation, or imprisoned. The most frequently mentioned form of violence is the capture of rioters during the confrontation (12 percent of cases). In addition, rebellion records sometimes include information on the number of rioters affected by each form of violence.⁶⁶

⁶³ See Appendix Table A28 for the distribution of the modes of expression variable.

⁶⁴ See Appendix Table A29 for the distribution of the weapons variable.

⁶⁵ See Appendix Table A30 for the distribution of the insults variable.

⁶⁶ See Appendix Tables A31 and A32 for the distributions of the forms of violence against rioters variables.

Violence against rioters' adversaries. The forms of violence against rioters' adversaries are documented in 49 percent of cases, across 9 categories: whether some adversaries were disheveled, had their clothing torn, were disarmed, detained, run off, wounded, killed, whether their corpses were desecrated, and whether there were exchanges of blows. The most frequently mentioned forms of violence are that rioters' adversaries were run off (22 percent of cases), wounded (20 percent), and subject to exchanges of blows (12 percent). In addition, rebellion records sometimes include information on the number of rioters' adversaries affected by each form of violence.⁶⁷

Violence against goods. The forms of violence against goods are documented in 23 percent of cases, across 17 categories: whether documents were burnt, stolen, or torn, whether posters were torn; whether buildings were stoned, invaded, demolished, or burnt, whether fences were broken or ditches filled, whether wood was looted from the forest, whether harvests were damaged, livestock attacked, and whether merchandise was misappropriated, looted, destroyed, burnt, or taxed. The most frequently mentioned form of violence against goods is the looting of merchandise (8 percent of cases).⁶⁸

Types of buildings attacked. Finally, the types of buildings attacked by rioters are documented in 15 percent of cases, across 5 categories: seats of a public authority, seigneurial or nobiliary authority, ecclesiastical authority, and the homes of a local officer or notable. The most frequently mentioned type of building attacked is the home of a notable (7 percent of cases).⁶⁹

3.1.7. Legal Consequences

The seventh set of variables concerns the legal consequences of rebellions: the jurisdictions that handled each case and the nature of sentencing. As before, these elements are composed of several not mutually exclusive categories, so I create an indicator variable for each category that takes the value of 1 if that category is mentioned in the rebellion record.

Jurisdictions. The jurisdictions that handled the rebellion's legal consequences are documented in 40 percent of cases, across 4 categories: simple seigneurial or royal jurisdictions, *bailliages*, *sénéchaussées*, or *présidiaux*, Parliaments or other sovereign courts, *prévôtés* or *maréchaussées*. The most frequently mentioned types of jurisdictions are

⁶⁷ See Appendix Tables A33 and A34 for the distributions of the forms of violence against adversaries variables.

⁶⁸ See Appendix Tables A35 and A36 for the distribution of the violence against goods variable.

⁶⁹ See Appendix Table A37 for the distribution of the types of buildings attacked variable.

prévôtés or *maréchaussées* (13 percent of cases), *bailliages*, *sénéchaussées*, or *présidiaux* (12 percent), and Parliaments or other sovereign courts (10 percent).⁷⁰

Sentencing. The forms of sentencing for rioters are documented in 19 percent of cases, across 9 categories: sentencing to breaking on the wheel, to hanging, to the galleys (either for life or for a fixed term), to banishment, to the pillory or whip, to a fine, to a reprimand, or to prison. The most frequently mentioned type of sentencing is a fine (7 percent of cases). In addition, rebellion records sometimes include information on the number of rioters affected by each form of sentencing.⁷¹

3.1.8. Sources

The events dataset contains a few variables extracted from the sources dataset, which provides detailed information on the sources used to create rebellion records (see Section 3.2): the number of sources used to create a given rebellion record, also broken down between archival and bibliographic sources, and indicator variables that take the value of 1 if a record cites at least one bibliographic source, one archival source, one national archival source, or one local archival source.⁷² All but six rebellion records have at least one source mentioned. On average, rebellion records cite 1.3 sources, with archival sources being four times more frequent than bibliographic ones. In addition, 87 percent of records mention at least one archival source—with 47 percent mentioning at least one national archival source and 42 percent, at least one local archival source—and 21 percent, at least one bibliographic source.

Based on the source composition of each record, I create a département-level quality grade variable that characterizes each record—the methodology of this grading scheme is detailed in Section 4.2.4. Grades range from higher quality (grade A) to lower quality (grade D). They are available for all records that describe rebellions that occurred on the current territory of France, i.e., all but 9 records. About one third of records have the grade of A (34 percent),

⁷⁰ See Appendix Table A38 for the distribution of the jurisdictions variable.

⁷¹ See Appendix Tables A39 and A42 for the distributions of the sentencing variables.

⁷² See Appendix Table A43 and A44. National archival sources are those from the National archives, the Bibliothèque nationale de France, the Service historique de la Défense, the archives of the préfecture of Police of Paris, the central archives of the Marine, the diplomatic archives, and the library of the Assemblée nationale. Local archival sources are those from départemental archives, municipal archives, municipal libraries, the Archivio di Stato di Torino, and the Bibliothèque historique de la Ville de Paris.

another third, the grade of B (34 percent), and the rest, the grades of C (18 percent) and D (14 percent).⁷³

3.1.9. Authorship

The events dataset provides a few variables extracted from the authors dataset, which provides detailed information on the researchers who created rebellion records (see Section 3.3): the identifier and name of the author of each record. A total of 64 authors contributed to the Jean Nicolas survey. Jean Nicolas himself produced 5,200 records (61 percent) of the events dataset. Several other authors also made substantial contributions: Sylvie Kleiber produced 351 records (4 percent), Elie Pélaquier, 309 records (4 percent), and Francis Loirette, 248 records (3 percent). In total, 16 authors contributed at least 70 records.⁷⁴

3.1.10. Description of Events

The final and tenth set of variables provides a description of the event in two ways: first, the terms used by contemporaries to characterize the rebellion—available under the label “Movement’s characterization by contemporaries” (“Qualification du mouvement par les contemporains”)—and second, a textual description—available under the label “Other information” (“Autres informations”).

Terminology used by contemporaries. The terms used by contemporaries are not mutually exclusive, so I create an indicator variable for each category that takes the value of 1 if that term is mentioned in the rebellion record. A terminology is documented in 54 percent of cases, with an average of 1.7 terms. I report the mention of 27 different terms that appear in at least 10 records: “rebellion,” “gathering,” “sedition,” “emotion,” “riot,” “tumult,” “disorder,” “excess,” “revolt,” “troubles,” “mutiny,” “assembly,” “violence,” “uprising,” “movement,” “carillon,” “rumor,” “cabal,” “conspiracy,” “noise,” “murmurs,” “tapage,” “brawl,” “insurrection,” “battery,” “hullabaloo,” and “effervescence.”⁷⁵ The most frequently mentioned terms are “rebellion”

⁷³ See Appendix Table A54 for the distribution of the grade variable. Appendix Figure A26 provides its spatial distribution.

⁷⁴ See Appendix Table A45 for the distribution of the authorship variable. The events dataset also provides the identifier of the research assistant who input the content of each record.

⁷⁵ The original terminology of these terms is rébellion, attroupement, sédition, émotion, émeute, tumulte, désordres, excès, révolte, troubles, mutinerie, assemblée, violences, soulèvement, mouvement, carillon, rumeur, cabale, complot, bruit, murmures, tapage, rixe, insurrection, batterie, charivari, and effervescence. I do not include the terms affrontements, rassemblement, révolution, agitation, commotion, crieries, fureur, or prise d’armes, since they appear in less than 10 records.

(21 percent of records), “gathering” (11 percent), “sedition” (9 percent), “emotion” (9 percent), and “riot” (8 percent).⁷⁶

Textual description. The textual description of events is available for 7,974 records (94 percent).⁷⁷ It is provided in the form of a long string, up to 9,491 characters long.

3.2. The Sources Dataset

The sources dataset (`nicolas_sources`) provides detailed information on the sources of records in the events dataset. It is a record-source-level dataset with 10,747 observations, where each source has a unique source identifier. This dataset can be matched to the events dataset through the `nicolas` identifiers. Each source is classified either as an archival source (80 percent of cases) or a bibliographic source (20 percent). The original source strings as transcribed from rebellion records are also included.⁷⁸ The 22 variables included in the sources dataset are listed in Panel B of Appendix Table A1.

3.2.1. Archival Sources

Archival sources are further classified across 12 holding services—this information is available for all but 22 archival sources.⁷⁹ Archival sources are drawn primarily from the National (43 percent of archival sources) and départemental (40 percent) archives. I also provide the main archival series of sources from National and départemental archives.⁸⁰ Among these, 28 percent are from the B series of départemental archives (Courts and jurisdictions), 18 percent, from the Z1A series of the National archives (Court of Aids of Paris), 17 percent, from the C series of départemental archives (Provincial administrations), and 13 percent, from the G7 series of the National archives (General control of finances). Finally, I provide information on the location of each archive holding service in 2021 geography, including its `insee` and `ign` identifiers, commune’s toponymy, spatial coordinates,

⁷⁶ See Appendix Tables A40 and A41 for the distribution of the terms used by contemporaries.

⁷⁷ These descriptions were entered by Renan Donnerh, the manager of the François Lebrun Library at the University of Rennes-2 where the Jean Nicolas records are preserved.

⁷⁸ These strings were also entered by Renan Donnerh. Links to the original sources can be found in the `url` variables in the HiSCoD database (Chambru & Maneuvrier-Hervieu, 2024).

⁷⁹ See Appendix Table A46 for the distribution of archival services.

⁸⁰ See Appendix Table A47 for the distribution of archival series from National and départemental archives.

and administrative setting—region, *département*, and *arrondissement* (see Section 4.2).

3.2.2. Bibliographic Sources

Bibliographic sources are classified into 4 categories: whether they correspond to books (69 percent of bibliographic sources), journal articles, book chapters, or conference proceedings (19 percent), master's theses (9 percent), or dissertations (4 percent).

3.3. The Authors Dataset

The authors dataset (*nicolas_authors*) provides detailed information on the 64 authors of records in the events dataset. Based on the content of Jean Nicolas's logbook, it includes each author's minimum, maximum, and average year of collaboration on the Jean Nicolas survey (available for 49 authors), as well as their professional location during this period in 2021 geography, with each author's location *insee* and *ign* identifier, commune's toponymy, spatial coordinates, and administrative setting—region and *département* (see Section 4.2). The 19 variables included in the authors dataset are listed in Panel C of Appendix Table A1.

4. RELIABILITY OF THE JEAN NICOLAS DATABASE

This section evaluates the reliability of the Jean Nicolas database. First, I compare the content of the database with the entries of the HiSCoD database (Section 4.1). Then, I conduct a quantitative source criticism exercise, analyzing potential biases introduced by the construction methodology of the Jean Nicolas survey (Section 4.2). I also provide recommendations on how to mitigate these biases for an empirical exploitation of the Jean Nicolas database.

4.1. Comparison with the HiSCoD Database

The Historical Social Conflict Database (HiSCoD) comprises about 21 thousand episodes of social conflict in Europe, spanning from the High Middle Ages to the late nineteenth century (Chambru & Maneuvrier-Hervieu, 2024). The Jean Nicolas survey is HiSCoD's most substantial data source, contributing 8,478 events—about 40 percent of the events in the HiSCoD database, and 54 percent of those located in France.^{81,82} While all but four

⁸¹ This portion of the HiSCoD database was derived from Jean Nicolas's personal data file, which he used to produce his book (Nicolas, 2002) and shared with Cédric Chambru and Paul Maneuvrier-Hervieu in 2012. I am grateful to Cédric Chambru for bringing this information to my attention.

⁸² Although Chambru and Maneuvrier-Hervieu (2024, Table 1, p. 1085) report that HiSCoD includes 8,528 events based on the Jean Nicolas survey, the corresponding

Jean Nicolas records in HiSCoD are included in the `nicolas_events` dataset, 42 events included in the `nicolas_events` dataset are absent from HiSCoD.⁸³ In addition, HiSCoD includes only a limited subset of variables for each rebellion—their typology, date, location, number and gender of participants, sources, and authorship. By contrast, the `nicolas_events` dataset retains the full set of variables available in each rebellion record, together with auxiliary data such as location population counts. An additional difference is that rebellion locations in the `nicolas_events` dataset are situated in both Ancien Régime and contemporary geographies, whereas HiSCoD locates them only in contemporary geography.

To assess the validity of the Jean Nicolas database, I compare its content with that of HiSCoD for the 8,477 rebellion records that are included in both databases.⁸⁴ For each record, I match all variables that are common to both datasets and manually check the relevant original record whenever a discrepancy arises. The results of this comparison are reported in Table 2. This analysis draws on the `nicolas_hiscod` dataset, which includes the original variables from both HiSCoD and `nicolas_events` datasets, as well as flag variables indicating discrepancies—the 41 variables included in the HiSCoD comparison dataset are listed in Panel D of Appendix Table A1. Overall, 1,619 HiSCoD records (19 percent) contain at least one error.

First, 7 records contain incorrect record identifiers, typically due to typographical errors—such as entering “7774” instead of “7744.” Second, the typology variables exhibit a substantial number of errors: 469 records (6 percent) have errors in their primary types and 1,152 records (14 percent), in their secondary types. In total, 1,268 HiSCoD records (15 percent) contain at least one error in their typology.⁸⁵ Third, 123 records include at least one

CSV file `db_hiscod_csv_v1_fr`, accessed from github.com/hiscod/hiscod-project in November 2025, contains only 8,478 such events.

⁸³ Of the four records present in HiSCoD but not in the `nicolas_events` dataset, one is an “effective” duplicate (4291), two are incomplete records (8780 and 8781), and one could not be located (8763). The reason for the absence of 42 records from HiSCoD remains unclear, as they do not appear to be duplicates.

⁸⁴ This comparison includes the “effective” duplicate record (4291) and the two incomplete records (8780 and 8781), but excludes the record that could not be located (8763).

⁸⁵ Note that HiSCoD also contains an alternative typology along nine types (Chambru & Maneuvrier-Hervieu, 2024, Table 2, p. 1086). However, Chambru and Maneuvrier-Hervieu (2024, p. 1086) “highlight here that [their] classification is first and foremost intended as a tool to guide users through the myriad of events reported. One can use the various information included in the database to establish her own typology and/or assign several categories to an event.”

error in the date variables, most of which concern the day component.⁸⁶ Fourth, 35 records present errors in the rebellion locations—an additional 37 records display further discrepancies between the two databases, e.g., due to uncertainty regarding the original location or because the referenced location was created after 1793.⁸⁷ Fifth, 98 records contain inaccuracies in the number of participants—an additional 257 records further display discrepancies between the two databases due to differences in the interpretation of participation counts, e.g., when reported as intervals rather than precise numbers.⁸⁸ Sixth, 152 records include errors regarding the gender of participants.⁸⁹ Finally, 30 records have incorrect author information—an additional 91 records further display discrepancies between the two databases due to typographical errors in author names, e.g., “Serge Dotenwill” instead of “Serge Dontenwill” or “Nicole Postel-Pellegrin” instead of “Nicole Pellegrin-Postel.”

Table 2 – Comparison with the HiSCoD Database.

	Variable names		Errors in HiSCoD	
	nicolas_events	HiSCoD	Freq.	Percent
Nicolas identifier	nicolas	id_riot_original_database	7	0.08
Primary type	type_prim_det	riot_type_original_database_1	469	5.53
Secondary type	type_sec_det	riot_type_original_database_2	1,152	13.59
Date	date	year, month_num, day	123	1.45
Location (2021 geography)	insee_2021	city_code	35	0.41
Number of participants	part_nb	nb_participants	98	1.16
Gender of participants	part_women	women_participation	152	1.79
Author	author	author	30	0.35
At least one error in HiSCoD			1,619	19.10
Total			8,477	100.00

Notes. This table reports the distribution of errors in HiSCoD for the overlapping set of events and variables in the `nicolas_events` dataset. It corresponds to the `nicolas_hiscod`, `type_prim_det_hiscod`, `type_sec_det_hiscod`, `date_hiscod`, `location_hiscod`, `part_nb_hiscod`, `part_women_hiscod`, `author_hiscod`, and `error_hiscod` variables of the `nicolas_hiscod` dataset.

4.2.Biases in the Jean Nicolas Survey

A key concern is whether rebellions included in the Jean Nicolas survey can be considered as representative both temporally (Figure 1) and spatially

⁸⁶ See Appendix Table A48 for the distribution of the date comparison variable.

⁸⁷ See Appendix Table A49 for the distribution of the location comparison variable.

⁸⁸ See Appendix Table A50 for the distribution of the number of participants comparison variable.

⁸⁹ This variable is constructed following the method described in Chambru and Maneuvrier-Hervieu (2024, Appendix, pp. 4–5). See Footnote 57.

(Figure 2). Several potential biases remain challenging to assess without extensive historical and archival research, particularly those stemming from the destruction, gaps, and reliability of archival sources (Nicolas, 2002, p. 15). Additionally, the observed upward trend in rebellions starting in the 1740s might reflect changes in state actors' recording practices or their enhanced capacity for repression associated with state consolidation during this period (see, e.g., Albertus & Gay, 2025).

Nevertheless, certain biases resulting from the survey methodology can be systematically evaluated. This section focuses specifically on three types of biases: those related to the selection of sources used to compile rebellion records (Section 4.2.1), those associated with the authorship of these records (Section 4.2.2), and those stemming from the incomplete examination of local archives (Section 4.2.3). After systematically exploring these biases and recommending strategies to mitigate them, this section introduces a département-level grading scheme that summarizes the quality of rebellion records in the Jean Nicolas database (Section 4.2.4).

4.2.1. Source-Driven Biases

Rebellion records draw on a wide range of sources, each with distinct characteristics that may introduce systematic biases into the temporal, spatial, and typological distribution of recorded events. I examine three primary dimensions of source variation: first, the difference between archival and bibliographic sources; second, the distinction between national and local archival materials; and third, the influence of specific recording actors, reflected in the sources' archival series.

Archival and bibliographic sources. The broad type of sources used to compile rebellion records may influence their temporal and spatial distribution. Indeed, records based exclusively on archival sources display time patterns similar to those observed in the full dataset (Figure 4). By contrast, records based exclusively on bibliographic sources show notable differences. In particular, they fail to capture the 1709 surge in rebellion activity and the upward trends in rebellion beginning in the 1740s.

Similarly, the spatial distribution of rebellions based on archival sources closely mirrors that of the full dataset (Panel a of Figure 5). By contrast, rebellions drawn from bibliographic sources exhibit much more clustered spatial patterns, likely reflecting their reliance on regional monographs (Panel b of Figure 5). For instance, the cluster of approximately 60 rebellions in the Cévennes region—located in the southeast of the Massif Central—is largely attributable to the reliance on Bosc's (1985–93) comprehensive study of the War of the Camisards.

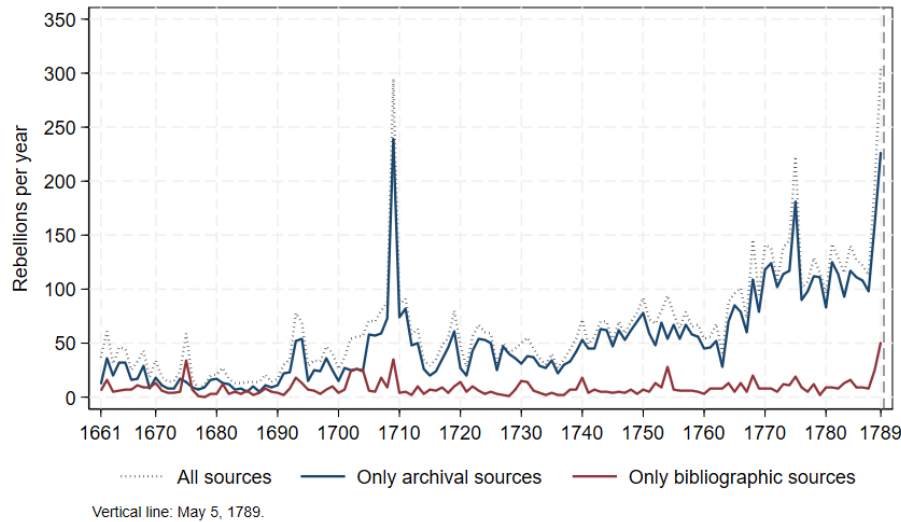


Figure 4 – Temporal Distribution of Rebellions by Type of Sources.

Notes. This figure displays the temporal distribution of rebellions by type of sources in the events dataset based on the `date_year` variable. The black dotted line includes the 8,510 rebellions with at least one source mentioned; the blue line, the 6,718 rebellions with exclusively archival sources; and the red line, the 1,113 rebellions with exclusively bibliographic sources.

National and local archival sources. The origin of archival sources—whether national or local—can also affect the temporal and spatial distribution of rebellions, as these sources may capture events of different scales or reflect the perspectives of distinct sets of actors. Reassuringly, both types of records exhibit time patterns closely aligned with those of the full dataset—though the 1709 surge in rebellion activity is not so pronounced for records drawing on local archival sources (Figure 6).

Nevertheless, the spatial distribution of rebellions differs somewhat between the two types of archival sources. National archival sources capture rebellions relatively evenly across the territory, particularly along the boundaries of the *Grandes Gabelles* frontiers in the northwest and in the north (Panel a of Figure 7).⁹⁰ This broader coverage is likely due to the archives of the *Ferme Générale*—the institution that managed the *gabelle* salt taxation—being preserved in the National Archives under the archival series of the Court of Aids (*AN-Z1A*). By contrast, local archival sources reveal more clustered patterns: rebellion records are for instance relatively

⁹⁰ Rebellions primarily motivated by salt smuggling—type 0204—account for 24 percent of all recorded rebellions (2,065 cases) in the events dataset (Appendix Table A3). Among these, 76 percent rely on sources drawn from national archival sources, compared to only 18 percent from local archival sources.

dense in southern Brittany and eastern Languedoc but essentially absent in northeastern France (Panel b of Figure 7). As discussed in Section 4.2.2, such patterns of rebellions drawing on local archival sources are likely driven by the regional stratification of collaborators to the Jean Nicolas survey.

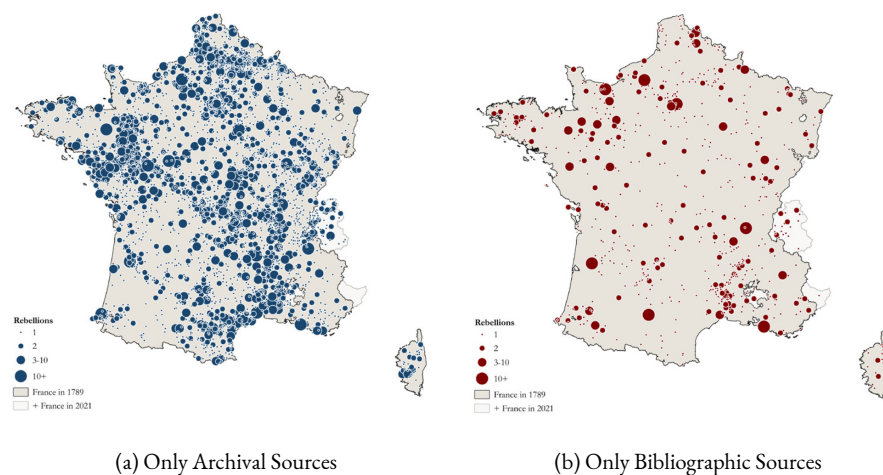


Figure 5 – Spatial Distribution of Rebellions by Type of Sources.

Notes. This figure displays the spatial distribution of rebellions by type of sources in RGF93 projection. Panel a includes the 6,718 rebellions with exclusively archival sources, and Panel b, the 1,113 rebellions with exclusively bibliographic sources. The underlying shapefile of the Kingdom of France as of 1789 is based on Gay et al.'s (2023) jurisdictions shapefile (Gay et al., 2024). The underlying shapefile of France as of 2021 is based on IGN's (2021) ADMIN-EXPRESS shapefile.

Types of archival series and recording actors. A third type of source-driven bias may stem from the over-representation of some archival series—reflecting different actors with heterogeneous recording practices and repression capacity—in the documentation of certain types of rebellions. To highlight this potential source of bias, I focus on the three main types of rebellions: rebellions linked to resistance to state taxation (3,380 cases), those linked to resistance to the state judiciary, military, or police (2,242 cases), and those linked to subsistence motives (1,510 cases).⁹¹ I find that some archival series account for a disproportionate share of rebellion records within these types.⁹² For instance, the (national) archives of the Court

⁹¹ These figures are larger than those displayed in either column of Table 1 because they comprise both primary and secondary types. Moreover, they only concern rebellions that mention at least one archival source.

⁹² See Appendix Table A51 for the distribution of the main archival series type of rebellion.

of Aids (AN-Z1A) account for one third of rebellions involving resistance to state taxation. Likewise, the (départemental) archives of courts and jurisdictions (AD-B) account for one quarter of rebellions involving resistance to the state judiciary, military, or police.

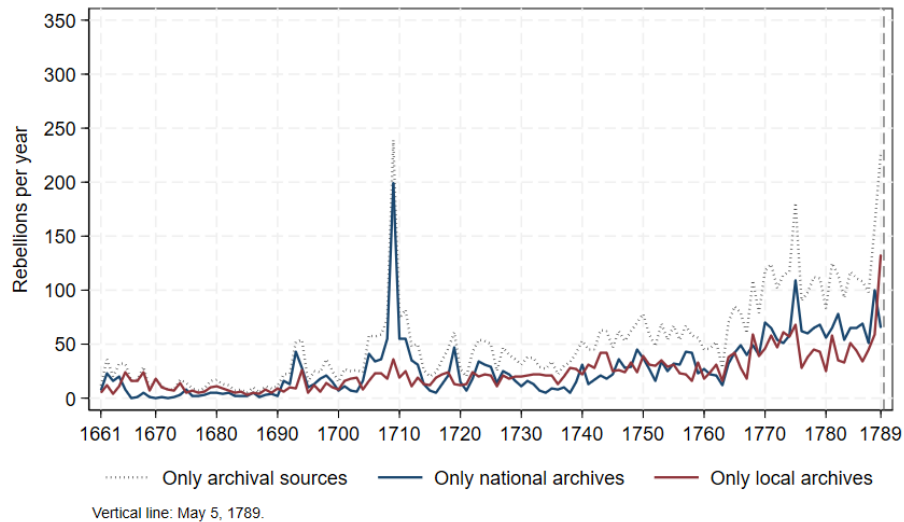


Figure 6 – Temporal Distribution of Rebellions by Type of Archival Sources.

Notes. This figure displays the temporal distribution of rebellions by type of archival sources in the events dataset based on the `date_year` variable. The black dotted line includes the 6,718 rebellions with exclusively archival sources; the blue line, the 3,556 rebellions with exclusively national archival sources; and the red line, the 3,045 rebellions with exclusively local archival sources.

Given these considerations, researchers using the Jean Nicolas database should carefully assess the robustness of their results to these potential biases. Proper guidelines include excluding rebellion records that rely exclusively on bibliographic sources, contrasting results across rebellion records that draw on national and local archival sources, and ensuring that findings are not disproportionately driven by any single archival series.

4.2.2. Author-Driven Biases

Another set of biases may arise from the authorship of rebellion records. I first examine differences in rebellion patterns across records authored by Jean Nicolas compared to his collaborators, then show that authors' workplace locations may introduce biases in the spatial distribution of recorded rebellions.

Jean Nicolas and his collaborators. While Jean Nicolas personally authored 61 percent of the records in the survey (5,200 cases), the remaining entries were contributed by 63 different collaborators (3,316 cases). Both

types of records exhibit similar time patterns, aligned with those of the full dataset—though rebellion records authored by Jean Nicolas’s collaborators exhibit a dip in the 1780s. However, the spatial distribution of rebellion records authored by Jean Nicolas differs from that of records authored by his collaborators.⁹³ Specifically, the patterns of rebellion records authored by Jean Nicolas resemble a combination of the patterns seen in records based on bibliographic and national archival sources, while those authored by his collaborators align more closely with records based on local archival sources.

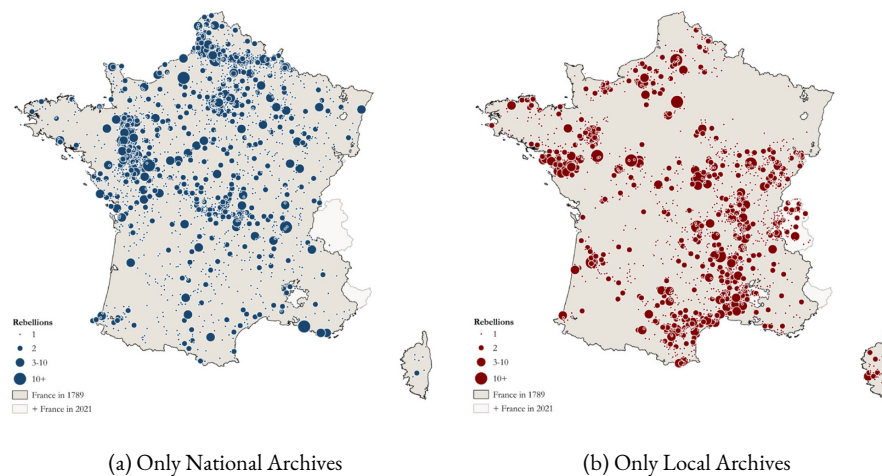


Figure 7 – Spatial Distribution of Rebellions by Type of Archival Sources.

Notes. This figure displays the spatial distribution of rebellions by type of archival sources in RGF93 projection. Panel a includes the 3,556 rebellions with exclusively national archival sources, and Panel b, the 3,045 rebellions with exclusively local archival sources. The underlying shapefile of the Kingdom of France as of 1789 is based on Gay et al.’s (2023) jurisdictions shapefile (Gay et al., 2024). The underlying shapefile of France as of 2021 is based on IGN’s (2021) ADMIN-EXPRESS shapefile.

These discrepancies can be explained by the types of sources used by Jean Nicolas compared to his collaborators. Jean Nicolas drew on both archival and bibliographic sources: among the records he authored, 81 percent cite at least one archival source, and 26 percent, at least one bibliographic source. By contrast, his collaborators relied more heavily on archival sources. Of the records they authored, 96 percent cite at least one archival source, while only 13 percent cite at least one bibliographic source. In addition, the archival material used by Jean Nicolas was primarily located in institutions near his home university in Paris: 69 percent of the records he

⁹³ See Appendix Figures A22 and A23 for the temporal and spatial distributions of authorship.

authored cite sources from the National archives, and 9 percent, sources from the Bibliothèque nationale de France. By contrast, archival sources of his collaborators were predominantly local, with 71 percent of their archival-based records drawing on archival services located within the département of their place of work—a figure that rises to 92 percent when considering archival services located within their broader region.⁹⁴

Hence, to alleviate this potential source of bias, researchers using the Jean Nicolas database should assess the robustness of their results by contrasting them across rebellion records authored by Jean Nicolas and his collaborators.

Regional stratification of the survey. Another type of bias may arise from the regional stratification of the Jean Nicolas survey. As discussed in Section 2.1.2, Jean Nicolas recruited his collaborators to cover specific geographic areas—generally, the département of their places of work. This recruitment process led to a pattern of spatial proximity between the workplace locations of these collaborators and the rebellion records they authored (Panel a of Figure 8).⁹⁵ Notable gaps in coverage are evident, similar to those highlighted in Section 4.2.1. For instance, the Grand Est region—highlighted in the northeast of the map—is strikingly underrepresented in the set of records collected by Jean Nicolas’s collaborators.⁹⁶ This gap does not necessarily imply a lack of rebellious activity in the region (see, e.g., Bischoff, 2009; Cabourdin, 1977; Gallet, 2009), but rather reflects the fact that the local archives of this region were seldom explored because Jean Nicolas was largely unsuccessful in recruiting collaborators in this area—this was not for a lack of effort on the part of Jean Nicolas, as his logbook reveals that he made 83 unsuccessful attempts at recruiting local collaborators throughout his survey. In the case of the Grand Est region, he reached out to 14 potential collaborators but only received two affirmative responses. The distribution of these unsuccessful contact attempts documented in his logbook is shown in Panel b of Figure 8.

This intuition is supported by a statistical analysis. Specifically, I examine whether the likelihood of observing a rebellion in a parish increases with proximity to the nearest collaborator of Jean Nicolas—I use the framework

⁹⁴ Excluding sources drawn from the National archives and the Bibliothèque nationale de France, these proportions increase to 76 and 94 percent, respectively.

⁹⁵ This map only shows rebellion records drawing on local archival sources and authored by Jean Nicolas’s collaborators.

⁹⁶ This region comprises 10 départements: Ardennes, Aube, Marne, Haute-Marne, Meurthe-et-Moselle, Meuse, Moselle, Bas-Rhin, Haut-Rhin, and Vosges. Only 4 rebellion records authored by Jean Nicolas’s collaborators drawing exclusively on local archival sources fall in this area, for a total of 419 rebellion records (5 percent) in the full dataset.

of Ancien Régime parishes. As shown in Figure 9, the probability of observing a rebellion decreases linearly with distance from the nearest collaborator. Focusing on rebellion records authored by collaborators and drawing exclusively on local archival sources—represented by red dots—the likelihood of observing a rebellion approaches zero when no collaborator is located within 75 kilometers.

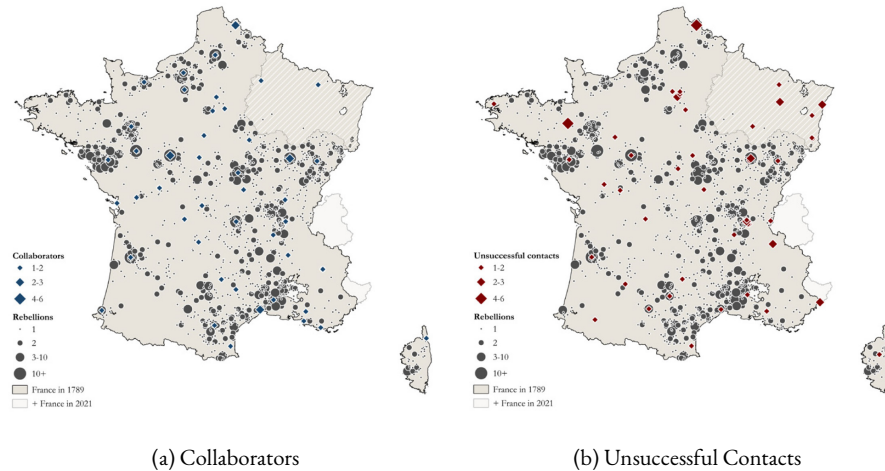


Figure 8 – Spatial Distribution of Rebellions by Type of Collaborator.

Notes. This figure displays the spatial distribution of rebellions by type of collaborator in RGF93 projection. Panel a includes Jean Nicolas's 63 collaborators based on their workplace at the time of their collaboration on the Jean Nicolas survey, and Panel b, the 83 researchers who were contacted by Jean Nicolas but who ultimately did not collaborate in the survey. The spatial distribution of rebellions corresponds to the 2,444 records authored by Jean Nicolas's collaborators that draw exclusively on local archival sources. The highlighted area corresponds to the Grand Est region. The underlying shapefile of the Kingdom of France as of 1789 is based on Gay et al.'s (2023) jurisdictions shapefile (Gay et al., 2024). The underlying shapefile of France as of 2021 is based on IGN's (2021) ADMIN-EXPRESS shapefile.

To assess the magnitude of this spatial bias, I investigate this pattern through a regression analysis. Specifically, I estimate the following linear regression model:

$$(1) \quad \text{Rebellion}_p = \alpha + \beta \log \text{Distance}_p + \mathbf{X}'_p \boldsymbol{\theta} + \varepsilon_p,$$

where Rebellion_p denotes an indicator variable that takes the value of 1 if parish p ever hosted a rebellion between 1661 and 1789, and Distance_p , the distance to the nearest collaborator—or unsuccessful contact—of Jean

Nicolas, in kilometers.⁹⁷ Of course, universities of Jean Nicolas's collaborators are generally located in urban centers, where the incidence of rebellions was presumably higher. Hence, vector **X** controls for parishes population in 1793 and for their representation on Cassini's map, included as fixed effects.⁹⁸ To account for spatial clustering arising from both the local co-occurrence of rebellions and the survey's regional stratification, I cluster standard errors at the level of départements.

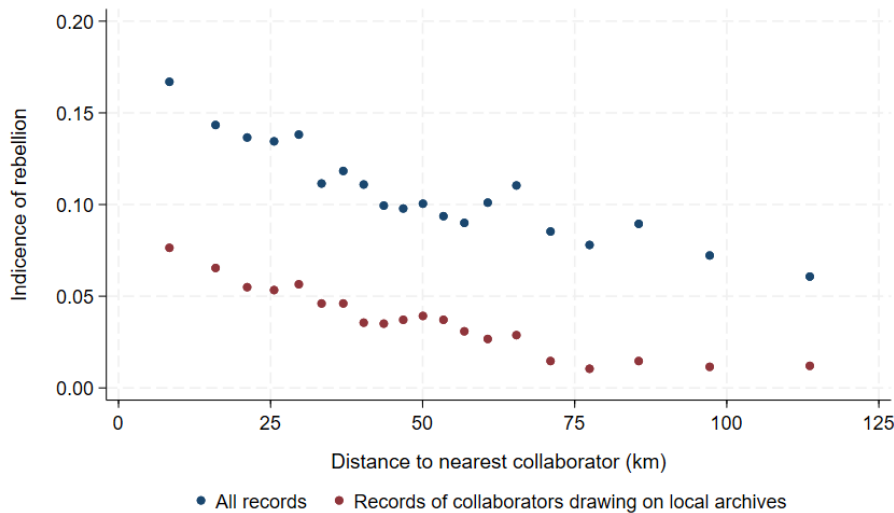


Figure 9 – Incidence of Rebellions and Distance to Nearest Collaborator.

Notes. This figure displays the relationship between the incidence of rebellions and the distance to the nearest collaborator of Jean Nicolas in kilometers. Each dot represents an equal-size bin of about two thousand parishes. Figure generated using Stepner's (2013) `binscatter` Stata command.

I report the results in Columns 1–4 of Table 3. In Panel A, I focus on all rebellion records. The coefficient in Column 1 is statistically significant at the one percent level and implies that a 10 percent increase in the distance to the nearest collaborator is associated with a small decrease in the probability of observing a rebellion of 0.4 percentage point. This coefficient remains stable after controlling for the distance to the nearest unsuccessful contact (Column 3) and parish-level characteristics (Column 4). Results are similar in Panel B when restricting the sample of rebellion records to those authored

⁹⁷ I use the log transformation of the distance measure because it approximates a normal distribution more closely than the untransformed measure, which resembles a right-skewed χ distribution (Appendix Figure A24).

⁹⁸ This vector also controls for the spatial coordinates of parishes to account for the local topography that may affect both the incidence of rebellion and the density of the urban landscape

by collaborators and drawing exclusively on local archival sources, though the relationship is weaker by a factor of one quarter. These results highlight the slight spatial dependence between the location of Jean Nicolas's collaborators and the likelihood of observing a rebellion, suggesting that researchers using the Jean Nicolas survey should be mindful of its spatial representativeness.

Table 3 – Incidence of Rebellions and Distance to Nearest Collaborator.

	Outcome: Indicator for rebellion in 1661–1789				
	(1)	(2)	(3)	(4)	(5)
A. All rebellion records (8,508 cases over 4,264 parishes)					
Log distance to nearest collaborator (km)	−0.041*** [0.010]		−0.040*** [0.010]	−0.041*** [0.008]	−0.016** [0.007]
Log distance to nearest unsuccessful contact (km)		−0.009 [0.009]	−0.002 [0.009]	0.012 [0.007]	0.009 [0.006]
Parishes	40,415	40,415	40,415	39,604	39,604
R-squared	0.008	0.000	0.008	0.160	0.197
B. Rebellion records by collaborators drawing on local archives (2,443 cases over 1,435 parishes)					
Log distance to nearest collaborator (km)	−0.029*** [0.006]		−0.028*** [0.005]	−0.030*** [0.005]	−0.020*** [0.005]
Log distance to nearest unsuccessful contact (km)		−0.012** [0.006]	−0.006 [0.005]	0.003 [0.005]	0.006 [0.005]
Parishes	40,415	40,415	40,415	39,604	39,604
R-squared	0.011	0.002	0.011	0.083	0.129
Controls	No	No	No	Yes	Yes
Département fixed effects (94)	No	No	No	No	Yes

Summary statistics. Incidence of all rebellions: 0.105 (mean), 0.306 (s. d.). Incidence of rebellions by collaborators drawing on local archives: 0.035 (mean), 0.184 (s. d.). Distance to nearest collaborator: 57 (mean), 36 (s. d.). Distance to nearest unsuccessful contact: 59 (mean), 32 (s. d.).

Notes. This table reports the OLS coefficients from estimating Equation 1. The unit of observation is a parish. Controls include parish population in 1793, parish representations in Cassini's map under the form of fixed effects, and parish spatial coordinates. Estimates are calculated using Correia's (2023) `reghdfe` Stata command. Standard errors are clustered at the département level and reported in brackets. *Statistical significance.* *** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.

Reassuringly, this modest spatial bias substantially diminishes once the survey's regional stratification is accounted for: including 94 département fixed effects reduces the correlation by more than one half (Panel A, Column 5) and by one third when focusing on rebellion records authored by collaborators (Panel B, Column 5). This suggests that the Jean Nicolas survey may be considered close to spatially representative for *within-département* comparisons. I therefore recommend that practitioners include

département fixed effects when conducting empirical analyses using the Jean Nicolas database.⁹⁹

4.2.3. Incomplete Examination of Local Archival Sources

Despite the (limited) spatial biases documented in Sections 4.2.1 and 4.2.2, the Jean Nicolas database still reveals notable gaps in the recording of rebellions *within* certain localized regions. To highlight this additional potential source of bias, this section develops a case study on two départements in the region of Eastern Brittany.

Eastern Brittany presents a striking contrast in recorded rebellious activity. In the département of Loire-Atlantique, 418 rebellions are recorded, while in the neighboring Ille-et-Vilaine, only 100 appear in the database.¹⁰⁰ Yet in 1793, both départements had comparable populations: 416 thousand in Loire-Atlantique and 511 thousand in Ille-et-Vilaine. This disparity raises an important question: does it reflect actual differences in levels of unrest, or does it instead point to inconsistencies in the underlying data collection process?

A close analysis of authorship and sources points to the latter explanation (Table 4). Most of the rebellion records for Loire-Atlantique (82 percent) were produced by Sylvie Kleiber during her master's thesis—written under the supervision of Jean Nicolas. She relied primarily on départemental archives, systematically examining the B series of judicial court records, which accounts for 74 percent of the archival sources cited for this département. By contrast, no researcher focused specifically on Ille-et-Vilaine: Jean Nicolas himself produced the majority of its records (77 percent), drawing heavily on secondary sources (33 percent) and relying on national archives for nearly half (45 percent) of the cited archival sources. Although Isabelle Coulangue—another of Nicolas's students—consulted the B series of Ille-et-Vilaine's départemental archives, her survey was limited to a narrow period (1700–40), resulting in a limited coverage (20 rebellion records).

To assess the magnitude of the bias resulting from the incomplete survey of local archives in Ille-et-Vilaine, Brice Évain (2022) conducted a systematic review of the entire B series at the départemental archives of this

⁹⁹ In doing so, and depending on the outcome of interest, practitioners should be mindful of the potential post-treatment bias introduced by controlling for twentieth-century département heterogeneity (Homola et al., 2024).

¹⁰⁰ See Appendix Figure A25. In fact, after Paris, the département of Loire-Atlantique is the one with the most rebellion records in the Jean Nicolas database.

département.^{101,102} Applying the same criteria used in the original Jean Nicolas survey, he identified 87 additional rebellions—nearly doubling the initially recorded number of rebellions for this département.¹⁰³ These findings indicate that about one-third of the discrepancy in recorded rebellions between Loire-Atlantique and Ille-et-Vilaine results from the limited examination of local archival sources. A proper exploitation of the Jean Nicolas database thus requires careful attention to the relative weight of local archives in the sources cited in rebellion records.

4.2.4. Département-Level Grading Scheme

Based on these findings, I propose a département-level grading scheme designed to summarize the quality of rebellion records in the Jean Nicolas database—it focuses on the département level due to the nature of the survey's stratification. This grading scheme retains three criteria aimed at capturing source- and author-driven biases: the share of rebellion records in a département citing at least one départemental archival source, the share citing at least one bibliographic source, and an indicator variable that takes the value of 1 if a collaborator was based in the département.¹⁰⁴

Consistent with previous analyses, each of these measures is strongly correlated with the share of parishes that experienced at least one recorded rebellion between 1661 and 1789. In other words, these measures help predict the incidence of recorded rebellions at the département level despite being unrelated to the underlying historical data generating process of rebellions.¹⁰⁵

¹⁰¹ This research was funded by the ANR-20-CE38-001 (ObARDI) grant and carried out during Summer 2022 with assistance from three master's students: Alix Janvier, Delphine Mialot, and Alicia Riou.

¹⁰² A similar analysis can be conducted for Normandy, drawing on Maneuvrier-Hervieu's (2020) comprehensive survey of subsistence riots from 1709 to 1817.

¹⁰³ See Appendix Figure A25 and Table 4.

¹⁰⁴ I exclude the share of rebellion records citing national archival sources to avoid collinearity with the first two measures. Moreover, I focus on départemental archival sources rather than on broader local sources because municipal sources could generate a spatial bias toward specific localities *within* départements.

¹⁰⁵ Results are similar when using the log number of rebellions as the outcome variable (see Appendix Table A52). Again, it is important to note that this analysis assumes away biases stemming from the destruction, gaps, and reliability of archival sources.

Table 4 – Breton Rebellions in the Jean Nicolas and Brice Évain Surveys.

	Nicolas survey				Evain survey	
	Loire-Atlantique		Ille-et-Vilaine		Ille-et-Vilaine	
	Freq.	Percent	Freq.	Percent	Freq.	Percent
Authors						
NICOLAS Jean	50	11.96	77	77.00		
COLLARD Nicole	13	3.11	2	2.00		
COULANGE Isabelle	14	3.35	20	20.00		
KLEIBER Sylvie	341	81.58	1	1.00		
EVAIN Brice					87	100.00
Rebellion records	418	100.00	100	100.00	87	100.00
Type of sources						
Archive	434	94.76	91	66.91	93	88.57
Bibliography	24	5.24	45	33.09	12	11.43
Sources	458	100.00	136	100.00	105	100.00
Archival service						
National archives	59	13.59	41	45.05		
Départemental archives	325	74.88	33	36.26	93	100.00
Municipal archives	44	10.14	0	0.00		
Other archival services	6	1.38	17	18.68		
Archival sources	434	100.00	91	100.00	93	100.00
Archival series						
AN-G7 General control of finances	20	5.21	7	9.46		
AN-Z1A Court of Aids	27	7.03	18	24.32		
AD-B Courts and jurisdictions	285	74.22	22	29.73	86	92.47
AD-C Provincial administrations	24	6.25	11	14.86	7	7.53
Other archival series	28	7.29	16	21.62		
Nat. and dép. archival sources	384	100.00	74	100.00	93	100.00

Notes. This table reports the distribution of rebellions and their sources in the Jean Nicolas and Brice Évain surveys across two départements of Brittany.

Specifically, Table 5 presents results from a series of linear regressions. In Column 1, the estimate is statistically significant at the one-percent level and implies that a 10 percent increase in the share of rebellions with départemental archival sources is associated with a 1.3 percent increase in the share of parishes with at least one recorded rebellion.¹⁰⁶ This single variable accounts for 16 percent of the variation in the outcome. Next,

¹⁰⁶ Because archival sources for the Île-de-France region are generally located in Paris, I group its eight départements: Paris (78), Seine-et-Marne (77), Yvelines (78), Essonne (91), Hauts-de-Seine (92), Seine-Saint-Denis (93), Val-de-Marne (94), and Val-d'Oise (95).

Column 2 shows that the share of rebellions with bibliographic sources is negatively correlated with the incidence of rebellions, though it explains only 6 percent of its variation. Finally, Column 3 indicates a positive correlation between the presence of a local collaborator and the incidence of rebellions, explaining 10 percent of its variation. Including all three measures simultaneously in Column 4 does not change the magnitude of these results, although the coefficient for bibliographic sources is no longer statistically significant. Together, these measures explain 25 percent of the variation in the incidence of rebellions. Of course, the number of parishes with recorded rebellions in a département should also be a function of its population and number of parishes. Indeed, estimates in Column 5 imply that these characteristics explain one-third of the incidence of rebellions. Nevertheless, even when controlling for these factors in Column 6, methodological biases remain significant predictors of the incidence of rebellions recorded in the Jean Nicolas survey.

Table 5 – Incidence of Rebellions and Département-Level Biases.

	Outcome: Share of parishes with at least one rebellion					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of rebellions with départemental sources	0.130*** [0.033]			0.106*** [0.034]		0.060** [0.030]
Share of rebellions with bibliographic sources		-0.119** [0.057]		-0.080 [0.051]		-0.139*** [0.039]
At least one author in département			0.054*** [0.018]	0.042** [0.016]		0.039*** [0.013]
Log population in 1793					0.076*** [0.023]	0.068*** [0.017]
Log number of parishes					-0.126*** [0.023]	-0.123*** [0.018]
Départements	88	88	88	88	88	88
R-squared	0.158	0.057	0.103	0.249	0.336	0.546

Summary statistics. Share of parishes with at least one rebellion: 0.124 (mean), 0.085 (s. d.). Share of rebellions with départemental sources: 0.349 (mean), 0.260 (s. d.). Share of rebellions with bibliographic sources: 0.255 (mean), 0.172 (s. d.). At least one author in département: 0.477 (mean), 0.502 (s. d.).

Notes. This table reports OLS coefficients. The unit of observation is a département. The eight départements of the Île-de-France region are grouped together. Shares are between 0 and 1. Robust standard errors are reported in brackets.

Statistical significance. *** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.

Based on these findings, I construct a département-level grading scheme to summarize the quality of rebellion records in the Jean Nicolas database. I develop three candidate grading schemes: the first relies exclusively on the measure with the strongest predictive power, i.e., the share of rebellions citing at least one départemental archival source; the second standardizes

and aggregates all three measures linearly; and the third aggregates them using weights proportional to their relative predictive powers.¹⁰⁷ Grades ranging from A to D are then assigned based on the interquartile distribution of each index, with A indicating the highest quality and D, the lowest.

Table 6 – Incidence of Rebellions and Département-Level Quality Grade.

	Outcome: Share of parishes with at least one rebellion					
	(1)	(2)	(3)	(4)	(5)	(6)
Quality grade						
A	0.078*** [0.026]	0.059** [0.028]	0.098*** [0.024]	0.056*** [0.019]	0.029 [0.020]	0.065*** [0.015]
B	0.035 [0.024]	0.034 [0.022]	0.062*** [0.022]	−0.011 [0.020]	0.011 [0.019]	0.038** [0.019]
C	0.002 [0.022]	0.024 [0.023]	0.035* [0.018]	−0.021 [0.019]	−0.006 [0.021]	−0.009 [0.017]
D = excluded						
Log population in 1793				0.088*** [0.022]	0.078*** [0.024]	0.075*** [0.021]
Log number of parishes				−0.130*** [0.021]	−0.123*** [0.022]	−0.122*** [0.022]
Grade type	AD	Flat	Weighted	AD	Flat	Weighted
Départements	88	88	88	88	88	88
R-squared	0.139	0.062	0.182	0.459	0.360	0.461

Summary statistics. Share of parishes with at least one rebellion: 0.124 (mean), 0.085 (s. d.).

Notes. This table reports OLS coefficients. The unit of observation is a département. The eight départements of the Île-de-France region are grouped together. Shares are between 0 and 1. *Quality grade* refers to the grade based on the measure of availability of départemental archival sources; *Flat*, to the grade based on the flat aggregation of all three measures; and *Weighted*, to the grade based on the weighted aggregation of all three measures. Robust standard errors are reported in brackets.

Statistical significance. *** $p \leq 0.01$. ** $p \leq 0.05$. * $p \leq 0.10$.

I then evaluate the predictive power of these grading schemes through fixed-effects regressions. The results in Columns 1–3 of Table 6 indicate that départements assigned with higher quality grades—regardless of the grading scheme—exhibit a greater share of parishes with recorded rebellions. The third grading scheme—based on a weighted aggregation of the three measures—demonstrates the highest predictive power. Controlling

¹⁰⁷ Specifically, weights correspond to each measure's associated R-squared from bivariate regressions with the outcome variable (see Columns 1–3 in Table 5), scaled by the total R-squared sum ($0.158 + 0.057 + 0.103 = 0.318$). The resulting weights are 50 percent for the share of rebellions citing départemental archival sources, 18 percent for bibliographic sources, and 32 percent for the presence of a collaborator.

for département population and number of parishes in Columns 4–6 does not substantively alter these results. Therefore, I adopt the third grading scheme, which I match with the `nicolas_events` dataset. Overall, about one third of records have the grade of **A** (34 percent), another third, the grade of **B** (34 percent), and the rest, the grades of **C** (18 percent) and **D** (14 percent).¹⁰⁸ Users of the Jean Nicolas database should contrast their results across rebellion records characterized by different quality grades.

5. GUIDELINES FOR A PROPER USE OF THE JEAN NICOLAS DATABASE

In summary, I propose a set of practical guidelines to help ensure that analyses based on the Jean Nicolas database yield robust and credible results:

1. **Typology.** Carefully consider the typology of events included in the analysis, as some may fall outside the scope of political unrest (see Section 2.2.2).
2. **Coverage.** Evaluate the representativeness of the analysis sample when using variables with limited coverage (see Sections 3.1.4–3.1.7).
3. **Sources.**
 - (a) **Archival and bibliographic sources.** Test the robustness of analytical result to excluding rebellion records based exclusively on bibliographic sources (see Section 4.2.1).
 - (b) **National and local sources.** Contrast analytical results across rebellion records that draw on national versus local archival sources (see Sections 4.2.1 and 4.2.3).
 - (c) **Archival series and recording actors.** Check that analytical results are not overly influenced by any single archival series representing a specific recording actor (see Section 4.2.1).
4. **Authorship.**
 - (a) **Jean Nicolas and his collaborators.** Examine whether analytical results change depending on whether rebellion records were authored by Jean Nicolas or his collaborators (see Section 4.2.2).

¹⁰⁸ Appendix Table A54 provides the statistical distribution of this grading scheme, and Appendix Figure A26, its spatial distribution.

- (b) **Regional stratification of the survey.** Leverage within-département variation by using département-level fixed effects (see Section 4.2.2).
- 5. **Quality grades.** Contrast analytical results across rebellion records with different quality grades (see Section 4.2.4).

The Jean Nicolas Database

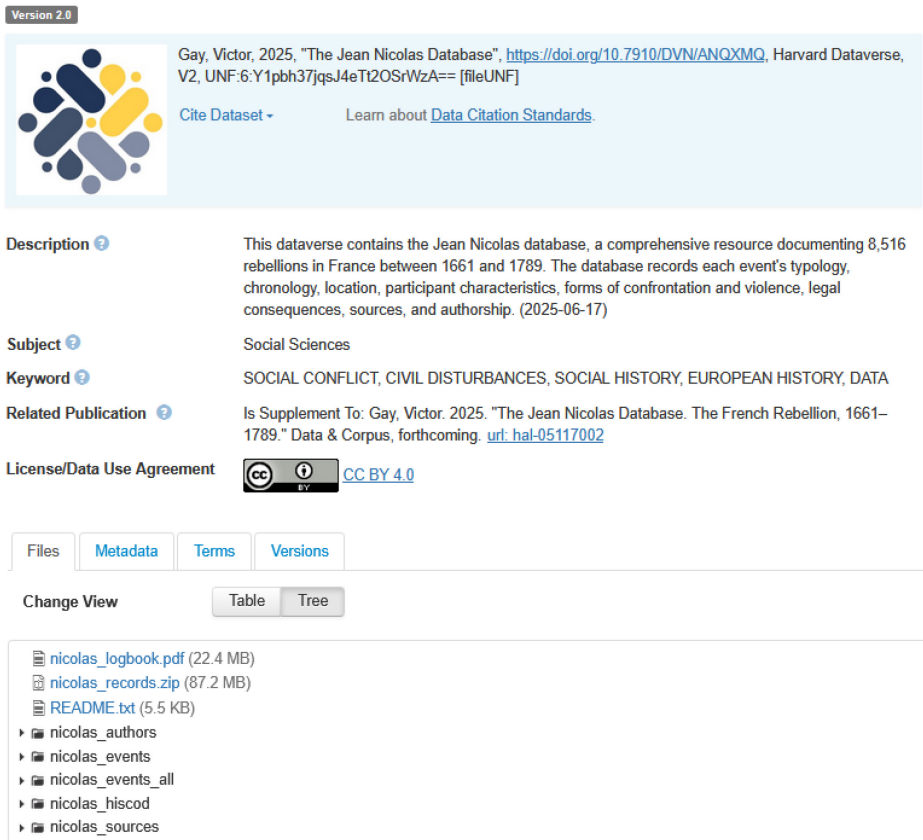


Figure 10 – The Jean Nicolas Database Repository.

Notes. This figure displays the file organization of the Jean Nicolas database repository (Gay, 2025) available on the Harvard Dataverse at <https://doi.org/10.7910/DVN/ANQXMQ>.

6. DATA AVAILABILITY

The content of the Jean Nicolas database is distributed through the Harvard Dataverse under the CC-BY 4.0 license at <https://doi.org/10.7910/DVN/ANQXMQ> (Gay, 2025). The file organization of the database is provided in Figure 10. It contains five sets of files: the 8,977 original records of the Jean Nicolas survey in PDF format (nicolas_records), a copy of Jean Nicolas’s logbook in PDF format

(`nicolas_logbook`), the events dataset (`nicolas_events`), the sources dataset (`nicolas_sources`), and the authors dataset (`nicolas_authors`).¹⁰⁹ These three datasets are provided in Stata DTA format as well as in CSV format, with and without value labels. I also provide codebooks in TXT format that contain the list of variables along with their labels and value labels. All the data and codebook files are available both in English and in French.¹¹⁰

7. POTENTIAL USES

The Jean Nicolas database offers a wide range of applications. Most notably, it provides a foundation for analyzing the spatial and temporal dynamics of popular uprisings in the decades leading up to the French Revolution (Aubert, 2015). It can also serve as a valuable entry point for regional studies of social conflict in early modern France (e.g., Évain, 2022).¹¹¹ More broadly, the Jean Nicolas database is of primary interest to scholars in historical political economy seeking to understand the roots of rebellion during this period (Jenkins & Rubin, 2024). Indeed, the HiSCoD version of the Jean Nicolas survey has already been widely used to examine factors underlying popular uprisings in early modern France, including adverse weather conditions (Chambru, 2019), ruling by local elites (Degrave, 2023), French combatants' exposure to the American Revolution (Jha & Wilkinson, 2023; Ottinger & Rosenberger, 2023), and taxation (Davoine et al., 2025; Giommoni et al., 2025).¹¹² Additionally, I have drawn on the Jean Nicolas database to investigate how the state's growing informational capacity shaped the dynamics of rebellion in the run-up to the Revolution (Albertus & Gay, 2025). Thanks to its fine-grained content, the Jean Nicolas database will contribute to improving our understanding of the roots of popular uprisings in early modern France.¹¹³

¹⁰⁹ The events dataset is also provided for all records of the Jean Nicolas survey under the `nicolas_events_all` folder.

¹¹⁰ Datasets and codebooks in French have the extension `_FR` in their file names. Note that while the labels and value labels are translated, I keep the original variable name to ensure interoperability between datasets.

¹¹¹ The availability of Jean Nicolas's logbook will also contribute to a history of this collective survey (Guirault, 2019).

¹¹² It has also been used to construct control variables for various empirical analyses (Chambru et al., 2024; Degrave et al., 2024; López Peceño, 2024).

¹¹³ The Jean Nicolas database is not intended to be expanded with contributions from contemporary researchers, primarily because the broad conceptual framework used by Jean Nicolas may blur the boundaries between political unrest and more ambiguous phenomena, such as ordinary criminal acts. Researchers wishing to

8. ACKNOWLEDGMENTS

This article is dedicated to the memory of Jean Nicolas (1928–2025), whose pioneering work made it possible. I am grateful to Michael Albertus, Gauthier Aubert, Pascal Bastien, Nicole Charley, Jamie Folsom, François Godicheau, Dominique Godineau, Sébastien Poublanc, Estefania Santacreu-Vasut for their input on this project. I am especially indebted to Renan Donnerh, Brice Évain, and Philippe Hamon for their invaluable contributions to this long-standing endeavor. I am also thankful to Loris Cuenot, Annie Dago, and Louis Vitrand for their outstanding research assistance. Finally, I thank Cédric Chambru for his valuable input about the construction of the HiSCoD database. I gratefully acknowledge the financial support of the Agence Nationale de la Recherche (ANR) under grants ANR-17-EURE-0010 (Investissements d’Avenir program) and ANR-20-CE38-0015 (ObARDI).

9. CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Victor Gay: Writing – review & editing; Writing – original draft; Visualization; Validation; Supervision; Software; Resources; Project administration; Methodology; Investigation; Funding acquisition; Formal analysis; Data curation; Conceptualization.

10. DECLARATION OF CONFLICTING INTERESTS

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

11. BIBLIOGRAPHY

- Aberdam, S. (2004). *Démographes et démocrates. L’œuvre du Comité de division de la Convention nationale*. Société des études robespierristes.
- Albertus, M., & Gay, V. (2025). State-building and rebellion in the run-up to the French Revolution. *American Political Science Review*.
- Andress, D. (Ed.). (2013). *The Oxford handbook of the French Revolution*. Oxford University Press.
- Armatte, M. (2020). Contextes et régimes de production et d’analyse des données. In L. Feller & A. Gramain (Eds.), *L’évident et l’invisible. Questions de méthodes en économie et en histoire* (pp. 101–114). Éditions de la Sorbonne.

contribute to a dynamic and evolving database are encouraged to turn instead to HiSCoD, which actively welcomes such contributions—see <https://mrsh.unicaen.fr/hiscod/contribuer.html>.

- Aubert, G. (2014). *Les révoltes du papier timbré, 1675. Essai d'histoire événementielle*. Presses universitaires de Rennes.
- Aubert, G. (2015). *Révoltes et répressions dans la France Moderne*. Armand Colin.
- Aubert, G. (2023). *1709: L'année où la révolution n'a pas éclaté*. Calype.
- Bély, L. (Ed.). (1996). *Dictionnaire de l'Ancien Régime*. Presses Universitaires de France.
- Bercé, Y.-M. (1974). *Histoire des Croquants. Étude des soulèvements populaires au XVII^e siècle dans le sud-ouest de la France*. Librairie Droz.
- Bideau, G., & Verdier, N. (2024). Une trame communale immuable ? La nécessaire géohistoire des communes françaises (1800–2024). *Cybergeo*, 1074. <https://doi.org/10.4000/12k5j>
- Bischoff, G. (2009). Contestations paysannes entre Vosges et Forêt Noire : La génération du Bundschuh (1493–1525). In G. Brunel & S. Brunet (Eds.), *Les luttes anti-seigneuriales* (pp. 61–78). Presses universitaires du Midi.
- Bosc, H. (1985–93). *La guerre des Cévennes: 1702–1710* (Vols. 1–6). Presses du Languedoc.
- Bourquin, L., Beauvalet-Boutouyrie, S., Brian, I., Le Gall, J.-M., & Pitou, F. (2005). *Dictionnaire historique de la France moderne*. Belin.
- Bouton, C. A. (1993). *The Flour War: Gender, class, and community in late Ancien Régime French society*. Pennsylvania State University Press.
- Brette, A. (1904). *Atlas des bailliages ou juridictions assimilées ayant formé unité électorale en 1789 dressé d'après les actes de la convocation conservés aux Archives nationales*. Imprimerie Nationale.
- Burguière, A. (1991). La déstabilisation de la société française. In J. Dupâquier (Ed.), *Histoire de la population française. 2, De la Renaissance à 1789* (pp. 475–493). Presses Universitaires de France.
- Cabourdin, G. (1977). *Terre et hommes en Lorraine (1550–1635). Toulous et comté de Vaudémont*. Université de Nancy II.
- Cabourdin, G., & Viard, G. (1978). *Lexique historique de la France d'Ancien Régime*. Armand Colin.
- Chambru, C. (2019). *Do the right thing! Leaders, weather shocks and social conflicts in pre-industrial France*. EHES Working Paper (No. 161). https://www.ehes.org/wp/EHES_161.pdf
- Chambru, C., Henry, E., & Marx, B. (2024). The dynamic consequences of state building: Evidence from the French Revolution. *American Economic Review*, 114(11), 3578–3622. <https://doi.org/10.1257/aer.20220110>

- Chambru, C., & Maneuvrier-Hervieu, P. (2024). Introducing HiSCoD: A new gateway for the study of historical social conflict. *American Political Science Review*, 118(2), 1084–1091. <https://doi.org/10.1017/S000305542300076X>
- Cirone, A. (2023). Data in historical political economy. In J. A. Jenkins & J. Rubin (Eds.), *The Oxford handbook of historical political economy* (pp. 31–54). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197618608.013.3>
- Correia, S. (2023). *REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects* (Version 6.12.3) [Software]. Statistical Software Components, Boston College Department of Economics. <https://ideas.repec.org/c/boc/bocode/s457874.html>
- Cristofoli, P., Vouloir, M.-C., Duménieu, B., & Motte, C. (2021). *Des chefs-lieux de Cassini aux communes de France (1756–1999): Données géo-historiques associant la position géographique des lieux représentés sur la Carte Générale de la France dite de Cassini (1756–1789) à l'histoire des communes de France (An III-1999)* [Spatial dataset]. Didómena. https://didomena.ehess.fr/concern/data_sets/6395wb092
- Davoine, E., Enguehard, J., & Kolesnikov, I. (2025). *The Political Costs of Taxation*. EU Tax Observatory Working Paper (No. 27). <https://www.taxobservatory.eu/publication/the-political-costs-of-taxation/>
- Dedieu, J.-P. (2024). Construire des bases de données: Introduction sous forme d'un retour d'expérience. *Histoire & Mesure*, 39(2), 7–26. <https://doi.org/10.4000/140kh>
- Degrave, A. (2023). Local rule, elites, and popular grievances: Evidence from Ancien Régime France. *Journal of Historical Political Economy*, 3(1), 1–29. <https://doi.org/10.1561/115.00000044>
- Degrave, A., Peceño, A. L., & Rozenas, A. (2024). *Peasants into citizens: Suffrage expansion and mass politics in France*. TSE Working Paper (No. 24–1529). <https://www.tse-fr.eu/publications/peasants-citizens-suffrage-expansion-and-mass-politics-france>
- Dennison, T., & Gehlbach, S. (2024). Interdisciplinary collaboration in historical political economy. In J. A. Jenkins & J. Rubin (Eds.), *The Oxford handbook of historical political economy* (pp. 17–30). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197618608.013.2>
- Évain, B. (2022). *Rébellions bretonnes* [Université Rennes 2]. <https://shs.hal.science/halshs-03850634>
- Foisil, M. (1970). *La révolte des Nu-pieds et les révoltes normandes de 1639*. Presses Universitaires de France.

- Gallet, J. (2009). Recherches sur les révoltes contre les seigneurs en Lorraine pendant la première modernité. In G. Brunel & S. Brunet (Eds.), *Les luttes anti-seigneuriales* (pp. 137–148). Presses universitaires du Midi. <https://books.openedition.org/pumi/9151>
- Gay, V. (2021). Mapping the Third Republic: A Geographic Information System of France (1870–1940). *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 54(4), 189–207. <https://doi.org/10.1080/01615440.2021.1937421>
- Gay, V., Gobbi, P. E., & Goñi, M. (2023). *Bailliages in 1789 France* (Version 2.0) [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/T8UXHK>
- Gay, V., Gobbi, P. E., & Goñi, M. (2024). The atlas of local jurisdictions of Ancien Régime France. *Journal of Historical Geography*, 84, 49–60. <https://doi.org/10.1016/j.jhg.2024.03.008>
- Gay, V. (2025). *The Jean Nicolas database* (Version 2.0) [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/ANQXMQ>
- Genet, J.-P. (1986). Histoire, Informatique, Mesure. *Histoire & Mesure*, 1(1), 7–18. <https://doi.org/10.3406/hism.1986.904>
- Genet, J.-P. (1997). La genèse de l'État moderne: Les enjeux d'un programme de recherche. *Actes de la recherche en sciences sociales*, 118(1), 3–18. <https://doi.org/10.3406/arss.1997.3219>
- Giommoni, T., Loumeau, G., & Tabellini, M. (2025). *Extractive taxation and the French Revolution*. CEPR Discussion Paper (No. 20110). <https://cepr.org/publications/dp20110>
- Girault, B. (2019). L'archive et le document : Matériaux pour une histoire des sciences sociales (note critique). *Annales. Histoire, Sciences Sociales*, 74(3–4), 779–800. <https://doi.org/10.1017/ahss.2020.58>
- Gorry, J.-M. (2008). Des communes de la Révolution aux communes actuelles. In E. Zadora-Rio (Ed.), *Des paroisses de Touraine aux communes d'Indre-et-Loire: La formation des territoires* (pp. 121–183). FERACF Editions. https://www.persee.fr/doc/sracf_1159-7151_2008_ant_34_1_1379
- Gossare, M. (1997). *Les Tards avisés, 1707: Histoire d'une révolte paysanne en Quercy, Périgord et Agenais*. L'Hydre Éditions. <https://gallica.bnf.fr/ark:/12148/bpt6k3322024w>
- Grenier, J.-Y., Béguin, K., & Bonzon, A. (2003). *Dictionnaire de la France moderne*. Hachette Supérieur.
- Hamon, P. (2021). L'enquête Jean Nicolas sur la rébellion en France (1661–1789) [Web page]. *L'atelier d'Obardi*. <https://doi.org/10.58079/sh4t>
- Hamon, P. (2022). Qu'est-ce qu'une rébellion ? [Web page]. *L'atelier d'Obardi*. <https://doi.org/10.58079/SH4Y>

- Homola, J., Pereira, M. M., & Tavits, M. (2024). Fixed effects and post-treatment bias in legacy studies. *American Political Science Review*, 118(1), 537–544. <https://doi.org/10.1017/S0003055423001351>
- INSEE. (2021). *Code officiel géographique au 1er janvier 2021* (Version 1.0) [Dataset]. INSEE. <https://www.insee.fr/fr/information/5057840>
- Jenkins, J. A., & Rubin, J. (2024). Historical political economy: What is it? In J. A. Jenkins & J. Rubin (Eds.), *The Oxford handbook of historical political economy* (pp. 3–16). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197618608.013.1>
- Jha, S., & Wilkinson, S. (2023). *Revolutionary Contagion*. Stanford University Graduate School of Business Research Paper (No. 4084). <https://www.gsb.stanford.edu/faculty-research/working-papers/revolutionary-contagion>
- Karila-Cohen, K., Lemerrier, C., Rosé, I., & Zalc, C. (2018). Nouvelles cuisines de l'histoire quantitative. *Annales. Histoire, Sciences Sociales*, 73(4), 773–783. <https://doi.org/10.1017/ahss.2019.90>
- LaDéHiS. (2021). Site Cassini [Web resource]. *Laboratoire de Démographie et d'Histoire Sociale*. <https://ladehis.ehess.fr/index.php?355>
- Le Roy Ladurie, E. (1966). *Les paysans de Languedoc* (Vols. 1–2). S.E.V.P.E.N.
- Lemerrier, C., Ollivier, C., & Zalc, C. (2013). Articuler les approches qualitatives et quantitatives : Plaidoyer pour un bricolage raisonné. In M. Hunsmann & S. Kapp (Eds.), *Devenir chercheur: Écrire une thèse en sciences sociales* (pp. 125–143). Éditions de l'École des hautes études en sciences sociales. <https://doi.org/10.4000/books.editionsehess.1301>
- Lemerrier, C., & Zalc, C. (2019). *Quantitative methods in the humanities: An introduction*. University of Virginia Press. (Original work published 2008)
- Levitsky, S., & Way, L. (2013). The durability of revolutionary regimes. *Journal of Democracy*, 24(3), 5–17. <https://doi.org/10.1353/jod.2013.0043>
- Litvine, A., Thévenin, T., Ségué, I., & Starzec, A. (2023). *French Historical GIS, 1700–2020. Administrative Units, Populations, Transports, Economy* (Version 1.0) [Dataset]. Zenodo. <https://doi.org/10.5281/ZENODO.3727273>
- López Peceño, A. (2024). *Does Education Make Subjects or Citizens? Primary Schooling and Anti-Authoritarian Mobilization in 19th-Century France*. SSRN. <https://doi.org/10.2139/ssrn.5076305>
- Maneuverier-Hervieu, P. (2020). *La Normandie dans l'économie Atlantique au XVIIIe siècle: Production, commerce et crises* [PhD Dissertation, Université de Caen Normandie]. <https://theses.fr/2020NORMC032>

- Markoff, J. (1990). Peasants protest: The claims of lord, church, and state in the cahiers de doléances of 1789. *Comparative Studies in Society and History*, 32(3), 413–454. <https://doi.org/10.1017/S0010417500016583>
- Meuriot, P. (1918). Le recensement de l'an II. *Journal de la Société de statistique de Paris*, 59, 34–56. https://www.numdam.org/article/JSFS_1918__59__34_0.pdf
- Motte, C., & Vouloir, M.-C. (2007). Le site cassini.ehess.fr. Un instrument d'observation pour une analyse du peuplement. *Bulletin du Comité français de cartographie*, 191, 68–84. <http://geoprodig.cnrs.fr/items/show/189128>
- Mours, S. (1958). *Les Eglises réformées en France*. Librairie protestante ; Oberlin.
- Mousnier, R. (1967). *Fureurs paysannes. Les paysans dans les révoltes du XVIIe siècle (France, Russie, Chine)*. Calmann-Lévy.
- Nicolas, J. (1973). Pour une enquête sur les émotions populaires au XVIIIe siècle. Le cas de la Savoie. *Annales historiques de la Révolution française*, 214(1), 593–607. <https://doi.org/10.3406/ahrf.1973.4155>
- Nicolas, J. (1974). Pour une enquête sur les émotions populaires au XVIIIe siècle. Le cas de la Savoie. *Annales historiques de la Révolution française*, 215(1), 111–153. <https://doi.org/10.3406/ahrf.1974.4311>
- Nicolas, J. (1978). *La Savoie au XVIIIe siècle. Noblesse et bourgeoisie*. Maloine S.A. Éditeur.
- Nicolas, J. (Ed.). (1985a). *Mouvements populaires et conscience sociale: XVIe–XIXe siècles. Actes du colloque de Paris, 24–26 mai 1984*. Maloine S.A. Éditeur.
- Nicolas, J. (1985b). Un chantier toujours neuf. In J. Nicolas (Ed.), *Mouvements populaires et conscience sociale: XVIe–XIXe siècles. Actes du colloque de Paris, 24–26 mai 1984* (pp. 13–20). Maloine S.A. Éditeur.
- Nicolas, J. (2002). *La rébellion française. Mouvements populaires et conscience sociale (1661–1789)*. Éditions du Seuil.
- Ottinger, S., & Rosenberger, L. (2023). The American origin of the French Revolution. *IZA Discussion Papers*, 15974. <https://docs.iza.org/dp15974.pdf>
- Pelletier, M. (1990). *La carte de Cassini. L'extraordinaire aventure de la carte de France*. Presses de l'École Nationale des Ponts et Chaussées.
- Pillorget, R. (1975). *Les mouvements insurrectionnels de Provence entre 1596 et 1715*. Éditions Pedone.
- Pillorget, R., & Pillorget, S. (1996). *France baroque, France classique. 1589–1715. Dictionnaire* (Vols. 1–2). Robert Laffont.

- Porchnev, B. F. (1963). *Les soulèvements populaires en France de 1623 à 1648*. S.E.V.P.E.N.
- Ribon, P. (2001). *D'Artagnan en Ardèche: La révolte de Roure en 1670* (Vol. 1–1). Éditions et Régions.
- Robert, D. (1961). *Les Églises réformées en France (1800–1830)*. Presses Universitaires de France.
- Ruggles, S. (2021). The revival of quantification: Reflections on old new histories. *Social Science History*, 45(1), 1–25. <https://doi.org/10.1017/ssh.2020.44>
- Sauvion, C. (2024). *Modèle Numérique de Terrain (MNT) France métropolitaine et DROM* (Version 1.0) [Dataset] data.gouv.fr. <https://www.data.gouv.fr/datasets/modele-numerique-de-terrain-mnt-france-metropolitaine-et-drom/>
- Scheltjens, W. (2023). Upcycling historical data collections. A paradigm for digital history? *Journal of Documentation*, 79(6), 1325–1345. <https://doi.org/10.1108/JD-12-2022-0271>
- Slater, D. (2010). *Ordering Power. Contentious Politics and Authoritarian Leviathans in Southeast Asia*. Cambridge University Press.
- Stepner, M. (2013). *BINSCATTER: Stata module to generate binned scatterplots* (Version 7.02) [Software]. Statistical Software Components S457709. <https://ideas.repec.org/c/boc/bocode/s457709.html>
- Thompson, E. P. (1971). The moral economy of the English crowd in the eighteenth century. *Past & Present*, 50(1), 76–136. <https://doi.org/10.1093/past/50.1.76>
- Tilly, C. (1986). *La France conteste: De 1600 à nos jours* (É. Diacon, Trans.). Fayard.

1.2 Revolutionary Inheritance Reforms and the Demographic Transition

This section provides the following analysis and data articles:

- Analysis article: Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2025. “Revolutionary Transition: Inheritance Change and Fertility Decline.” *Journal of Political Economy*, forthcoming.³
- Data article 1: Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2024a. “The Atlas of Local Jurisdictions of Ancien Régime France.” *Journal of Historical Geography*, 84: 49–60.⁴
- Data article 2: Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2024b. “The Customary Atlas of Ancien Régime France.” *Explorations in Economic History*, 93: 101588.⁵

³The appendix to this article is available at <https://doi.org/10.1086/739821>. The replication material for this article is available at <https://doi.org/10.7910/DVN/PTZRGN>.

⁴The original dataset associated with this article is available at <https://doi.org/10.7910/DVN/T8UXHK>.

⁵The appendix to this article is available at <https://doi.org/10.1016/j.eeh.2024.101588>. The original dataset associated with this article is available at <https://doi.org/10.7910/DVN/HYE209>.



Revolutionary Transition: Inheritance Change and Fertility Decline

Victor Gay, Paula Gobbi, Marc Goñi

► To cite this version:

Victor Gay, Paula Gobbi, Marc Goñi. Revolutionary Transition: Inheritance Change and Fertility Decline. 2025. <hal-04285818v3>

HAL Id: hal-04285818

<https://hal.science/hal-04285818v3>

Preprint submitted on 5 May 2025 (v3), last revised 4 Dec 2025 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Revolutionary Transition: Inheritance Change and Fertility Decline*

Victor Gay[†] Paula E. Gobbi[‡] Marc Goñi[§]

April 14, 2025

Abstract

We test Le Play’s (1875) hypothesis that the French Revolution contributed to France’s early fertility decline by imposing equal partition of inheritance among all children, including women. We combine new data on local inheritance rules before the Revolution and individual-level demographic data from historical sources and crowdsourced genealogies. Difference-in-differences and regression-discontinuity estimates show that the inheritance reforms enacted during the Revolution reduced completed fertility by 0.5 children. A key mechanism was the desire to avoid land fragmentation across generations. These reforms closed the fertility gap between regions with different historical inheritance rules and crucially contributed to France’s demographic transition.

JEL classification: D10, J10, K11, N33, O10.

Keywords: *Demographic transition, Fertility, French Revolution, Inheritance.*

*We thank George Alter, Dan Bogart, Bruno Caprettini, Bastien Chabé-Ferret, Matthew Curtis, Neil Cummins, David de la Croix, Mathias Dewatripont, Martin Dribe, James Feigenbaum, Martin Fiszbein, Oded Galor, Vincent J. Geloso, Claudia Goldin, Edward Glaeser, Noel D. Johnson, Mark Koyama, Satomi Kuroso, Volha Lazuka, Robert A. Margo, Stelios Michalopoulos, Faustine Perrin, Thomas Piketty, Gilles Postel-Vinay, Jean-Marc Robin, Lukas Rosenberger, Elyce Rotella, Louis Rouanet, Edson Severnini, Pierre Sicsic, Yannay Spitzer, Maria Stanfors, Philip Ushchev, and seminar participants at AMSE, Bergen, Bologna, Brown, BU, CY Cergy, CMI, ECARES, EIEF, GMU, Harvard, IESEG, Konztanz, LSE, Luxembourg, Mannheim, Pittsburgh, PSE, SDU, TSE, UCLA, UC Louvain, UM, Uppsala, UTEP, Tinbergen, UCI, and conference participants at AFSE 2023 (Paris), ASSA 2024 (San Antonio), CNEH 2023 (Toronto), HHCSI 2023 (Solstrand), LORDE 2022 (Rouen), SAEe 2021 (Barcelona), and SSHA 2023 (Washington) for their suggestions. This research was co-funded by the the Agence Nationale de la Recherche (ANR) under grant agreements ANR-17-EURE-0010 (Investissements d’Avenir program) and ANR-20-CE38-0015 (ObARDI), and by the European Union (ERC) under grand agreement 947654 (IDED). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

[†]Department of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole. Email: victor.gay@tse-fr.eu.

[‡]Université Libre de Bruxelles (ECARES) and CEPR. Email: paula.eugenia.gobbi@ulb.be.

[§]University of Bergen and CEPR. Email: marc.goni@uib.no.

1 Introduction

The demographic transition is a critical stage in the development of any society. Historically, sustained economic growth began only after large segments of the population limited their fertility. Low fertility rates prevented gains in output from being largely offset by rapid population growth, paving the way for modern economic growth (Galor 2012).¹

Because of its importance for the economy and society, a large literature has studied the causes of historical and modern fertility transitions (Guinnane 2011). Most research by economists highlights changes in the economic incentives for having children based on the trade-off between the quantity and the quality of children (Becker and Lewis 1973). Under this framework, the fertility transition emerged as a result of technological progress and increased demand for education after the Industrial Revolution (Galor and Moav 2002; Delventhal et al. 2021). Alternatively, other changes concomitant to industrialization have been linked to the decline in fertility, such as health improvements (Cervellati and Sunde 2015), advances in contraception technology, or changes in social norms about their use (Beach and Hanlon 2022).

Although some of these factors played an important role in England, where the Industrial Revolution preceded the fertility transition, they cannot rationalize why the first large-scale fertility transition started in eighteenth-century France, more than 50 years before its industrialization, and subsequently spread to neighboring countries (Perrin 2022; Melki et al. 2024). Figure 1 illustrates France’s early fertility decline by showing the crude birth rate, the Princeton I_g index of marital fertility, and women’s completed fertility at age 40 from 1700 to 1850. These three measures indicate that fertility had been slowly declining throughout the eighteenth century, long before France’s industrialization in the 1850s. As a result, a recent literature has linked France’s early fertility decline to deep-rooted social norms, culture, or religiosity, rather than to sharp changes in economic incentives following the Industrial Revolution (Spolaore and Wacziarg 2022; Blanc 2024a). That said, some crucial features of the French fertility transition remain unexplained. Most notably, Figure 1 also shows that a sharp reduction in fertility began shortly after the French Revolution. The speed of this change is difficult to rationalize with deep-rooted, slowly-evolving cultural factors, while its timing seems to rule out economic factors linked to industrialization.

In this article, we offer a novel explanation for the sharp fertility decline in late eighteenth-century France illustrated in Figure 1. Instead of changes associated with the Industrial Revolution or cultural factors, we show that institutions affected the economic incentives for having children. Specifically, we revisit the longstanding hypothesis

¹In the twentieth century, all developing countries that reached medium income levels first experienced a drop in fertility (Chesnais 1992), with the exception of oil producing countries.

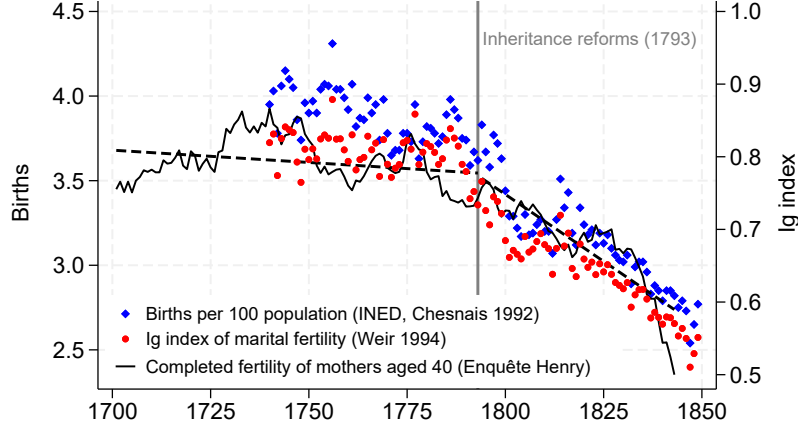


Figure 1: Fertility decline in France, 1700–1850.

Notes: I_g is the Princeton index of marital fertility, i.e., the number of legitimate births relative to the expected number of births in a Hutterite population with the same number and age structure of married women. *Completed fertility* is a 5-year moving average plotted in the year when mothers were aged 40. Dashed lines are fitted values relative to completed fertility before and after 1793.

of [Le Play \(1875\)](#), which links France’s early fertility decline to the new inheritance laws introduced during the French Revolution.² Motivated by the equality principle, a series of reforms passed in 1793 ruled that inheritance had to be partitioned equally among all children including women (henceforth, egalitarian inheritance). These reforms effectively harmonized inheritance systems across France which, before the Revolution, had been regulated by a myriad of local customary and written laws. In some areas, inheritance was already egalitarian before 1793. In others, the reforms meant that parents could no longer nominate a unique heir or exclude daughters from inheritance (henceforth, inegalitarian inheritance). [Le Play \(1875, pp. 26, 314\)](#) argued that these egalitarian inheritance reforms contributed to France’s early fertility decline, as households had to curbe family size to avoid the fragmentation of their land among multiple heirs.

We test Le Play’s hypothesis by examining whether these legal changes on inheritance affected fertility decisions. To do so, we create a new atlas mapping the 141 different customs and laws that regulated inheritance across France before the Revolution. This enables us to identify, at a highly-disaggregated level, which areas were treated by the 1793 inheritance reforms and which already had an inheritance system that complied with the new egalitarian rules. We then link this information to individual-level demographic data from the *Enquête Louis Henry* ([Séguy 2001](#)) and from online genealogies in [geni.com](#) (henceforth, Henry and Geni databases, respectively). These databases enable us to reconstruct the completed fertility, childlessness rate, and age at marriage of women born in the eighteenth and early nineteenth centuries. Our main identification strategy is

²Despite the many reforms introduced during the Revolution, only the reforms on inheritance have received attention in the literature as a potential drivers of fertility decline.

a difference-in-differences (DD) approach based on comparing cohorts of fertile age to cohorts too old to be fertile in 1793 between municipalities where the reforms altered and did not alter the existing inheritance system. We further augment our analysis with a regression-discontinuity difference-in-differences (RD-DD) design that compares women born very close to – but on opposite sides of – the border delimiting judicial districts with different inheritance rules, before and after the reforms that harmonized inheritance laws across this border.

Consistent with Le Play’s hypothesis, we find that the 1793 inheritance reforms substantially reduced fertility. Women exposed to the reform reduced their completed fertility by half a child, or by 17 percent relative to the average number of surviving children in treated areas before the reforms (2.92). This magnitude is similar to the pre-reform fertility gap between areas with egalitarian and inegalitarian inheritance. We obtain consistent estimates across our DD and RD-DD estimation strategies, which use data obtained from very different methodologies: the family reconstruction method (Henry database) and crowdsourced genealogies (Geni database). Results are also consistent along the extensive and intensive margins of fertility: before 1793, childlessness was lower and the fertility of mothers was higher in inegalitarian-inheritance areas; after 1793, the reforms contributed to close these fertility gaps. Taken together, our results imply that the 1793 inheritance reforms brought large areas of the country to the low-fertility regime that was predominant where egalitarian inheritance was already at place before the reforms. This crucially accelerated the decline in fertility in late eighteenth-century France. We present a counterfactual exercise suggesting that the 1793 inheritance reforms accounted up to 46 percent of the fertility decline during the first phase of the demographic transition (1760s–1850s) (van de Walle 1986) and to 36 percent over the entire fertility transition (1760s–1936) (Delventhal et al. 2021).³

Despite the numerous discussions around Le Play’s hypothesis among social scientists, data limitations have until now prevented a conclusive verdict (Chesnais 1992, p. 339). An important contribution of this article is to provide the first complete map of inheritance rules over France’s territory at the onset of the Revolution. Our map is at the highly-disaggregated judicial district level. Specifically, we georeference and vectorize the boundaries of the 435 judicial districts in which different customary rules applied based on Brette’s (1904a) *Atlas des Bailliages*. We then match these districts to their relevant customary law and classify inheritance customs into egalitarian and inegalitarian inheritance. To do so, we primarily use the *Nouveau Coutumier Général* (Bourdot de Richebourg 1724), which provides the original text of each custom.

The empirical setting we examine offers a number of advantages. The timing of the 1793 reforms was unexpected, as inheritance was not at the core of the popular grievances

³We do not claim that the French Revolution was the only cause of the fertility transition. There were forces limiting fertility before then, e.g., secularization (Murphy 2015; Blanc 2024a).

raised in the *cahiers de doléances* in 1789 (Goy 1988). Likewise, concerns about fertility were not instrumental for policy makers, who instead aimed at enforcing the equality principle (Shaffer 1982), unifying the legal system across France (Hyslop 1934), and consolidating the younger generations’ support for the Revolution (Lataste et al. 1901, pp. 681–3). Inheritance reforms were quickly upheld after 1793 by family tribunals (Desan 2006) and there was substantial regional variation in the set of laws and customs that were affected, which we exploit in our DD and RD-DD strategies. Furthermore, we present aggregate trends in fertility and individual-level event-study estimates to show that, prior to 1793, fertility followed common trends in areas with different inheritance systems. To deal with potential unobservable confounders, we extend our baseline DD design to control for differential fertility trends in municipalities that varied along a range of economic, religious, political, and economic-geography characteristics. These flexible-trend DD specifications account for the possibility that economic conditions – the primary determinant of fertility in Malthusian societies – might have evolved differently across areas with different inheritance systems. Finally, the validity of our augmented RD-DD estimates rests on the assumption that unobservable factors vary smoothly over space. We document that, by 1793, areas affected and not affected by the reforms were balanced on numerous individual- and local-level covariates. This finding is consistent with the historical origins of France’s heterogeneous inheritance systems. These different systems were rooted in the laws of the Germanic peoples that ruled different parts of the territory after the fall of the Roman Empire, and, therefore, are largely unrelated with environmental or structural factors.⁴

Next, we explore the mechanisms behind our results. The fertility decline following the 1793 inheritance reforms could be the result of two complementary forces. First, both by imposing equal partition and by extending the right to inherit to women, the reforms increased the number of heirs. This reduced economic incentives for having children to avoid the fragmentation of land into inefficiently small parcels. Second, by extending the right to inherit to women, the reforms potentially increased their bargaining power within the household, thereby reducing their fertility. The ideal experiment to disentangle these mechanisms would be to estimate the effect of granting inheritance rights to women where the reform did not change the number of heirs. Unfortunately, this is not possible, as extending the right to inherit to women roughly doubles the number of heirs by construction. In addition, inheritance systems that excluded women also tended to nominate a single heir. That said, we present evidence suggesting that the

⁴We show that our DD and RD-DD estimates are robust to a variety of validity checks and alternative specifications—including controlling for migration, changes in mortality associated with the demographic transition, or terrain characteristics; placebo and permutation tests; alternative definitions of the treatment group; alternative inference methods correcting for spatial correlation in the error term; and adjusting for under-reported children deaths using the first-name repetition technique (Cummins 2020).

French Revolution, despite granting women access to inheritance, did not empower wives (Desan 2006; Tudor 2021). This is mainly because the 1804 Civil Code gave husbands absolute control over jointly or independently owned property, including inheritances. Consistently, we find no effect of the reforms on spouses’ relative age and education, two distribution factors typically associated with wives’ bargaining power in the household (Browning et al. 2014).

Qualitative and quantitative evidence suggest that land fragmentation played a larger role. At the eve of the Revolution, France was populated by farmers with small landholdings living close to subsistence levels (de Brandt 1901). Such farming structure is characterized by indivisibility constraints, that is, further dividing landholdings can result in production falling below the subsistence level. Hence, the egalitarian inheritance reforms incentivized smaller family sizes to avoid land fragmentation among sons and daughters. We document that discussions of inheritance reform in the National Constituent Assembly, as well as contemporary observers and existing case-studies, frequently mention this hypothesis. We then evaluate this mechanism through two quantitative exercises. First, we digitize information on 36 thousand parcels listed in the Napoleonic cadastre for a subset of nine municipalities under different inheritance rules. We document three facts about land fragmentation, including its strong association with egalitarian inheritance practices. Second, we formalize this mechanism in a model and bring its testable implications to the data. We show that the effects of the 1793 inheritance reforms were circumscribed to areas where soil and terrain characteristics favored land fragmentation, while fertility remained unaffected in areas where these characteristics favored large farms.⁵

Relative to the existing literature, we make the following contributions. First, we study a novel and overlooked determinant of fertility choices: legal institutions regulating inheritance. Second, we are the first to provide empirical support for Le Play’s hypothesis. Third, we provide a complete and highly disaggregated atlas of inheritance customs and laws in Ancien Régime France. And fourth, our study sheds new light to the economic consequences of inheritance systems.

Although the literature on fertility determinants is vast, legal factors are overlooked (Doepke et al. 2022). Instead, since Becker and Lewis (1973)’s seminal quantity-quality trade-off theory, human capital is widely studied to understand fertility differences over time or across individuals.⁶ Under this framework, the rise of skilled labor and urbanization can reduce fertility (Ager et al. 2020; Baudin and Stelter 2022). In addition, the literature has considered other triggering factors for fertility reductions, such as health improvements and a decline in child mortality (Sah 1991; Bar and Leukhina 2010; Bhat-

⁵We use soil texture and terrain ruggedness, two (exogenous) environmental characteristics typically associated with different degrees of fragmentation in the landownership distribution.

⁶See, e.g., Galor and Weil (2000), Galor and Moav (2002), Murtin (2013), Baudin et al. (2015), Murphy (2015), Vogl (2016), and de la Croix and Perrin (2018).

tacharya and Chakraborty 2012; Herzer et al. 2012; Cervellati and Sunde 2015), female empowerment (de la Croix and Vander Donckt 2010; Brée and de la Croix 2019; Hazan et al. 2022), old-age security (Boldrin et al. 2015; Rossi and Godart 2022), cultural transmission from migrants (Daudin et al. 2019), as well as information and usage of family planning methods (Cavalcanti et al. 2020; Beach and Hanlon 2022). However, France’s early fertility transition is at odds with most of these explanations, as human capital (Blanc and Wacziarg 2020), health improvements (Brée and de la Croix 2019; Perrin 2022), or advances in contraception are unlikely to have contributed to the initial phases of France’s fertility decline. In contrast, deep-rooted cultural factors and secularization have been highlighted (Spolaore and Wacziarg 2022; Blanc 2024a). Closer to our article, Cummins (2013) shows that the French fertility transition started in places where inequality was low, and Weir (1984) and Rosental (1991) argue that the French Revolution and the introduction of the Civil Code drove France’s fertility decline. Our article analyzes a novel factor – legal institutions on inheritance – hence bringing this new piece to the unresolved puzzle of France’s fertility transition.

In addition, this paper provides the first empirical test of Le Play’s hypothesis. Formulated in 1884, it was the first theory on France’s demographic transition. Social scientists have been skeptical about the possibility that the French Revolution and its inheritance reforms contributed to the fertility decline (Chesnais 1992, p. 338).⁷ However, as Chesnais (1992, p. 339) acknowledges, empirical evidence for or against Le Play’s hypothesis is lacking. The main reason is that there exists no complete, local-level atlas of inheritance customs and laws before the Revolution. One of our contributions is to construct such an atlas. Beyond this article, our novel atlas will enable scholars to delimit customary regions in Ancien Régime France and study the legacies of historical customs. We provide more details on our customary atlas and on the spatial distribution of judicial districts in Gay et al. (2024a, 2024b).

Finally, we contribute to the literature studying how inheritance systems affect economic outcomes. In the long run, inheritance systems can shape income and wealth inequality (Piketty 2011). Bartels et al. (2024) show that, in Germany, locations with equal inheritance later exhibited higher income, education, labor productivity, and entrepreneurship. Inheritance rules can also affect gender gaps and education (Bertocchi and Bozzano 2015) or pension systems (Galasso and Profeta 2018). However, despite

⁷Le Play’s hypothesis has been subject to two main critiques: first, fertility was relatively low in the southwest before the Revolution. The critique is based on the misconception that primogeniture predominated everywhere in the south. Our inheritance atlas shows that this region had heterogeneous inheritance laws (see Figure 3). Moreover, we use granular data from Geni – which includes the southwest – to corroborate Le Play’s hypothesis. Second, other countries adopted similar inheritance reforms without experiencing fertility drops (Chesnais 1992, p. 338). We believe that this critique lacks empirical support. That said, our findings can rationalize why similar reforms will reduce fertility where landownership is fragmented (e.g., France) but will not affect fertility where landownership is concentrated (e.g., England and Prussia).

Habakkuk (1955)’s work, the literature has mostly overlooked the effects of inheritance rules on family decisions and fertility.⁸ As for short-term effects, several studies have analyzed laws that extend inheritance rights to women or forbid dowries in developing countries.⁹ France was the first country to pass a national law on inheritance. Hence, we shed new light to the potential consequences of large-scale inheritance reforms.

2 Historical background

2.1 Inheritance in Ancien Régime France

At the eve of the Revolution, France was predominantly populated by small landholding farmers (de Brandt 1901; Bourdieu et al. 2013). Between 40 and 80 percent of households owned their land.¹⁰ Hence, inheritance laws on immovable property were relevant for much of the French population. Much as other aspects of daily life, inheritance practices in Ancien Régime France were regulated by a blend of written laws and customs. Written laws, derived from Roman Law (the Justinian Code), were prevalent mostly in the south. Customs represented a set of long-established local rules that emerged from traditional practices (Chénon 1926). Initially oral, customs were codified following Charles VII’s Ordinance of Montils-lès-Tours in 1454 (Grinberg 2006, p. 66).

While Roman law gave complete freedom to assign a unique heir through a testament, customary laws contained specific inheritance provisions. Despite their complexity, inheritance rules can be classified into four categories: partible and impartible rules, and rules where women had the right to inherit or not. Under partible rules, family wealth was divided among several offspring. This division could take the form of “strict equality,” where married offspring had to return any donation received before inheriting (e.g., a dowry) or “option,” which allowed them to keep such donations and opt-out from inheriting. Under impartible rules, inheritance could not be broken and parents could favor one child over other offspring. These rules included primogeniture, ultimogeniture, and unigeniture. Inheritance systems further varied in whether women could inherit or not. One of the justifications for excluding women, as well as for impartible rules, was to prevent the family wealth from breaking down and diluting among descendants (Yver 1966, p. 39). We refer to inheritance systems affected by the reform – those with impartible rules *and/or* that excluded women – as *inegalitarian inheritance* systems. Conversely, we refer to inheritance systems that already complied with the reform before 1793 – those with partible rules *and* that included women – as *egalitarian inheritance* systems.

The origins of these different inheritance systems are not well understood. The prevailing theory is that inegalitarian inheritance arises in farming economies where land is the

⁸Exceptions are Casari et al. (2019) and Gobbi and Goñi (2020).

⁹See Anderson and Genicot (2015), Roy (2015), and Bahrami-Rad (2021) for India, Aldashev et al. (2012) and La Ferrara and Milazzo (2017) for Ghana, and Harari (2019) for Kenya.

¹⁰See Appendix F for an overview of estimates on the share of landowners in France.

primary source of wealth and is subject to indivisibilities (Bertocchi 2017). Other explanations highlight concerns over mortality and the lineage’s survival (Chu 1991), economic uncertainty (Grieco and Ziebarth 2015), housing markets, and limited access to wood to build new houses (Zink 1993). Motivated by these theories, we show that areas with different inheritance rules were balanced across an extensive set of covariates—including the key hypothesized determinants of inequalitarian inheritance (see Table 1). This aligns with the dominant view on the origins of the different legal systems in Ancien Régime France. These origins were rooted in the laws of the peoples – namely the Burgundians, Visigoths, Salian Franks, and Ripuarian Franks – that moved across the territory upon the fall of the Western Roman Empire in 476 (Chénon 1926), and hence, are uncorrelated with any late eighteenth-century factors affecting the sharp fertility decline after 1793.

2.2 The 1793 inheritance reforms

Inheritance was not among the main popular grievances in the run-up to the French Revolution. For example, the *cahiers de doléances* written in preparation of the Estates General of 1789 seldom mentioned the issue: of the 571 cahiers analyzed by Goy (1988), only 8 mentioned inheritance rules.

Despite the lack of popular demand, a series of decrees enacted throughout 1793 and resulting in the Loi de Nivôse, an II (January 6, 1794), established egalitarian inheritance among all offspring – including women – throughout the territory. These inheritance reforms shattered regional differences in inheritance rules that had prevailed for centuries.¹¹ Three elements help explain why the revolutionaries reformed inheritance. First, equality concerns were central to the Revolution (Shaffer 1982). Second, the revolutionaries aimed at unifying France’s fragmented legal system and administrative geography (de Tocqueville 1856).¹² Third, there was mounting concerns that those who joined the Revolution would be disinherited by their parents (Lataste et al. 1901, pp. 681–3).¹³

The consensus in the historiography is that the population quickly abode by the new rules. Indeed, offspring felt they had the right to an equal share in the family patrimony and were soon eager to protect their new rights (Shaffer 1982, p. 95). Family tribunals ensured the enforcement of the new law. In the late 1790s, disputes over inheritances were the most common cases these tribunals had to resolve (Desan 2006, ch. 4).

¹¹The text of this law is available in the *Lois et Actes du Gouvernement*, Tome VIII (pp. 214–29) at <https://gallica.bnf.fr/ark:/12148/bpt6k56370f/f219>.

¹²Voltaire (1829, pp. 229–30) complained that “[a] man who travels in this country changes laws almost as many times as he changes horse post [9–12 kilometers].”

¹³Philippeaux, a representative at the National Convention, declared in 1793: “There are a hundred thousand younger sons waiting for this law to fly at the borders, but the fear of being reduced to poverty, by being disinherited from their parents, [...] prevents them from leaving” (*Archives Parlementaires de 1787 à 1860*, Tome LIX, p. 681, March 7, 1793).

3 Data

3.1 Inheritance rules in Ancien Régime France

To construct an atlas of inheritance rules in France at the eve of the Revolution, we start by creating a shapefile of the 435 judicial districts within which inheritance rules applied. This shapefile is based on Brette’s (1904a) atlas. Specifically, we georeference all 32 large-scale maps of this historical atlas and manually attribute a judicial district to each municipality.¹⁴ We then record the 141 customs or written laws that existed before the Revolution and assign them to their corresponding judicial district. Here, we rely on Brette (1904b, 1915) for about half of judicial districts, which we complement with Bourdot de Richebourg (1724), Joignon (1989) for the region of Lorraine, and Zink (1993) for the southwest. We display the resulting shapefiles of customary and written-law regions on the left-hand side of Figure 2.¹⁵

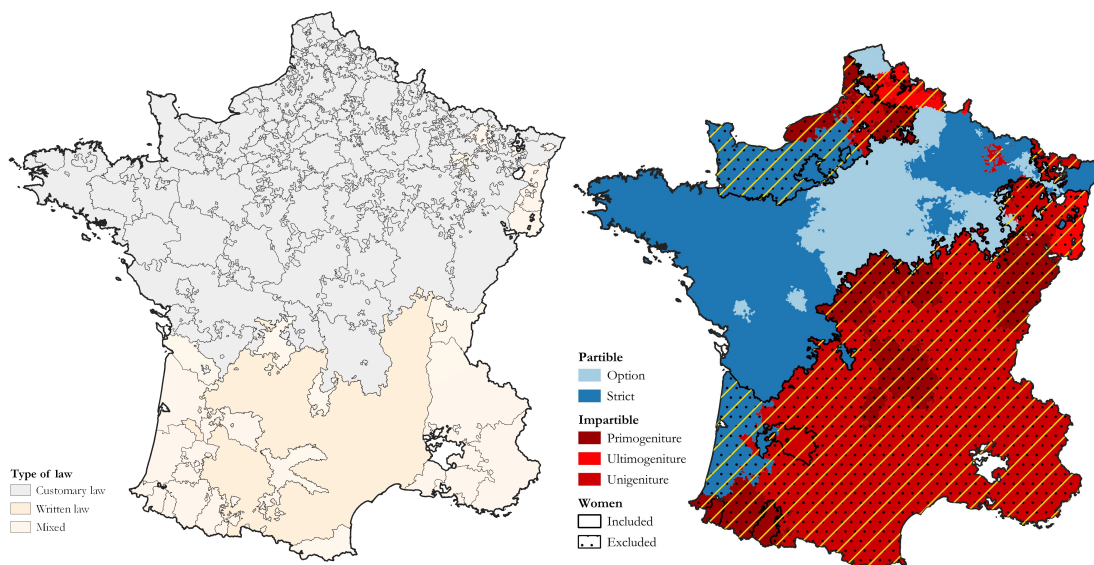


Figure 2: Customs and inheritance systems in Ancien Régime France.

Notes: The left panel displays the spatial distribution of customary-law and written-law areas in Ancien Régime France. The right panel displays the spatial distribution of inheritance systems across partibility and women inclusion. It also highlights (in yellow) areas treated by the 1793 inheritance reforms.

The figure shows that France was broadly divided into two legal regimes. In the north – the *pays de droit coutumier* – customary law prevailed, with substantial local variation in customs. In contrast, in the south – the *pays de droit écrit* – Roman law and

¹⁴We focus on France as of 1789, which corresponded to its current territory with the exceptions of the Duchy of Savoy, the County of Nice, the Comtat Venaissin, and Corsica.

¹⁵See Gay et al. (2024a) and (2024b) for details on the construction of these shapefiles, which are openly available on the Harvard Dataverse. The spatial distribution of underlying judicial districts is provided in Appendix Figure B1.

the *corpus iuris civilis* of Justinian prevailed.¹⁶ In some areas, customary and written laws coexisted. There, written law was often supplementary and applied only when a relevant customary rule was absent (Chénon 1926, pp. 331–4).¹⁷ That said, regarding inheritance, these regions followed customary law (Poumarède 1972).¹⁸

Finally, we code and classify the specific rules on inheritance in each custom or written law. To do so, we resort to the *Nouveau Coutumier Général* (Bourdote de Richebourg 1724), a compendia of the customs and written laws prevalent in Ancien Régime France. For customs not listed in this source, we use original documents that codified customs (see Table A1 in Gay et al. 2024b). We classify inheritance rules along the four dimensions described in Section 2.1: partible versus impartible rules, and rules that included versus excluded women from inheritance. Our map of inheritance systems is the first to cover the entire territory of France, as other work has focused on specific areas: Yver (1966) for the north, Joignon (1989) for Lorraine, and Zink (1993) for the southwest. Moreover, while previous work has focused exclusively on the partibility dimension, we provide the first map that also displays inheritance systems that included versus excluded women. This enables us to distinguish, at a fine-grain level, between areas that complied with egalitarian-inheritance rules imposed by the 1793 reforms – those with partibility including women – from areas with inegalitarian-inheritance rules that were treated by the reforms – those with impartibility and/or excluding women.

The right panel of Figure 2 displays our map of inheritance, highlighting partible rules (light and dark blue), impartible rules (light and dark red), areas excluding women from inheritance (dots), as well as areas treated by the 1793 inheritance reforms (yellow lines). Consistent with previous work on partial areas of France (Yver 1966; Joignon 1989; Zink 1993), our map shows that the southeast followed impartible rules whereas the north and the southwest had areas under either partible or impartible rules. Moreover, by codifying women’s inheritance rights, our novel map uncovers a strong spatial correlation between impartible rules and rules excluding women. In other words, in most areas in the treatment group, the 1793 inheritance reforms simultaneously abolished impartible rules and extended the right to inherit to women. This spatial correlation is intuitive as the historical rationale behind impartibility and the exclusion of women was similar: to prevent the family wealth from diluting among many descendants (Yver 1966, p. 39). That said, the overlap is not perfect in the southwest and in Normandy, where women were excluded under partible rules. This geography of inheritance was stable from the codification of the customs in the fifteenth century up to the Revolution (Gilissen 1979, p. 250), when impartible rules and rules excluding women were abolished, affecting the

¹⁶Le Bris (2019) studies the consequences of these two regimes in the long run. He finds that Roman law led to higher economic development than customary law.

¹⁷“Mixed” areas include the Basque country, Provence, the Dauphiné (Poumarède 1972; Zink 1993), and parts of Lorraine and Alsace (Ganghofer and Levresse 1977; Joignon 1989).

¹⁸See Appendix Figures B2 and B3 for a comparison to previous work.

areas highlighted in yellow.

3.2 Fertility from the *Enquête Louis Henry*

Our main individual-level fertility data are from the *Enquête Louis Henry* (Fleury and Henry 1958; Séguy 2001). This database was generated through the method of reconstitution of families, i.e., linking birth, marriage, and death records within and between individuals. It relies on parish and hospital records before 1792 and civil registers afterwards.¹⁹ We use the nominative part of the Henry database, which contains individual-level information on the fertility of 34,812 women born between 1604 and 1803 in 39 rural municipalities.²⁰ The left panel of Figure 3 displays the locations of these municipalities. Our baseline sample consists of 20,332 women born between 1700 and 1803. We consider women born in the eighteenth century as this provides a balanced sample of cohorts exposed and not exposed to the 1793 inheritance reforms: 52.9 percent of women in our sample were fertile after 1793 (aged below 40) and 47.1 percent had completed their biological reproductive span before 1793 (aged over 40).

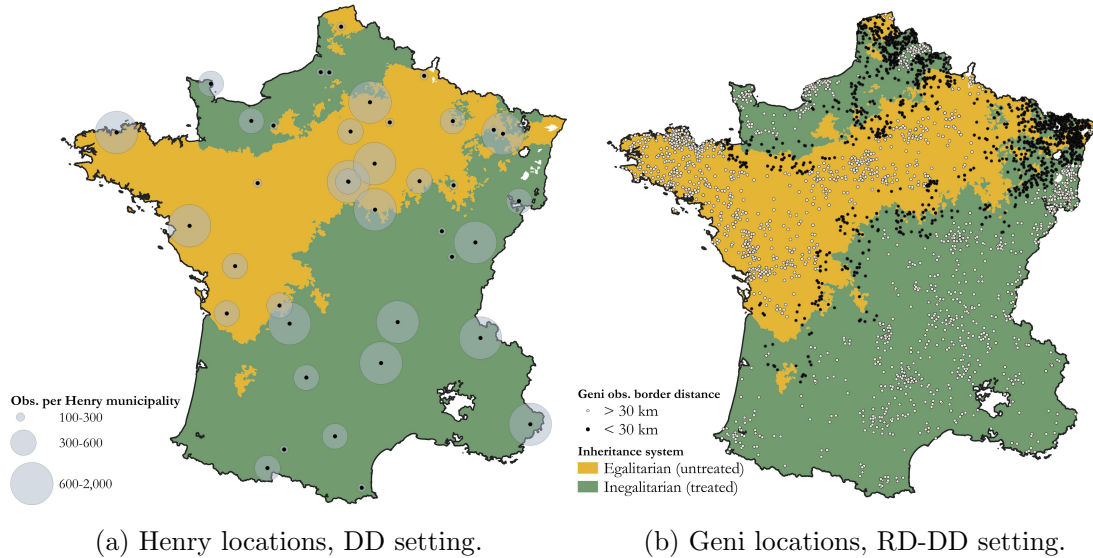


Figure 3: Observations across Egalitarian and Inegalitarian Inheritance.

Notes: Yellow areas denote egalitarian inheritance systems before the Revolution, and green areas, inegalitarian inheritance systems. The left panel displays the distribution of observations in the Henry database used in the DD setting. The right panel displays the distribution of observations in the Geni database used in the RD-DD setting.

The Henry database lists the total number of births over a woman's lifetime and the

¹⁹The high quality of the Henry database is well known among historical demographers, as reflected by the more than 300 studies using this database (Renard 1997).

²⁰Starting from the 41 randomly drawn municipalities of the nominative part of the Henry database, we exclude two municipalities: Massongy, which was in the Duchy of Savoy in 1789, and Suze-sur-Sarthe, which only contains a few observations.

year of death of her children who died young (see Appendix Figure B4 for an example of record). We use this information to construct three fertility variables: completed fertility, completed fertility of mothers, and a childlessness indicator. These variables focus on net fertility, that is, they are restricted to children who reached age 6. Child mortality was high until the 1800s, with 1–1.5 children per mother dying before age 6. Nevertheless, both crude and net fertility declined sharply for mothers born in the 1750s, i.e., for those who were fertile after the 1793 inheritance reforms (see Appendix Figure B5). In addition, the Henry database includes years of birth, marriage, and death, literacy indicators based on whether marriage certificates were signed, information on whether different relatives were alive at the time of marriage, and the accuracy of each entry.²¹

Appendix Table A1 provides summary statistics for our baseline sample. Net completed fertility has a sample average of 2.35 children, the completed fertility of mothers is 3.19, while one in four women are childless. Mean age at marriage is about 26 for women; their husbands are 3.44 years older. On average, women had children from age 26 (first birth) to 35 (last birth). The sample is balanced in terms of the treatment: 47.7 percent of women were born in inegalitarian-inheritance areas which were affected by the reform. Roughly half of parents-in-law were alive at their children’s marriage, suggesting that family wealth was passed down around this time.

3.3 Fertility from online genealogies in `geni.com`

We complement our fertility data with information from crowdsourced genealogies available in [geni.com](https://www.geni.com), a MyHeritage company. The main advantage of this crowdsourced data relies in its sheer size—the underlying database contains over 153 million observations. This data provides information on the timing and locations of birth, baptism, marriage, death, and burial, as well as family links between parents, spouses, and children (Alburez-Gutierrez et al. 2023).

To construct our sample, we keep profiles of individuals born in France between 1700 and 1810.²² Next, we assign latitudes and longitudes to birth locations using the GeoNames database, and verify each match.²³ Based on the municipality of birth, we assign a pre-reform inheritance rule and a treatment status.

We measure individual fertility using the family links of each Geni profile to reconstruct their family trees. Genealogical data often suffer from substantial bias as most users provide information on their direct ancestors (parents, grandparents, ...) but not

²¹Accuracy measures comprise ten categories based on the availability of birth and marriage dates (see Appendix Table A2).

²²To do so, we search for words that can refer to France in the variables containing the birth and baptism location information, e.g., “francia” or “frankrijk.”

²³To verify matches, we project the georeferenced points onto the contemporary shapefile of France’s municipalities. We then attribute the corresponding municipality information and compare it to the information from GeoNames. The automatic georeferencing has an accuracy rate close to 70 percent. We correct the remaining inaccuracies manually.

on their collateral ancestors (uncles, granduncles, ...). As a result, families with single children are over-represented. To overcome this issue, we follow the horizontal restriction proposed by [Blanc \(2024b\)](#) and keep an observation if at least one of the four preceding generations has more than one recorded offspring. Applying this restriction, our Geni sample contains 11,649 women born in France between 1700 and 1810 spread over 2,966 different locations (see the right panel of Figure 3). Hence, despite its lower quality relative to the Henry database, the Geni database provides a broader coverage of the territory. The average mother in the Geni database had 3.55 children who reached age 6. Moreover, fertility trends for mothers born in 1700–1810 are consistent across the Henry and Geni databases, and Familinx—the sub-sample of the Geni database used by [Blanc \(2024b\)](#) (see Appendix Figure B6).

4 Empirical strategy

4.1 Difference-in-differences specification

Our aim is to assess the effect of the 1793 inheritance reforms on women’s fertility. Our identification strategy consists in a difference-in-differences approach based on comparing cohorts of women who were fertile in 1793 to cohorts of women who were too old to be fertile in 1793, between municipalities where the reforms altered versus did not alter the inheritance system. Our main specification is:

$$Y_{icm} = \alpha + \beta I_m \times F_c + \mu_c + \delta_m + \mathbf{X}_i' \theta + \epsilon_{icm} , \quad (1)$$

where Y_{icm} is the completed fertility recorded at the end of the reproductive span of women i born in municipality m in cohort c ; μ_c and δ_m are fixed effects for cohorts and municipalities; and $I_m \times F_c$ is an interaction between a reform treatment indicator equal to one for women from municipalities with an inheritance system affected by the 1793 reforms (I_m), and an exposure indicator equal to one for cohorts of fertile age, i.e., below 40, after 1793 (F_c). We consider a woman’s biological reproductive span to be in ages 15–40. We also evaluate the robustness to using spans over ages 15–45 and 15–30, and to allowing for non-linear effects at different ages.²⁴ The treatment group ($I_m = 1$) comprises women born in municipalities with unequalitarian inheritance, i.e., where assets were impartible and/or where women were excluded from inheritance before 1793. The control group ($I_m = 0$) comprises women born in municipalities with egalitarian inheritance among all children, including women, before the 1793 reforms. Hence, the parameter β captures the effect of the 1793 egalitarian inheritance reforms on fertility, which is the combined effect of abolishing impartible rules and extending

²⁴The 15–40 reproductive spans fits our data, where the average woman had her last child at age 35 with a 6-years standard deviation (see Table A1). Additionally, evidence from pre-1800 European populations shows a sharp decline in marital fertility rates after age 40 ([Clark 2007](#)).

inheritance rights to women. It captures both the reduced economic incentives for having children to avoid land fragmentation among sons and daughters, as well as the potential effect of increased women’s bargaining power.

In extended specifications, vector $\mathbf{X}_i$ includes individual-level controls that are potentially correlated with fertility: an indicator for women’s literacy, fixed effects for the different accuracies of the Henry form, and indicators for whether parents were alive at the time of marriage. We cluster standard errors at the municipality level. In Section 5.2, we use Conley (1999) standard errors to show that results are robust to allowing for various degrees of spatial correlation in the error term.

4.2 Identifying assumptions

Parallel trends. The identifying assumption in our difference-in-difference strategy is that, in the absence of the inheritance reforms, average fertility would have followed the same trend in municipalities under different inheritance systems.

In Panel (a) of Figure 4, we consider whether trends in women’s fertility varied systematically across areas with different pre-reform inheritance rules. We plot average fertility by birth cohort in areas with impartible rules excluding women (red) and in areas with partible inheritance rules including women (blue).²⁵ For cohorts that completed their reproductive span before 1793, trends are declining in a parallel fashion. In detail, note that average fertility was lower by about 0.3 children for women born in the mid-eighteenth century relative to women born in the early 1700s – a pattern consistent with previous evidence on the early decline in birth rates in France (van de Walle 1986; Guinnane 2011; Blanc 2024b). That said, we observe no systematic difference in this early fertility decline between areas with different inheritance rules. The fertility gap between areas with different inheritance rules is constant around 0.7 children for both women born in the early 1700s (3 children versus 2.3) and women born in the early 1750s (2.7 children versus 2)—the latter being the last cohorts to complete their reproductive span before the 1793 reforms. Among cohorts that were fertile after 1793, the figure shows a steeper fertility decline in areas affected by the reforms: the fertility gap closes from 0.7 children to 0 for women born after 1790—the cohorts whose entire reproductive span occurred after the reforms.

In Panel (b) of Figure 4, we use an event-study design to test for differential pre-trends across areas treated by the reforms versus areas in the control group. Specifically, we plot the interaction coefficients β_t for 16 birth-cohort bins with our treatment indicator for municipalities where the inheritance system was affected by the 1793 reforms:

$$Y_{icm} = \alpha + \sum_{t \neq 40-44} \beta_t a(t) \times I_m + \mu_c + \delta_m + \mathbf{X}_i' \theta + \epsilon_{icm} , \quad (2)$$

²⁵Appendix Figure B5 shows the aggregate trends for the fertility of mothers.

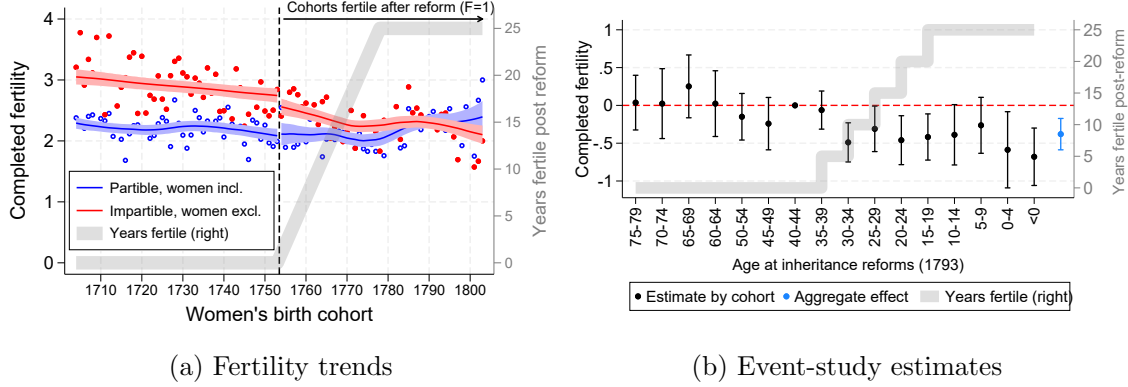


Figure 4: Completed fertility by cohort across inheritance systems.

Notes: Panel (a) shows average completed fertility (dots), trends (lines), and 95-percent confidence intervals (shaded areas) from a local polynomial regression on each side of the 1753 birth cohort. Panel (b) shows β_t coefficients and 95-percent confidence intervals from Equation (2) on a sample of all women aged 0–79 at the time of the reforms or born after them (1794–1803). Individual-level controls are as in Table 2, Column (5). Gray lines show the remaining fertile years after 1793 for each cohort (right axis).

where Y is completed fertility, $a(t)$ are indicators for age t in 1793 (in 4-year groups), and the omitted group is women aged 40–44, i.e., those who completed their biological reproductive span in 1793. The treatment indicator, I_m , is defined as in Equation (1). We also include cohort and municipality fixed effects, μ_c and δ_m , and the individual-level controls described above, $\mathbf{X}_i$.

The results provide no evidence of a differential fertility trend between treated and control municipalities before the 1793 inheritance reforms. Most pre-reform estimates on cohorts older than 40 in 1793 are close to zero and none is statistically significant, so trends in fertility before 1793 are unlikely to confound our results. In contrast, for most cohorts younger than 35 at the time of the reforms, completed fertility drops by around 0.5 children and estimates are statistically significant. This provides preliminary evidence that the 1793 inheritance reforms contributed to France’s fertility decline.

In addition, the event study supports our specification in Equation (1) instead of a design where a woman’s exposure to the reforms is an increasing function of her number of remaining fertile years after 1793. Except for the cohort aged 35–39 in 1793, estimates for all post-treatment cohorts are statistically indistinguishable from each other, independently of the number of remaining fertile years, and are very similar to the aggregate effect obtained by pooling all age groups into a post-reform dummy as in Equation (1). This pattern is biologically plausible and consistent with the demography literature showing that, at the start of the demographic transition, women limited their fertility by not having their “last child” (Knodel 1987; Cinnirella et al. 2017).

Specifications with flexible trends. Our identification strategy also requires that there exists no omitted time-varying and municipality-specific characteristics correlated

with both pre-reform inheritance systems and fertility. The most serious threat to this assumption is that, before industrialization, good economic conditions were associated with high fertility, and these economic conditions might have evolved differently across areas with different initial inheritance systems. To address this concern, we report extended specifications that control for trends in economic conditions using wheat prices and population density:

$$Y_{icm} = \alpha + \beta I_m \times F_c + \mu_c + \delta_m + \mathbf{X}_i' \theta + \gamma p_{md(c)} + \sum_t \mathbf{1}[c = t] \times \mathbf{Z}_m' \Gamma_t + \epsilon_{icm}, \quad (3)$$

where $p_{md(c)}$ is the average wheat price in municipality m in decade $d(c)$, and $\mathbf{Z}_m$ includes the pre-reform wheat price (in the 1780s) and population density (in 1793) in municipality m , interacted with birth-cohort fixed effects, $\mathbf{1}[c = t]$, so as to account for flexible trends in fertility along these two pre-reform characteristics.²⁶

We further account for the possibility that fertility followed differential trends in areas that varied along religious, political, and economic-geography characteristics. Specifically, $\mathbf{Z}_m$ also includes two variables that capture religiosity at the local level, a strong predictor of fertility in pre-industrial France (Murphy 2015). One is proximity to Church authorities (*évêchés*); the other is the proportion of marriages during lent and advent between 1792 and 1815. This variable exploits the fact that the Catholic Church did not perform marriages during lent and advent, while civil marriages – introduced in 1792 – could take place year-round. Therefore, the prevalence of marriages during lent and advent reflects how secularized a municipality was. We also include proxies for the support to the Revolution, the adhesion to the equality principle that inspired inheritance reforms, and information about the reforms and other revolutionary events. To do so, we use distances to the closest political society and rebellion against state authorities in the decade preceding the Revolution. Finally, we consider institutional and economic-geography factors by including distances to judicial district seats (*bailliages*), which resolved inheritance legal cases, to territorial administrative centers (*subdélégations*), which provided better access to public infrastructures, markets, innovation, and economic development, to tax collection centers (*recettes des finances*), and to paved roads and horse-post relays, which capture economic and information networks. All distances are included as $\log(1 + \text{km})$.²⁷

Balancedness. We show in Table 1 that before the 1793 inheritance reforms, a wide range of individual- and municipality-level characteristics were balanced across areas with egalitarian- and inegalitarian-inheritance systems. The only characteristic that differed systematically was fertility (Panel A): before the harmonization of inheritance rules, there

²⁶High wheat prices can proxy bad economic conditions, as demand was inelastic, markets, poorly integrated, and price variations, primarily driven by weather conditions (see Appendix H for more details).

²⁷We construct our religiosity measure from marriage dates in the Henry database. For more information on the sources of our other control variables, see Appendix H.

Table 1: Balancedness of pre-reform characteristics.

	Pre-reform Inheritance				Difference	
	Egalitarian		Inegalitarian			
	Mean	S.d.	Mean	S.d.	Mean	S.e.
<i>Panel A. Pre-reform individual-level characteristics</i>						
Completed fertility	2.22	(2.28)	2.84	(2.72)	0.62	[0.16]**
Wife's age at death	56.96	(17.08)	58.51	(17.57)	1.55	[1.56]
Husband's age at death	59.72	(15.53)	61.59	(15.80)	1.87	[1.27]
Wife died before age 40	0.14	(0.34)	0.12	(0.32)	−0.02	[0.02]
Wife's mother alive at marriage	0.51	(0.50)	0.47	(0.50)	−0.04	[0.06]
Husband's mother alive at marriage	0.45	(0.50)	0.41	(0.49)	−0.04	[0.05]
Wife's father alive at marriage	0.42	(0.49)	0.41	(0.49)	−0.01	[0.05]
Husband's father alive at marriage	0.38	(0.48)	0.36	(0.48)	−0.01	[0.04]
Known birth year or in GPC	0.63	(0.48)	0.63	(0.48)	−0.00	[0.05]
Known union end date	0.73	(0.44)	0.67	(0.47)	−0.06	[0.03]
Wife's literacy	0.16	(0.36)	0.09	(0.29)	−0.06	[0.06]
Husband's literacy	0.37	(0.48)	0.29	(0.45)	−0.08	[0.10]
Observations	5,870		4,882		10,752	
<i>Panel B. Pre-reform municipality-level characteristics</i>						
Wheat price (log)	1.98	(0.09)	2.07	(0.23)	0.09	[0.06]
Population density (log)	3.97	(0.54)	3.87	(0.59)	−0.10	[0.18]
Religiosity index	0.56	(0.30)	0.53	(0.28)	−0.03	[0.09]
Distance to religious centre	3.18	(0.61)	3.15	(0.83)	−0.04	[0.23]
Distance to judicial district seat	2.37	(0.78)	2.66	(0.65)	0.29	[0.23]
Distance to territorial administration	2.55	(0.69)	2.47	(0.55)	−0.07	[0.20]
Distance to tax center	2.69	(0.74)	2.83	(0.70)	0.14	[0.23]
Distance to political society	1.74	(0.64)	1.67	(0.66)	−0.07	[0.21]
Distance to rebellion in 1780–89	2.97	(1.05)	2.71	(0.77)	−0.26	[0.29]
Distance to paved road	0.81	(0.62)	0.90	(0.67)	0.09	[0.21]
Distance to horse-post relay	2.16	(0.60)	2.63	(0.68)	0.46	[0.21]**
Average caloric suitability of land	1902	(305)	1868	(396)	−34	[114]
Terrain ruggedness (in 100s of meters)	0.51	(0.31)	0.76	(0.52)	0.25	(0.15)
Share sandy soils	0.17	(0.32)	0.13	(0.25)	−0.04	(0.09)
Observations	20		19		39	

Notes: In Panel A, the sample are the 10,752 women in the Henry database who completed their reproductive span before the 1793 inheritance reforms. In Panel B, the sample are the 39 municipalities in the Henry database. Age at death is missing for about 30 percent of wives and 40 percent of husbands. *Known birth year* is recovered from parish records or from age in the general population census (GPC). Distances are in $\log(1 + \text{km})$ units. For terrain ruggedness, the sample excludes three villages in the Alps and Pyrenees with extreme values (twice as large as the 90th percentile value). Standard deviations are in parenthesis. Standard errors are in brackets. They are clustered by municipality in Panel A and are robust-standard errors in Panel B. * $p < .10$; ** $p < .05$; *** $p < .01$.

was a large fertility gap of 0.6 children between egalitarian- and inegalitarian-inheritance areas. In contrast, we find no significant differences for 23 of the 24 individual- and municipality-level characteristics. Before the reforms, the age at death of wives and husbands and the proportion of women who died before completing their reproductive span

were not significantly different across areas with different inheritance systems. So was the probability of receiving an inheritance before or after marriage, i.e., the probability that a woman’s parents or in-laws were alive when she married. This suggests that the reform treatment was orthogonal to mortality and to economic uncertainty over the timing of inheritances—two hypothesized historical determinants of inequalitarian inheritance (Chu 1991; Grieco and Ziebarth 2015). Moreover, using a sample of 48,308 eighteenth-century children, we find that child mortality did not evolve differently after 1793 in areas with different inheritance rules (see Appendix Table A4). This alleviates the concern that, because mortality and fertility often go hand-in-hand in demographic transitions, the fertility decline after 1793 may simply reflect differences in mortality across inheritance areas. In addition, the accuracy of the data – which we capture by whether birth and marriage dates are known – is similar across areas with different inheritance rules. Finally, women in egalitarian-inheritance areas were more likely to be literate and to marry a literate man, but these differences are not statistically significant.

Panel B of Table 1 shows that municipalities with different inheritance systems were also comparable in terms of wheat prices, population density, religiosity, and in their distance to religious, judicial, fiscal, and territorial centers, political societies, rebellions, and paved roads. The only exception is that horse-post relays were about 500 meters further away in inequalitarian-inheritance areas. Finally, terrain ruggedness, average caloric suitability for post-1500 crops, and soil texture (the share of sandy soils) are balanced. This is important because climatic and soil characteristics have been linked to the origins of inequalitarian inheritance (Bertocchi 2017).²⁸

The 1793 inheritance reforms. Our identification strategy relies on a relatively rapid take-up of the reforms and on their exogeneity to fertility decisions. As discussed in Section 2.2, the historiography supports this, as the new inheritance rules were quickly upheld by family tribunals, inheritance was not a major popular grievance in the 1789 *cahiers de doléances*, and the reforms aimed at enforcing the equality principle, unifying the legal system, and consolidating support for the Revolution rather than affecting fertility. Finally, identification relies on the regional variation in the laws and customs regulating inheritance before 1793. Although inequalitarian inheritance was more prevalent in the south, there was significant local variation.²⁹ The spatial heterogeneity in pre-reform inheritance rules is reflected in the sample covered by the Henry database (Figure 3). Indeed, 47.7 percent of the 20,332 women in our sample were born in municipalities in inequalitarian-inheritance areas, which were affected by the reforms (Appendix Table A1).

²⁸We also show balance in these characteristics along the border between egalitarian and inequalitarian inheritance areas (see Appendix Table E1).

²⁹For instance, the southern administrative centers of Marmande, Meilhan, Villandraut, and Langon, despite being within a 20-kilometer radius, each had a different inheritance system: impartible rules that included (Marmande) or excluded women (Meilhan) and partible rules that included (Langon) or excluded women (Villandraut) (see Appendix Figure B7).

5 Empirical results

5.1 Difference-in-differences estimates

In this section, we analyze the effects on fertility decisions of the 1793 egalitarian inheritance reforms, which imposed equal division of assets and extended the right to inherit to women. According to Le Play’s hypothesis, we expect women who were fertile after 1793 in affected areas to limit their fertility in order to avoid fragmenting land among many heirs. In addition, the reform also potentially reduced fertility through increased women’s bargaining power in the household.

Table 2 reports estimates from Equation (1). In Panel A, the dependent variable is completed fertility, i.e., the number of children who reached age 6 ever born to a woman. We find a large negative and significant effect of the 1793 inheritance reforms on completed fertility. The coefficient on *Reformed inheritance* $\times$ *Fertile post-reform* implies that cohorts exposed to the reforms – i.e., cohorts of fertile age after 1793 in areas where inheritance was reformed – reduced their fertility by 0.5 children over their entire reproductive span. Given a sample average of 2.35, this corresponds to a 20-percent drop in fertility. The magnitude of the effect is similar to the pre-reform fertility gap of 0.6 children between egalitarian- and inegalitarian-inheritance areas (see Table 1). This suggests that the 1793 egalitarian inheritance reforms brought large areas of the country toward the low-fertility regime that had already prevailed in egalitarian-inheritance areas before the Revolution.

Coefficients are similar across all specifications. Column (1) considers a parsimonious two-way fixed-effects model with cohort and municipality fixed effects, which controls for unobservable factors affecting average differences in fertility across birth cohorts and municipalities. In Column (2), we include literacy, i.e., whether women signed their marriage certificate, which was a strong predictor of fertility (Becker et al. 2010). In Column (3), we control for the accuracy of the data through fixed effects for each of the ten form types available in the Henry database, which are based on the availability of birth and marriage dates (see Appendix Table A2). In Column (4), we further include two indicator variables for whether a woman’s father or mother were alive upon her marriage. These capture family-specific health conditions and genetic endowments transmitted from both the maternal and paternal lines. In addition, whether parents were alive at the time of marriage captures potential inheritance transfers, which could trigger income effects.

Panels B and C of Table 2 report estimates for the effect of the 1793 inheritance reforms on the intensive and extensive margins of fertility. The dependent variable is the completed fertility of mothers in Panel B and an indicator variable equal to one if a woman was childless in Panel C. As before, both variables exclude children who did not reach age 6. Our estimates suggest that the 1793 inheritance reforms reduced the

Table 2: Difference-in-differences estimates, Henry database.

	(1)	(2)	(3)	(4)
<i>Panel A. Dep. Variable is completed fertility</i>				
Reformed inheritance × Fertile post-reform	−0.492*** (0.117)	−0.493*** (0.118)	−0.475*** (0.113)	−0.520*** (0.127)
Observations	20,332	20,332	20,331	20,322
Adjusted R-squared	0.044	0.044	0.152	0.175
<i>Panel B. Dep. Variable is completed fertility of mothers</i>				
Reformed inheritance × Fertile post-reform	−0.354** (0.142)	−0.356** (0.148)	−0.371** (0.151)	−0.392** (0.160)
Observations	14,066	14,066	14,065	14,057
Adjusted R-squared	0.057	0.058	0.083	0.091
<i>Panel C. Dep. Variable is =1 if childless</i>				
Reformed inheritance × Fertile post-reform	0.067*** (0.017)	0.067*** (0.017)	0.062*** (0.013)	0.076*** (0.017)
Observations	20,332	20,332	20,331	20,322
Adjusted R-squared	0.023	0.024	0.160	0.186
Cohort FE	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y
Literacy (0/1)	.	Y	Y	Y
Accuracy of Henry form FE	.	.	Y	Y
Father alive at marriage (0/1)	.	.	.	Y
Mother — ” —	.	.	.	Y
N clusters	39	39	39	39

Notes: The sample is women born in 1700–1803 in the Henry database. The dependent variable is the number of children ever born to all women (Panel A), to mothers (Panel B), and the probability to be childless (Panel C). All variables consider “net” fertility, i.e., they are based on the number of children who reached age 6. *Literacy* is an indicator variable equal to one if the woman signed the marriage certificate. *Accuracy of Henry form* includes fixed effects for each category listed in Appendix Table A2. Standard errors in parentheses are clustered by municipality. *p<.10; **p<.05; ***p<.01.

completed fertility of mothers by 0.35–0.39 children and increased women’s likelihood of being childless within marriage by 6.2–7.6 percentage points.³⁰

³⁰In historical and developing contexts such as eighteenth-century France, childlessness within marriage is mostly the result of high child mortality and infecundity due to poverty. It rarely reflects a preference to forego motherhood or increased female empowerment (Baudin et al. 2020). Hence, the reforms’ impact on childlessness was primarily mechanical, resulting from a sudden drop in desired fertility in a context of high and unchanged child mortality (see Appendix Figure B5 and Table A4). Appendix Table A5 provides further evidence for this by showing that the effect of the 1793 inheritance reforms on childlessness remains if we restrict the sample to women who gave birth to at least one child—regardless of whether that child survived infancy or not.

Next, we show that these effects are driven by the 1793 inheritance reforms and not by heterogeneous trends in fertility across different areas. Table 3 reports the results from estimating Equation (3). This extended specification allows fertility to follow different trends in each municipality, depending on their local economic, religious, political, and economic-geography characteristics. In Column (1), we include municipality-level wheat prices by decade to account for the fertility effects of time-varying and municipality-specific economic conditions. We also include two measures of pre-reform economic conditions – wheat price (in the 1780s) and population density (in 1793) – interacted with birth-cohort fixed effects to account flexibly for trends in fertility along initial economic conditions. In Column (2), we consider the distance to Church administrative centers and the proportion of marriages during lent and advent after the introduction of civil marriage in 1792. The interaction of these variables with cohort fixed effects captures the possibility that cohorts in more religious or secular municipalities could have been on a different trajectory relative to the demographic transition. Likewise, in Column (3) we include the interaction between cohort fixed effects and the distance to political societies (in 1789–93) and to rebellion against state authorities (in the 1780s), two proxies for access to information about revolutionary reforms and local support to the Revolution before 1793. We also add distances to legal, fiscal, and territorial administrative centers prior to the Revolution, and to paved roads and horse-post relays, interacted with birth-cohort fixed effects. This accounts flexibly for institutional and economic-geography factors, as well as transport and information networks, that could affect women’s fertility over time.

Our estimates are stable across specifications. Estimates in Columns (1)–(4) imply that cohorts exposed to the 1793 egalitarian inheritance reforms reduced completed fertility by 0.43–0.52 children. This effect is very similar to that found in our baseline specification and to the estimates for the pre-reform fertility gap between areas with egalitarian and inegalitarian inheritance. Likewise, the estimated effects of the reform on the completed fertility of mothers and on childlessness also point toward a drop in fertility. Overall, our main conclusions are robust to allowing fertility to follow different trends in municipalities that differed in economic conditions, religiosity, political factors, and economic geography. This suggests that our estimates effectively reflect local changes in fertility resulting from the 1793 inheritance reforms.

Finally, Column (7) of Table 3 shows that these reductions in fertility were partly achieved by delaying age at marriage, an important preventive check in pre-industrial societies (Cinnirella et al. 2017). Specifically, the 1793 inheritance reforms are associated with an increase in age at marriage by 1.9 years. Other fertility-control strategies involved increasing the span between marriage and first birth, and reducing the span between first and last birth (see Appendix Table A3). This is consistent with previous work showing that women started limiting their fertility by not having their “last child” (Knodel 1987;

Table 3: Flexible-trend difference-in-differences estimates, Henry database.

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dep. Variable:</i>	Completed fertility	Completed fertility	Completed fertility	Completed fertility of mothers	= 1 if childless	Age at marriage
Reformed inheritance × Fertile post-reform	−0.478*** (0.174)	−0.504*** (0.159)	−0.434*** (0.154)	−0.432** (0.201)	0.046** (0.021)	1.938*** (0.572)
Observations	20,322	20,322	20,322	14,057	20,322	20,321
N clusters	39	39	39	39	39	39
Adjusted R-squared	0.178	0.179	0.182	0.103	0.190	0.255
Cohort & municipality FE	Y	Y	Y	Y	Y	Y
Individual-level controls	Y	Y	Y	Y	Y	Y
Wheat price series	Y	Y	Y	Y	Y	Y
Cohort FE						
× Wheat price 1780s	Y	Y	Y	Y	Y	Y
× Population density 1793	Y	Y	Y	Y	Y	Y
× Religiosity index	.	Y	Y	Y	Y	Y
× Distance religious center	.	Y	Y	Y	Y	Y
× Distance political society	.	.	Y	Y	Y	Y
× Distance rebellion	.	.	Y	Y	Y	Y
× Eco. geography variables	.	.	Y	Y	Y	Y
× Network variables	.	.	Y	Y	Y	Y

Notes: The sample is women born in 1700–1803 in the Henry database. The dependent variable is the number of children ever born to all women in Columns (1)–(4), to mothers in Column (5), the probability to be childless in Column (6), and age at marriage in Column (7). All variables consider “net” fertility, i.e., they are based on the number of children who reached age 6. All specifications include cohort fixed effects and the full-set of individual-level controls in Table 2. The remaining covariates capture flexible trends in fertility by municipality-level characteristics. *Economic geography variables* include the distance to legal centers, tax centers, and territorial administrative centers. *Network variables* include the distance to paved roads and horse-post relays. Wheat prices are in log and distances are in log(1 + km). Standard errors in parentheses are clustered by municipality. *p<.10; **p<.05; ***p<.01.

Cinnirella et al. 2017).³¹ In addition, we find that the fertility decline in areas affected by the reforms was not the result of increased child mortality rates (see Appendix Table A4).

Implications for the demographic transition. Taken together, our results suggest that the 1793 inheritance reforms contributed to the large fertility decline in late-eighteenth century France. The effect of the reforms on women’s completed fertility (0.43–0.52 children) is similar to the pre-reform fertility gap between areas with egalitarian and inegalitarian inheritance (0.62). This shows that the harmonization of inheritance systems during the Revolution accelerated France’s fertility transition by driving high-fertility, inegalitarian-inheritance regions (covering roughly half of the country) toward the low-fertility regime already prevailing in egalitarian-inheritance areas.

To what extent did the inheritance reforms contribute to the demographic transition? France’s fertility transition is characterized by a first phase – primarily rural – that

³¹In contrast, the reform is positively but not significantly associated with the minimum spacing between consecutive births.

spanned the eighteenth and the mid-nineteenth centuries (van de Walle 1986) and a second phase that ended by 1939 (Delventhal et al. 2021).³² We use the Henry database together with later sources – vital statistics (*Mouvement de la population*) for 1855 and the 1954 *Enquête famille* – to reconstruct how completed fertility evolved through these two phases (see Appendix C). In the first quarter of the eighteenth century (1700–25), completed fertility was 2.59 for women at the end of their reproductive span. Fertility declined to 2.13 by the end of the first phase and to 2.01 children by the end of the transition. For women who completed their reproductive span in 1793–1818, the first quarter century after the inheritance reforms, fertility was 2.16 (see Appendix Figure C1).

We then run a counterfactual exercise where we use estimates of Equation (3) with the full set of controls (Table 3, Column 3) to predict each individual’s fertility in the absence of the reforms, i.e., setting $F_c = 0$. Based on these predicted individual fertilities, we calculate the counterfactual average fertility across France. The counterfactual fertility of cohorts who completed their reproductive span in 1793–1818 would have been 2.37 children, substantially larger than the observed 2.16.³³ Altogether, these numbers imply that the 1793 inheritance reforms contributed to 46 percent of the fertility decline during the first phase of the transition and to 36 percent of the decline over the entire demographic transition.

Moreover, data from the vital statistics for 1855 and the 1954 *Enquête famille* shows that the large fertility gap that had existed before 1793 between egalitarian and inegalitarian inheritance areas had disappeared by the end of the first and second phase of the demographic transition. Indeed, the completed fertility of women in areas with different pre-reform rules was not anymore significantly different after the 1850s (see Appendix Table C1). This further suggests that the harmonization of inheritance rules in 1793 accelerated the demographic transition and fostered fertility convergence across regions of France.

5.2 Difference-in-differences robustness

We perform several robustness checks and extensions of the analysis. This section briefly describes them. Detailed results are available in Appendix D.

Permutation tests. Appendix Figure D1 reports 10,000 estimates of β in Equations (1) and (3), where we reshuffle the pre-reform inheritance system among municipalities and interact it with the “true” exposure indicator for cohorts of fertile age after 1793. In addition, we report results that, together with the pre-reform inheritance system, also reshuffle the set of municipality-level flexible trends. Estimated coefficients

³²The exact beginning of the transition within the eighteenth century is debated (Blanc 2024a). Hence, the starting point for this exercise are the earliest eighteenth-century cohorts.

³³Note that this counterfactual estimate captures the average treatment effect across France, while our DD estimates are larger, as they capture average treatment effects on the treated.

in these placebo regressions have a distribution centered around zero, and our baseline estimate is larger in magnitude than all but 0.07–0.82 percent of the estimates resulting from the permutations. This suggests that our main results are not due to random chance or to general trends in fertility, but that they effectively reflect the impact of the 1793 inheritance reforms.

Sensitivity to outliers. We perform a jackknife test to show that our results are not driven by outliers. Estimates obtained by sequentially omitting one of the 39 municipalities in the Henry database are indistinguishable from our baseline DD estimates (see Appendix Figure D2).

Placebo test. We conduct a placebo test on the cohorts that had all their children before the 1793 inheritance reforms. Estimates for the effect of the placebo reform are close to and not significantly different from zero (see Appendix Table D1). This further suggests that our baseline estimation captures the effect of the 1793 inheritance reforms and not that of pre-trends in completed fertility.

Alternative sample, treatment, and control group. Appendix Table D2 examines the robustness of our results to modifying sample and treatment definitions. In detail, our results are robust to restricting the sample to women born between 1720 and 1780. Estimates are also almost identical when we define the cohorts exposed to the reforms as those aged below 30, below 40 (baseline), or below 45 at the time of the reforms. In addition, because the 1793 inheritance reforms simultaneously abolished impartible rules and extended the right to inherit to women, we consider three alternative definitions of the treatment group based on these two aspects of the reform. First, we remove the few areas that followed partible rules excluding women, and hence, that were only partially affected by the reform. Second, we compare areas with pre-reform impartible versus partible inheritance, regardless of women’s inheritance rights (see Appendix Table D3). Third, we compare areas that excluded versus included women from inheriting, regardless of the impartibility of inheritance (see Appendix Table D4). Results are consistent, suggesting that the fertility decline in eighteenth-century France was driven by cohorts exposed to the reforms in areas where the inheritance system was altered. Finally, we account for the possibility that the treatment effect may be heterogeneous across post-reform cohorts through the event-study presented above, which shows that estimates are similar and not statistically different from each other for all cohorts with five or more reproductive years remaining after the reforms (Panel b of Figure 4).

Migration and mortality. Because the Henry database was constructed through the family reconstitution method, our estimates are potentially biased by migration patterns. Similarly, diverging mortality trends across pre-reform inheritance systems could bias our estimates. We address these concerns by showing that our results are robust to restricting the sample to women who were alive at age 40, i.e., whose records were not

missed because of emigration, and to including municipality-specific trends in mortality (see Appendix Table D5).

First-name repetition technique. We account for the possibility that children deaths are under-reported in the Henry database (Houdaille 1984) by applying the first-name repetition technique of Cummins (2020). This technique is based on the fact that parents often used the name of a deceased child to name a newborn. Our results are robust to using adjusted fertility measures (see Appendix Table D6).

Soil, climate, and terrain characteristics. Climate and the soil suitability for different crops affect land values – a hypothesized historical determinant of inequalitarian inheritance (Bertocchi 2017) – as well as average farm sizes (Bhalla 1988). In Table 1, we examine the balance of the average caloric suitability of the soil for post-1500 crops, terrain ruggedness, and the share of sandy soils. Our estimates are also robust to adding the interaction between these factors and cohort fixed effects, allowing for flexible trends in fertility along soil characteristics (see Appendix Table D7). Moreover, results are unaffected by excluding three villages in the Alps and Pyrenees with extreme ruggedness values – twice as large as the 90th percentile (see Appendix Tables D7 and D8).

Few clusters. In our baseline specification, we cluster standard errors at the municipality level. Because our sample comprises 39 municipalities, in Appendix Table D9, we report p -values from the wild bootstrap procedure suggested by Cameron et al. (2008) to address the problem of few clusters. Difference-in-differences estimates with and without individual-level controls and flexible trends remain precisely estimated and statistically significant at conventional levels for all main outcomes: completed fertility, completed fertility of mothers, childlessness, and age at marriage.

Spatially-adjusted errors. Appendix Figure D5 reports spatially-adjusted z -statistics estimated from Equations (1) and (3). We account for various degrees of spatial dependence in the error term using Conley (1999) standard errors with different distance cutoffs, from 50 to 200 kilometers, in 50-kilometer increments. That is, we allow for observations as far as 200 kilometers away to be spatially dependent. All spatially-adjusted errors are close to our baseline errors. This is consistent with the corrections proposed by Kelly (2020), who finds that spatially-adjusted standard errors tend to fall for applications which, like ours, use longitudinal data with time and space fixed effects.

5.3 Spatial RD-DD with crowdsourced genealogies

In this section, we extend our analysis by using crowdsourced genealogies from geni.com. The substantial spatial coverage in this database enables us to implement a spatial regression discontinuity design to account for unobservable factors that vary smoothly across space—e.g., soil characteristics, land values, historical experiences, the north-south divide in Roman law, or how much the rule of law prevailed in remote areas. Using this

data is also appealing because it covers the southwest of France, an area often used to question Le Play’s hypothesis. Finally, it is also helpful to validate our results with data obtained from crowdsourced genealogies, a different methodology from Henry’s family reconstruction method.

Empirical strategy. We use a regression discontinuity difference-in-differences (RD-DD) design to study the effect of the 1793 inheritance reforms on fertility. We restrict our Geni sample of mothers (1700–1810) to those born close to a border between judicial districts, where one district had egalitarian and the other inegalitarian inheritance prior to the reforms. Panel (b) of Figure 3 illustrates this setup, displaying the 1,291 birth locations of 5,640 mothers within 30 kilometers of the inheritance borders. The RD-DD strategy exploits the fact that women living close to – but on opposite sides of these borders – were subject to different inheritance rules before 1793, and to the same egalitarian rule after 1793. Specifically, we compare the fertility gap at the border for cohorts who were fertile before versus after the 1793 inheritance reforms. This strategy accounts for unobservable factors that vary smoothly across space, and differences-out time-invariant unobservables and general trends.

The basic regression-discontinuity (RD) estimating equation in our setup is:

$$Y_{icm} = \alpha + \beta \mathbf{1}[d_m \geq 0] + \mu_c + \phi_b + \mathbf{1}[d_m \geq 0] \times f_I(d_m, \gamma_I) + \mathbf{1}[d_m < 0] \times f_E(-d_m, \gamma_E) + \epsilon_{icm}, \quad (4)$$

where Y_{icm} is the completed fertility of mother i , born in cohort c , in municipality m , and d_m is the distance to the border, with positive values for inegalitarian- and negative values for egalitarian-inheritance areas. Hence, $\mathbf{1}[d_m \geq 0]$ is equal to one for the treatment group. f_I and f_E are unknown polynomial functions with parameter vectors γ_I and γ_E . They capture location-specific factors on both sides of the inheritance border that can affect fertility. We also include fixed effects for cohorts, μ_c , and for 50-kilometer border segments, $\phi_{b(m)}$. We use triangular kernel functions and polynomial fits of orders 1 (linear) and 2 (quadratic). We avoid higher-order polynomials to limit the overfitting bias (Gelman and Imbens 2019). The sample includes mothers whose Geni record satisfies the horizontal restriction explained in Section 3.3 and who were born within a bandwidth of the inheritance border. We follow Calonico et al. (2014) and use mean squared error (MSE) optimal bandwidths. We provide estimates of β , the discontinuity in completed fertility at the border, for two sub-samples. First, for cohorts who completed their reproductive span before the reforms, i.e., over 40 in 1793, we expect higher fertility in inegalitarian areas ($\beta > 0$). Second, for cohorts who were fertile after the reforms harmonized inheritance systems, i.e., below 40 in 1793, we expect fertility differences at the border to disappear ($\beta \approx 0$).

We augment this RD setup with a difference-in-differences model. Specifically, we pool the two sub-samples described above and extend Equation (4) by interacting each RD variable with our reform exposure indicator, F_c :

$$Y_{icm} = \alpha + \beta_1 \mathbf{1}[d_m \geq 0] + \beta_1 \mathbf{1}[d_m \geq 0] \times F_c + \mu_c + \phi_{b(j)} + \mathbf{Z}'_{mc} \delta_c + \sum_{s=0}^1 \mathbf{1}[F_c=s] \times \left\{ \mathbf{1}[d_m \geq 0] \times f_I(d_m, \gamma_{Is}) + \mathbf{1}[d_m < 0] \times f_E(-d_m, \gamma_{Es}) \right\} + \epsilon_{icm}. \quad (5)$$

Equation (5) is essentially a fully interacted version of Equation (4) with separate effects for cohorts above ($F_c=0$) versus below ($F_c=1$) age 40 in 1793. In extended specifications, we add fixed effects for judicial districts, $\delta_{j(m)}$ – the level at which inheritance rules varied prior to the reforms – and an analogous set of flexible trends as before, $\mathbf{Z}_{mc}$.³⁴ These allow fertility to follow different trajectories by local economic, religious, political, and economic-geography characteristics.³⁵

Our RD-DD approach requires two identifying assumptions. The first is that unobservables vary smoothly across borders. We evaluate this assumption by conducting a balancing test for the full set of covariates in the flexible-trends vector $\mathbf{Z}_{mc}$. In detail, Appendix Table E1 reports estimates of Equation (4) using as dependent variables wheat prices by municipality and decade, the municipality population density in 1793, the district-level share of refractory clergy in 1791, and indicator variables equal to one if the municipality was within 15 kilometers of, respectively, a political society, a rebellion against state authorities in the 1780s, a religious, legal, fiscal, and territorial administrative center, or a horse-post relay, and within 7.5 kilometers of a paved road.³⁶ For each of these 11 variables, the RD estimate is small and, for 10 of them, the discontinuity at the border is not statistically different from zero (see Appendix Figure E1 for the corresponding RD plots). The second assumption is the standard no pre-trends assumption. Using the full Geni database in France, as well as a sample of Geni observations around the inheritance border, we show that prior to the reforms, fertility evolved in a parallel fashion across areas under egalitarian and inegalitarian inheritance rules (see Appendix Figure E2).³⁷

Results. As a first step, we present graphical evidence on the relationship between inheritance rules and fertility. Figure 5 plots the completed fertility of mothers, excluding child deaths before age 6, around the egalitarian-inegalitarian inheritance border. Each dot represents average fertility within a bin, partialled out of cohort and border-segment

³⁴When included, $\delta_{j(m)}$ s absorb the direct effect of the treatment group indicator, $\mathbf{1}[d_m \geq 0]$.

³⁵Differently from before, we proxy for religiosity with the *district*-level share of refractory clergy who refused the oath of loyalty to the state in 1791, interacted with cohort fixed effects. In addition, because the RD-DD specification already accounts for running variables in distance, we now consider dichotomous rather than distance-based location variables.

³⁶We use 15 kilometers as it is close to the MSE optimal bandwidth on each side of the border. We use a smaller cutoff for paved roads because no observation in our border sample is further than 15 kilometers from them.

³⁷Appendix Figure E1 also suggests that fertility converged after the 1793 inheritance reforms, although the visual evidence is less obvious than in Figure 4. The main reason is that the smaller sample size in the Geni data is less suited for such a granular breakdown of fertility over time and space.

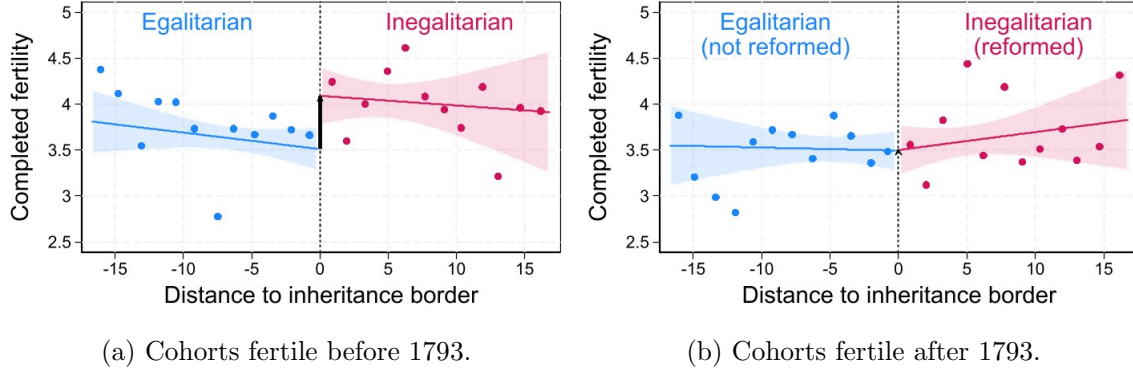


Figure 5: Fertility and distance to the inheritance border.

Notes: This figure displays RD plots. The sample is mothers whose Geni record satisfies the horizontal restriction, born in France (1700–1810) within about 17 kilometers of the inheritance border. Completed fertility is the number of births to mothers, excluding child deaths before age 6. The border is normalized at 0 with positive values for treated areas. Points show average fertility within bins, where fertility is partialled out of cohort and border-segment fixed effects, and where the 24 bins are based on the IMSE-optimal evenly-spaced selector. Lines show a polynomial fit of order 1 and shaded areas are 95-percent confidence intervals. The bandwidth is based on the MSE optimal selector (Table 4, Column 1).

fixed effects, for 24 bins based on the IMSE-optimal evenly-spaced selector.³⁸ Lines shows a linear local-polynomial fit within a MSE-optimal bandwidth. A discontinuity at the border is apparent only for the pre-reform sub-sample (Panel a). For the post-reform sub-sample who were fertile after the harmonization of inheritance rules, we find no evidence of a discontinuity at the border (Panel b). This provides some initial evidence that the 1793 inheritance reforms led to a convergence in fertility across pre-reform egalitarian- and inegalitarian-inheritance areas in France.

Table 4 presents RD estimates of β and RD-DD estimates of β_1 . Columns (1) and (2) show the result from estimating the basic RD setup from Equation (4) on mothers who completed their reproductive span before and after the reforms, respectively. Column (3) reports estimates from the RD-DD model from Equation (5) pooling these two samples, and Columns (4) and (5) add judicial district fixed effects, flexible trends, and quadratic polynomial fits. The dependent variable is the completed fertility of mothers, excluding children who did not reach age 6. All bandwidths are based on the MSE-optimal selector.

Estimates suggest a sharp discontinuity in fertility at the inheritance border of about one child, which vanishes almost entirely after the reforms harmonized inheritance rules. In detail, the RD estimate for mothers who completed their reproductive span before the 1793 inheritance reforms is 1.04, suggesting that inegalitarian inheritance is associated with a higher fertility (Column 1). In contrast, the RD estimate for mothers who were fertile after the reforms is 0.19, close to zero and not statistically significant, suggesting that fertility differences at the border disappeared after the inheritance reforms (Column 2). Pooling these two discontinuities together, the RD-DD estimate shows that

³⁸We follow the recommendations of Korting et al. (2023) to use a large number of small bins.

Table 4: Spatial regression-discontinuity estimates, Geni database.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	RD pre-reform	RD post-reform	RD-DD all	RD-DD all	RD-DD all	RD-DD all	RD-DD all
Estimate	1.044*** (0.368)	0.186 (0.354)	−0.848** (0.365)	−0.958** (0.382)	−1.187** (0.537)	−1.197*** (0.445)	−1.118** (0.541)
Observations	2,385	1,848	4,325	3,797	3,527	5,410	5,090
N clusters	585	624	959	862	778	1,214	1,070
Mean dep. variable	3.81	3.58	3.69	3.77	3.85	3.63	3.72
Cohort FE	Y	Y	Y	Y	Y	Y	Y
Border segment FE	Y	Y	Y	Y	Y	Y	Y
Judicial district FE	.	.	.	Y	Y	Y	Y
Flexible trends	.	.	.	.	Y	.	Y
Order polynomial	Linear	Linear	Linear	Linear	Linear	Quadratic	Quadratic
MSE-bandwidth (km)	19.47	19.47	19.47	16.91	17.49	29.29	30.53

Notes: This table reports estimates of β and β_1 in Equations (4) and (5). The sample is mothers whose Geni record satisfies the horizontal restriction, born in France (1700–1810) within a MSE-optimal bandwidth on each side of the inheritance border. We use local-polynomial fits of orders 1 and 2, and triangular kernel functions for local-polynomial estimation. The dependent variable is the number of children ever born to mothers, excluding child deaths before age 6. Flexible trends include municipality-level wheat prices by decade, and cohort fixed effects interacted with a vector of initial municipality-level variables: wheat price in 1780, population density in 1793, the share of refractory clergy in 1791, indicator variables if, within 15 kilometers, there was a political society, rebellions against state authorities in the 1780s, religious, legal, fiscal, territorial administrative centers, horse-post relay, or paved roads (within 7.5 kilometers). Standard errors in parentheses are clustered by municipality. * $p < .10$; ** $p < .05$; *** $p < .01$.

the 1793 inheritance reforms reduced fertility by 0.85 children for exposed cohorts (Column 3). Compared to the difference-in-difference estimates of Section 5.1, the RD-DD estimates are larger in magnitude. One reason is that, because women with no children are underreported in the Geni database, the dependent variable is by construction the completed fertility of mothers, which has a higher sample mean. In fact, compared to the sample mean, the RD-DD estimate implies a 23-percent reduction in fertility, similar to the 20-percent reduction estimated when using the Henry database. Finally, RD-DD estimates are similar across specifications including judicial district fixed effects (Columns 4–7), using linear (Columns 4–5) or quadratic polynomial fits (Columns 6–7), and allowing fertility to follow different trends by local economic, religious, political, and economic-geography factors (Columns 5 and 7).

Robustness checks. We examine the sensitivity of our RD-DD estimates to alternative specifications of Equation (5) in Appendix Table E2. In Panel A, we use two-dimensional running variables in latitude, longitude, and their interaction. In Panel B, we use as running variables the distance to the border interacted with 26 indicator variables for cohorts with 0–25 remaining fertile years after the reforms. In Panel C, we use a uniform kernel. In Panels D–F, the sample is analogous to that in the Henry database: respectively, eighteenth-century cohorts, rural municipalities, and excluding *arrondissement* centers. In Panel G, we use 100-kilometer border-segment fixed effects. The estimated effect of the reforms remains negative, similar in magnitude, and is precisely estimated in almost

all specifications. In addition, we expand and narrow the MSE optimal bandwidths by 1, 2, . . . , 5 kilometers in our linear-polynomial specifications, and by 2, 4, . . . , 10 kilometers in our quadratic-polynomial specifications. Results are robust to the choice of bandwidth (see Appendix Figure E3). We also show that RD-DD results are robust to correcting for spatial correlation in the error term (see Appendix Figure E4). As before, we do so by computing Conley (1999) standard errors and account for various degrees of spatial dependence by using distance cutoffs from 50 to 200 kilometers.

6 Mechanisms

6.1 Land fragmentation

The prevailing explanation for the fertility decline following the 1793 inheritance reforms is that families had fewer children to avoid the progressive fragmentation of their landholdings (Le Play 1875). In detail, when land is a key asset for production but parcels are small and fragmented – as was the case in eighteenth-century France (Bourdieu et al. 2013) – subdividing land plots among multiple heirs can reduce their capacity to sustain livelihoods (Habakkuk 1955).³⁹ The 1793 inheritance reforms, both by abolishing impartibility and by extending the right to inherit to women, increased the number of heirs.⁴⁰ Hence, it intensified the incentives to have fewer children to avoid dividing the land among sons and daughters to the point where production would fall below subsistence levels. This section provides qualitative and quantitative evidence suggesting that this mechanism played an important role.

Qualitative evidence. Several contemporaries of the reforms highlighted the close relationship between inheritance rules and land fragmentation at the core of this mechanism (Beckerts 2008, pp. 29–32). For instance, de Cazalès, a representative in the National Constituent Assembly (1798–91), opposed inheritance reforms arguing that “through the effect of necessarily equal partitions from generation to generation, property becomes so divided [...] that it can no longer feed its owner.”⁴¹ This echoes the claims of Normandy’s representative Lambert de Frondeville that the reforms, both by imposing equal inheritance and extending inheritance rights to women, could exacerbate land fragmentation, as “[t]he inequality of inheritance [...] and the exclusion of daughters from land division aim at the success of agriculture.”⁴² Some case-study analysis suggests that the egalitarian inheritance reforms increased the fragmentation of landholdings, with many

³⁹Habakkuk (1955) argued that “the regions of [inheritance] division were regions of [...] small peasants, anxious to add acre to acre [whose] savings [...] were not used to improving their properties but to extending them” (p. 11).

⁴⁰In fact, with an average net completed fertility of about 3 children in treated areas before the reform – 1.5 of each gender – including daughters in inheritances increased the number of heirs by as much as abolishing impartible inheritance.

⁴¹*Archives Parlementaires de 1787 à 1860*, Tome XXIV, pp. 572–4 (April 5, 1791).

⁴²*Archives Parlementaires de 1787 à 1860*, Tome XXIV, pp. 48–50 (March 12, 1791).

farmers owning several scattered parcels, each just a few hectares (Sagnac 1903, p. 465; Moriceau 2002, pp. 139–41). For instance, de Serviez (1801, pp. 125–6), describing the département of Basses-Pyrénées, noted that the division of land was “disastrous in this mountainous and barren region” and that “the children sharing the inheritance would be unable to make their lots productive, and the subdivision in the second generation would plunge all their descendants into indigence.”⁴³

Notwithstanding these local case studies, no study to date has quantified the implications of the 1793 inheritance reforms for the aggregate distribution of landholdings in France. The reason is that data on landownership before the Revolution only exists for a very limited sample of towns (see Appendix F.2). That said, aggregate agricultural productivity remained low but stable throughout the Napoleonic era (Newell 1973), suggesting that the inheritance reforms did not deteriorate agricultural production as dramatically as anticipated. One possibility is that, despite the egalitarian laws of 1793, the subsequent fertility decline slowed down the dilution of landholdings below subsistence levels. Indeed, the opponent of the reform de Cazalès argued that fertility would be reduced to avert land fragmentation: “[t]his equal share that one would be obliged to give to their younger siblings might even prevent them from being born.”⁴⁴

De Cazalès’ statement foreshadowed Le Play’s hypothesis linking inheritance reforms, land fragmentation, and fertility decline. Specifically, Le Play argued that the “forced division” of inheritance among sons and daughters had three social consequences: it reduced fertility, hindered economic development by fragmenting land, and weakened paternal authority. The 1793 inheritance reforms turned extended families into nuclear units, which had to “seek in the sterility of marriage a way to avoid the fragmentation of small properties and the unfortunate consequences that result from it” (Le Play 1875, p. 314). Since larger estates are less subject to such indivisibility constraints, equal division disproportionately harmed small landowners (Le Play 1875, pp. 331, 355).

Altogether, the historical evidence suggests that the 1793 inheritance reforms, both by imposing equal partition and by extending the right to inherit to women, created strong incentives to limit fertility so as to prevent land fragmentation and production from falling below subsistence levels.

Quantitative evidence from cadastre data. We provide further support for this narrative through two quantitative exercises. First, we digitize part of the Napoleonic cadastre (1807–47), a source that, to the extent of our knowledge, remains unexplored by economists. This is the first available source that allows to measure land fragmentation

⁴³Case studies on the effect of the reforms for land fragmentation exist for Basses-Pyrénées, Orne, the Nivernais, and Franche-Comté—see the sources in Appendix Table A6 and Green (2024) for a broader discussion. More generally, Finley et al. (2021) and Deseau (2023) highlight the role of land indivisibilities by showing that the redistribution of monastic lands increased agricultural productivity, in part by reducing land fragmentation.

⁴⁴*Archives Parlementaires de 1787 à 1860*, Tome XXIV, pp. 572–4 (April 5, 1791).

systematically. Because the cadastre was collected only a few decades after 1793, the land distribution documented therein still (partly) reflects the incidence of pre-reform inheritance rules. Specifically, we collect information for 36 thousand parcels of land in a subset of nine municipalities in the Henry database for which cadastral data is available: four with pre-reform egalitarian inheritance and five with pre-reform inegalitarian inheritance.

Armed with this data, we document three important facts (see Appendix F.2 for details). The first fact is that productive land was very fractionalized in early nineteenth-century France. Across municipalities, the median parcel was below 0.1 hectare. This is consistent with the narrative that land fragmentation was a serious concern for most farmers in France. The second fact is a strong association between inheritance rules and land fragmentation: under pre-reform egalitarian (inegalitarian) inheritance, the average parcel size was 0.16 hectares (2.33), the Gini index, 0.65 (0.91), and the share of very small parcels – less than 0.1 hectare – in a municipality, 61 (36) percent. A Kolmogorov-Smirnov test confirms that the parcel-size distribution under inegalitarian inheritance first-order stochastically dominated that under egalitarian inheritance. The third fact is that, a few decades after the 1793 inheritance reforms, the parcel-size distribution had not fully converged, which is consistent with the idea that fertility adjustments effectively slowed down land fragmentation in treated areas.

Heterogeneity analysis. For our second quantitative exercise, we construct a simple model that rationalizes the link between inheritance, land fragmentation, and fertility choices. We then use this model to derive testable predictions, which we bring to the data. Specifically, we show that the effects of the 1793 inheritance reforms were confined to areas where soil characteristics favored small farms, that is, where land fragmentation could have dire consequences.

We present the details of our model in Appendix G. In short, our model holds that households make fertility decisions under “warm glow” altruism whereby they care about their own consumption and their children’s endowments (de la Croix and Doepke 2003). Production is determined by land and labor under a Stone-Geary technology with a minimum land input threshold for subsistence (Beattie and Aradhyula 2015). Because land markets are incomplete (Finley et al. 2021), access to land is determined by inheritance: under inegalitarian inheritance, a single heir receives all the land and his siblings are employed as laborers in an extended family household. Under egalitarian inheritance, land is divided equally among siblings and each household operates as a nuclear unit. The model shows that households limit their fertility under egalitarian inheritance to avoid it falling below the subsistence threshold and that the egalitarian-inegalitarian fertility gap is larger under more fragmented landownership. Hence, a testable prediction of the model is that the 1793 inheritance reforms reduced fertility relatively more where households were endowed with small farms.

We test this prediction of the model by leveraging two soil characteristics which arguably provide exogenous variation in farm size: the geological composition of the soil (soil texture) and terrain ruggedness. Soil texture is based on the relative proportions of sand, silt, and clay particles. Sandy soils do not retain well storm water (Leeper and Uren 1993). Therefore, historically these soils were settled relatively late and were subject to a lower demand for land, which resulted in larger farm sizes.⁴⁵ In addition, terrain ruggedness is associated with small farms in mid nineteenth-century France (Montalbo 2023). The reason is that rugged terrains are harder to exploit and historically “deterred local elites from cornering large proportions of land” (Montalbo 2023, p. 220).

We corroborate these relationships using data from the 1852 agricultural census. Consistent with previous studies in agricultural economics, we find that sandier soils and flatter terrains are associated with larger farms (see Appendix F.3). Moreover, using soil characteristics is appealing for two reasons. First, while data on farm size is not available before the Revolution, soil texture and terrain ruggedness provide a proxy that is largely exogenous and time-invariant, as these soil characteristics cannot be altered by human intervention.⁴⁶ Second, these soil characteristics also affect land productivity. That said, they do so in opposite directions: flat terrains are generally of better quality than sandy soils, even if both are associated with large farms. Hence, by using these two distinct proxies for farm size, we address concerns that the differential effects of the reform may be driven by variation in land quality.⁴⁷

Next, we investigate how the effects of the 1793 inheritance reforms differed along farm size. Based on our theory, the reforms should have stronger effects where soil characteristics favor small farms (non-sandy soils and rugged terrains) than where they favor large farms (sandy soils and flat terrains). To test this, we extend Equation (3) to allow for heterogeneous effects along these characteristics. Specifically, we estimate a cell-means model where we omit the constant and interact our main variable of interest, $I_m \times F_c$, with indicator variables for soil conditions favoring large versus small farms in each municipality. We also account for lower-order effects by including these soil conditions (soil sandiness and terrain ruggedness) interacted with cohort fixed effects.⁴⁸

Table 5 reports estimates obtained using the Henry database. In Columns (1) and (2), we compare the effects across municipalities with sandy and non-sandy soils based on the classification from the soil geographical database of France (see Appendix F.3).⁴⁹ In

⁴⁵Indeed, the association between sandy soils and larger farms is an empirical regularity in agricultural economics across space and time; e.g., in nineteenth-century Prussia (Cinnirella and Hornung 2016) and England (Clark and Gray 2014; Goñi 2023), 1945 Italy (Martinelli and Pellegrino 2024), and modern-day India (Bhalla 1988).

⁴⁶Importantly, these characteristics are balanced across pre-reform inheritance areas (see Table 1).

⁴⁷Moreover, Montalbo (2023) shows that ruggedness was not associated with literacy, urbanization, industrial production, or agriculture mechanization in mid nineteenth-century France.

⁴⁸The municipality fixed effects absorb both the direct effect of the I_m and F_c indicator variables, as well as the lower-order interaction between I_m and soil conditions favoring large versus small farms.

⁴⁹One in four municipalities in the Henry database falls under the sandy soil category.

Table 5: Heterogeneous effects by soil conditions for small versus large farms.

	(1)	(2)	(3)	(4)
<i>Dep. Variable: Completed fertility</i>	Soil conditions based on:			
	Texture		Ruggedness	
Reform $\times$ Fertile (soil conditions for large farms)	−0.015 (0.220)	−0.093 (0.262)	−0.133 (0.218)	0.063 (0.216)
Reform $\times$ Fertile (soil conditions for small farms)	−0.505*** (0.144)	−0.456*** (0.150)	−0.575*** (0.204)	−0.614*** (0.194)
Differential effect	0.490	0.363	0.442	0.677
<i>p</i> -value difference	0.031	0.218	0.097	0.008
Cohort and municipality FE	Y	Y	Y	Y
Individual-level controls	Y	Y	Y	Y
Flexible trends	Y	Y	Y	Y
Texture $\times$ cohort FE	.	Y	.	Y
Ruggedness $\times$ cohort FE	.	Y	.	Y
Observations	20,322	20,322	20,322	20,322
Clusters	39	39	39	39

Notes: In Columns (1) and (2), soil conditions are based on texture, where sandy soils favor large farms and non-sandy soils favor small farms. In Columns (3) and (4), soil conditions are based on ruggedness, where flat terrains favor large farms and rugged terrains favor small farms. The sample is women born in 1700–1803 in the Henry database. Individual-level controls and flexible trends are defined as in Table 3. Columns (2) and (4) account for lower-order interactions by including soil sandiness and terrain ruggedness interacted with cohort fixed effects. Constants are omitted. Standard errors in parentheses are clustered by municipality. * $p < .10$; ** $p < .05$; *** $p < .01$.

Columns (3) and (4), we compare municipalities above and below median ruggedness (Nunn and Puga 2012). Odd columns consider our baseline set of controls, even columns control for lower-order interactions as explained above. Across specifications, estimates suggest that the 1793 inheritance reforms reduced completed fertility in areas where soil characteristics favored small farms. In these areas, the reform is associated with a drop in completed fertility by 0.46–0.61 children for exposed cohorts. In contrast, the reforms had negligible effects where soil characteristics favored larger farms, and hence, where the threat of excessive land fragmentation was weaker. There, the estimates are close to zero and statistically insignificant. For comparison, relative to the sample mean, the reforms reduced fertility by 20–26 percent in areas with soil conditions favoring small farms, and by up to 6 percent in areas with soil conditions favoring large farms. A test for the equality of both coefficients can be rejected at conventional levels except in the specification of Column (2).⁵⁰

In Appendix Table A7, we expand this analysis by looking at heterogeneous effects by four major soil texture categories: coarse (sandy), medium, medium-fine, and fine soils.⁵¹

⁵⁰The sample includes three villages in the Alps and Pyrenees with extreme ruggedness. When dropping these, we obtain coefficients of −0.56 to −0.74 for rugged terrains and of −0.00 to 0.04 for flat terrains, along with a p -value on the difference of 0.004 to 0.07.

⁵¹See Appendix Figure F6 for definitions of these categories.

For all soil types, we find a large, negative effect of the 1793 inheritance reforms, with the sole exception of sandy soils—the texture associated with large farms in the literature. We also show that the completed fertility of mothers and childlessness were affected more strongly by the 1793 inheritance reforms where soil conditions favored small farms (see Appendix Table A8). Finally, we find consistent results when conducting an analogous heterogeneity analysis using our RD-DD design with the Geni database (see Appendix Table A9).

Taken together, the empirical evidence is consistent with the idea that, after the 1793 inheritance reforms, families faced a powerful incentive to curbe their fertility to avoid land fragmentation, especially where small farms were prevalent.

6.2 Women’s bargaining power

So far, we have shown evidence suggesting that the 1793 inheritance reforms, both by abolishing impartible rules and by extending the right to inherit to women, created strong incentives to limit fertility so as to prevent land fragmentation among sons and daughters. A second, complementary mechanism is women’s empowerment (Hazan et al. 2022). By extending inheritance rights to women, the reforms may have improved their outside option in the marriage market and their bargaining position within the household, which in turn could have reduced their fertility (Baudin et al. 2015). The ideal experiment to disentangle these two mechanisms – women’s bargaining power and land fragmentation – would be to estimate the effect of extending inheritance rights to women where the reform did not change the number of heirs. Unfortunately, this is not possible, as areas where the reform extended inheritance to women also increased the number of heirs. Instead, we provide qualitative and quantitative evidence to assess whether women’s bargaining power increased after 1793.

Qualitative evidence. Historical evidence suggests that the French Revolution, despite granting women access to inheritance in 1793, did not empower wives. This is mainly because the 1804 Civil Code reformed marital property law, granting husbands absolute control over jointly or independently owned property, including assets inherited by women (Desan 2006, pp. 300–1). Indeed, the Civil Code effectively declared women to be civilly incapable, stating that “the wife owes obedience to her husband” (art. 213). This is in line with opinions expressed by the Revolutionaries in 1793 or by Napoleon in 1801, who declared that “[husbands] must have absolute power over the conduct of their wives” (quoted in Tudor 2021, p. 333). It was not until 1985 that French women gained equal control over marital property. Hence, without control over their assets, it is unlikely that women – despite gaining the right to inherit in 1793 – also gained bargaining power in the couple’s decisions on, for instance, fertility. Similarly, the impact of the 1793 inheritance reforms on women’s outside options in the marriage market was limited by the nature of divorce law. Although equal divorce rights were first granted in 1792,

women’s right to remarry was restricted in 1793, and the Civil Code further hindered women’s ability to file for divorce (Desan 2006, pp. 326–8). This was part of a broader policy aimed at creating a new society in which the so-called “republican mother” had no rights in public life and limited power in the household. According to the majority of revolutionary legislators, a woman’s role in society was to bear and raise children. From 1793, the newly established public schools educated women in these values throughout France (Fayolle 2022, pp. 31, 37–40). Similarly, women’s political rights were also limited. From 1789, they were ineligible to vote, and from 1793, they were prohibited from participating in political associations (Tudor 2021, pp. 191, 231).

Taken together, husbands’ control over inherited assets, the civil incapability of married women, education policies, and political exclusion all contributed to diminishing women’s status, limiting the bargaining power that the new inheritance laws could have given them.

Quantitative evidence. To further evaluate this conjecture, we estimate the relationship between the inheritance reforms and spouses’ relative age and education—two common distribution factors which determine women’s bargaining power within the household (Browning et al. 2014, Table 5.1, p. 204). If the 1793 egalitarian inheritance reforms empowered women and improved their bargaining position, we expect exposed cohorts to marry more similar spouses in terms of age and education. Instead, we expect null effects if the concomitant reforms on marital property law, divorce, or education prevented women from gaining bargaining power despite their newly-acquired inheritance rights. Estimates based on Equation (3) for the effect of the reforms on spouses’ relative age and education are close to and not significantly different from zero (see Appendix Table A10). Compared to the 20-percent reduction in fertility, effect sizes relative to the mean are substantially smaller (2 and 6.6 percent, respectively).

In summary, the historical and quantitative evidence suggest that women did not gain bargaining power within the household after the French Revolution. That said, extending the right to inherit to women likely reduced fertility by roughly doubling the number of heirs, and hence, by creating strong incentives to limit fertility so as to prevent land fragmentation among sons and daughters.

7 Conclusion

The revolutionary change in inheritance laws in 1793 was one of the causes of the French demographic transition. France was the first country to experience a demographic transition, at least a century before any other European country. We show that legal institutions crucially reduced the economic incentives for having children. Egalitarian inheritance rules imposed during the French Revolution had a strong causal effect on the fertility of affected regions. This effect remains even after controlling for other potential determinants of the fertility decline, such as human capital, secularization, changes in

economic conditions, distance to administrative centers or information networks, exposure to political pressure or rebellions linked to the Revolution. Results are also robust to using data obtained from very different methodologies: the family reconstruction method and crowdsourced genealogies.

France’s demographic transition was not only the first but was also among the longest in the world. Hence, multiple factors beyond inheritance contributed to the decline in fertility, either at the beginning, or at later stages of the transition. Yet, traditional explanations for the demographic transition in the literature cannot be reconciled with key features of the French case: its early timing is at odds with theories based on industrialization, human capital, and the quantity-quality trade-off, while the sharp fertility decline observed in the late eighteenth century cannot be rationalized with slowly evolving cultural norms. Changes in inheritance rules had been seen as a plausible driver of fertility decline since [Le Play \(1875\)](#) first put forward his theory, but empirical support to prove it was lacking. Our findings suggest that the 1793 inheritance reforms played a significant role in the French demographic transition, accounting for 36 percent of the aggregate fertility decline. By providing the first empirical evidence to support this long-standing hypothesis, we place legal institutions at the forefront of the puzzle, offering an explanation that aligns with both the early timing and the abrupt nature of the fertility decline in late eighteenth-century France.

Our results may also have important implications for the diffusion of the demographic transition from France to the rest of Europe. The Napoleonic invasions contributed to the propagation of egalitarian inheritance laws devised by the French revolutionaries to neighboring countries. In fact, by the 1850s, most European countries had introduced egalitarian inheritance laws. Whether this was also responsible for their demographic transitions remains an open question.

Beyond providing support for Le Play’s theory of fertility decline, this article unveils a potentially important contributor to modern fertility transitions: legal institutions. Legal factors have been overlooked as potential determinants of fertility decline ([Doepke et al. 2022](#)). Our finding that legal institutions regulating inheritance can have substantial effects on fertility may be relevant for boosting fertility transitions in developing countries, in particular for those experiencing stalls. The extent to which inheritance reforms toward more equality can help grasping the benefits of a demographic dividend in developing countries is an intriguing question for future research.

References

- Ager, P., B. Herz, and M. Brueckner. 2020. “Structural Change and the Fertility Transition.” *Review of Economics and Statistics* 102 (4): 806–22.
- Alburez-Gutierrez, D., T. Baudin, M. Curtis, P. E. Gobbi, and S. Moriconi. 2023. “Learning from Our Ancestors: Using Crowdsourced Historical Data for Empirical Research.” Manuscript.

- Aldashev, G., I. Chaara, J.-P. Platteau, and Z. Wahhaj. 2012. "Formal Law as a Magnet to Reform Custom." *Economic Development and Cultural Change* 60 (4): 795–828.
- Anderson, S., and G. Genicot. 2015. "Suicide and Property Rights in India." *Journal of Development Economics* 114:64–78.
- Bahrami-Rad, D. 2021. "Keeping it in the Family: Female Inheritance, Inmarriage, and the Status of Women." *Journal of Development Economics* 153:102714.
- Bar, M., and O. Leukhina. 2010. "Demographic Transition and Industrial Revolution: A Macroeconomic Investigation." *Review of Economic Dynamics* 13 (2): 424–51.
- Bartels, C., S. Jäger, and N. Obergruber. 2024. "Long-Term Effects of Equal Sharing: Evidence from Inheritance Rules for Land." *Economic Journal* 134 (664): 3137–72.
- Baudin, T., D. de la Croix, and P. E. Gobbi. 2015. "Fertility and Childlessness in the United States." *American Economic Review* 105 (6): 1852–82.
- . 2020. "Endogenous Childlessness and Stages of Development." *Journal of the European Economic Association* 18 (1): 83–133.
- Baudin, T., and R. Stelter. 2022. "The Rural Exodus and the Rise of Europe." *Journal of Economic Growth* 27:365–414.
- Beach, B., and W. Hanlon. 2022. "Culture and the Historical Fertility Transition." *The Review of Economic Studies* 90 (4): 1669–700.
- Beattie, B. R., and S. Aradhyula. 2015. "A Note on Threshold Factor Level(s) and Stone-Geary Technology." *Journal of Agricultural and Applied Economics* 47 (4): 482–93.
- Becker, G. S., and G. H. Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy* 81 (2): S279–88.
- Becker, S. O., F. Cinnirella, and L. Woessmann. 2010. "The Trade-Off Between Fertility and Education: Evidence from Before the Demographic Transition." *Journal of Economic Growth* 15 (3): 177–204.
- Beckerts, J. 2008. *Inherited Wealth*. Princeton: Princeton University Press.
- Bertocchi, G. 2017. "Primogeniture." In *The Wiley-Blackwell Encyclopedia of Social Theory*, edited by B. S. Turner. New York: John Wiley & Sons.
- Bertocchi, G., and M. Bozzano. 2015. "Family Structure and the Education Gender Gap: Evidence from Italian Provinces." *CESifo Economic Studies* 61 (1): 263–300.
- Bhalla, S. S. 1988. "Does Land Quality Matter?: Theory and Measurement." *Journal of Development Economics* 29 (1): 45–62.
- Bhattacharya, J., and S. Chakraborty. 2012. "Fertility Choice Under Child Mortality and Social Norms." *Economics Letters* 115 (3): 338–41.
- Blanc, G. 2024a. "The Cultural Origins of the Demographic Transition in France." Manuscript.
- . 2024b. "Demographic Transitions, Rural Flight, and Intergenerational Persistence: Evidence from Crowdsourced Genealogies." Manuscript.
- Blanc, G., and R. Wacziarg. 2020. "Change and Persistence in the Age of Modernization: Saint-Germain-d'Anxure, 1730–1895." *Explorations in Economic History* 78:101352.
- Boldrin, M., M. De Nardi, and L. E. Jones. 2015. "Fertility and Social Security." *Journal of Demographic Economics* 81 (3): 261–99.
- Bourdieu, J., L. Kesztenbaum, and G. Postel-Vinay. 2013. *L'enquête TRA: histoire d'un outil, outil pour l'histoire*. Paris: INED.
- Bourdot de Richebourg, C.-A. 1724. *Nouveau coutumier général ou Corps des coutumes générales et particulières de France et des provinces connues sous le nom des Gaules*. Paris: Michel Brunet.

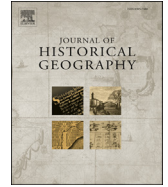
- Brée, S., and D. de la Croix. 2019. "Key Forces Behind the Decline of Fertility: Lessons from Childlessness in Rouen Before the Industrial Revolution." *Cliometrica* 13 (1): 25–54.
- Brette, A. 1904a. *Atlas des bailliages ou juridictions assimilées ayant formé unité électorale en 1789 dressé d'après les actes de la convocation conservés aux Archives nationales*. Paris: Imprimerie nationale.
- . 1904b. *Recueil de documents relatifs à la convocation des États généraux de 1789. Tome troisième*. Paris: Imprimerie nationale.
- . 1915. *Recueil de documents relatifs à la convocation des États généraux de 1789. Tome quatrième*. Paris: Imprimerie nationale.
- Browning, M., P.-A. Chiappori, and Y. Weiss. 2014. *Economics of the Family*. Cambridge: Cambridge University Press.
- Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6): 2295–326.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* 90 (3): 414–427.
- Casari, M., M. Lisciandra, and C. Tagliapietra. 2019. "Property Rights, Marriage, and Fertility in the Italian Alps, 1790–1820." *Explorations in Economic History* 71:72–92.
- Cavalcanti, T., G. Kocharkov, and C. Santos. 2020. "Family Planning and Development: Aggregate Effects of Contraceptive Use." *Economic Journal* 131 (634): 624–57.
- Cervellati, M., and U. Sunde. 2015. "The Economic and Demographic Transition, Mortality, and Comparative Development." *American Economic Journal: Macroeconomics* 7 (3): 189–225.
- Chénon, E. 1926. *Histoire générale du droit français public et privé des origines à 1815. Volume 1*. Paris: Recueil Sirey.
- Chesnais, J.-C. 1992. *The Demographic Transition*. Oxford: Oxford University Press.
- Chu, C. Y. C. 1991. "Primogeniture." *Journal of Political Economy* 99 (1): 78–99.
- Cinnirella, F., and E. Hornung. 2016. "Landownership Concentration and the Expansion of Education." *Journal of Development Economics* 121:135–52.
- Cinnirella, F., M. Klemp, and J. Weisdorf. 2017. "Malthus in the Bedroom: Birth Spacing as Birth Control in Pre-Transition England." *Demography* 54 (2): 413–36.
- Clark, G. 2007. *A farewell to alms: a brief economic history of the world*. Princeton: Princeton University Press.
- Clark, G., and R. Gray. 2014. "Geography is not Destiny: Geography, Institutions and Literacy in England, 1837–63." *Oxford Economic Papers* 66 (4): 1042–69.
- Conley, T. G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45.
- Cummins, N. 2013. "Marital Fertility and Wealth During the Fertility Transition: Rural France, 1750–1850." *The Economic History Review* 66:449–76.
- . 2020. "The Micro-Evidence for the Malthusian System. France, 1670–1840." *European Economic Review* 129:103544.
- Daudin, G., R. Franck, and H. Rapoport. 2019. "Can Internal Migration Foster the Convergence in Regional Fertility Rates? Evidence from 19th Century France." *Economic Journal* 129 (620): 1618–92.
- de Brandt, A. 1901. *Droit et coutumes des populations rurales de la France en matière successorale*. Paris: Larose.
- de la Croix, D., and M. Doepke. 2003. "Inequality and Growth: Why Differential Fertility Matters." *American Economic Review* 93 (4): 1091–113.

- de la Croix, D., and F. Perrin. 2018. “How Far Can Economic Incentives Explain the French Fertility and Education Transition?” *European Economic Review* 108:221–45.
- de la Croix, D., and M. Vander Donckt. 2010. “Would Empowering Women Initiate the Demographic Transition in Least Developed Countries?” *Journal of Human Capital* 4 (2): 85–129.
- Delventhal, M. J., J. Fernández-Villaverde, and N. Guner. 2021. “Demographic Transitions Across Time and Space.” NBER Working Paper 29480.
- Desan, S. 2006. *The Family on Trial in Revolutionary France*. Berkeley: University of California Press.
- Deseau, A. 2023. “Land Reform and Productivity: Evidence from the Dissolution of the French Monasteries.” LIDAM Discussion Papers IRES 2023009.
- de Serviez, E. G. R. 1801. *Statistique du département des Basses-Pyrénées*. Paris: Imprimerie des Sourds-Muets.
- de Tocqueville, A. 1856. *L’Ancien Régime et la Révolution*. Paris: Michel Lévy Frères.
- Doepke, M., A. Hannusch, F. Kindermann, and M. Tertilt. 2022. “The Economics of Fertility: A New Era.” In *Handbook of Family Economics*, edited by S. Lundberg and A. Voena, 151–254. Amsterdam: Elsevier.
- Fayolle, C. 2022. *La femme nouvelle. Genre, éducation, Révolution (1789–1830)*. Paris: CTHS.
- Finley, T., R. Franck, and N. D. Johnson. 2021. “The Effects of Land Redistribution: Evidence from the French Revolution.” *The Journal of Law and Economics* 64 (2): 233–67.
- Fleury, M., and L. Henry. 1958. “Pour connaître la population de la France depuis Louis XIV. Plan de travaux par sondage.” *Population* 13 (4): 663–86.
- Galasso, V., and P. Profeta. 2018. “When the State Mirrors the Family: The Design of Pension Systems.” *Journal of the European Economic Association* 16 (6): 1712–63.
- Galor, O. 2012. “The Demographic Transition: Causes and Consequences.” *Cliometrica* 6 (1): 1–28.
- Galor, O., and O. Moav. 2002. “Natural Selection and the Origin of Economic Growth.” *The Quarterly Journal of Economics* 117 (4): 1133–91.
- Galor, O., and D. N. Weil. 2000. “Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond.” *American Economic Review* 90 (4): 806–28.
- Ganghofer, R., and P. Levresse. 1977. *Le droit romain en Alsace du XIIIe au XVIe siècle*. Milan: Giuffrè.
- Gay, V., P. Gobbi, and M. Goñi. 2024a. “The Atlas of Local Jurisdictions of Ancien Régime France.” *Journal of Historical Geography* 84:49–60.
- Gay, V., P. E. Gobbi, and M. Goñi. 2024b. “The Customary Atlas of Ancien Régime France.” *Explorations in Economic History* 93:101588.
- Gelman, A., and G. Imbens. 2019. “Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs.” *Journal of Business & Economic Statistics* 37 (3): 447–56.
- Gilissen, J. 1979. *Introduction historique au droit*. Bruxelles: Bruylant.
- Gobbi, P. E., and M. Goñi. 2020. “Childless Aristocrats: Inheritance and the Extensive Margin of Fertility.” *Economic Journal* 131 (637): 2089–118.
- Goñi, M. 2023. “Landed Elites and Education Provision in England: Evidence from School Boards.” *Journal of Economic Growth* 28 (1): 125–71.
- Goy, J. 1988. “Transmission successorale et paysannerie pendant la Révolution française: un grand malentendu.” *Études rurales* 110–2:45–56.

- Green, N. 2024. “Longing for the Beheaded Father Inheritance and Departmental Statistics under the Directory and the Consulate, 1795–1804.” *French Historical Studies* 47 (1): 71–102.
- Grieco, P. L. E., and N. L. Ziebarth. 2015. “Unigeniture in an Uncertain World.” *Cliometrica* 9 (2): 139–66.
- Grinberg, M. 2006. *Écrire les coutumes. Les droits seigneuriaux en France*. Paris: Presses universitaires de France.
- Guinnane, T. W. 2011. “The Historical Fertility Transition: A Guide for Economists.” *Journal of Economic Literature* 49 (3): 589–614.
- Habakkuk, J. 1955. “Family Structure and Economic Change in Nineteenth-Century Europe.” *The Journal of Economic History* 15 (1): 1–12.
- Harari, M. 2019. “Women’s Inheritance Rights and Bargaining Power: Evidence from Kenya.” *Economic Development and Cultural Change* 68 (1): 189–238.
- Hazan, M., D. Weiss, and H. Zoabi. 2022. “Women’s Liberation and the Demographic Transition.” CEPR Discussion Paper 16838.
- Herzer, D., H. Strulik, and S. Vollmer. 2012. “The Long-Run Determinants of Fertility: One Century of Demographic Change 1900–1999.” *Journal of Economic Growth* 17 (4): 357–85.
- Houdaille, J. 1984. “La mortalité des enfants dans la France rurale de 1690 à 1779.” *Population* 39 (1): 77–106.
- Hyslop, B. 1934. *French Nationalism in 1789 According to the General Cahiers*. New York: Columbia University Press.
- Joignon, L. 1989. “Coutumes, familles, successions et alliances en Lorraine: 1670–1900.” Ph.D. diss., Paris: EHESS.
- Kelly, M. 2020. “Understanding Persistence.” CEPR Discussion Paper 15246.
- Knodel, J. 1987. “Starting, Stopping, and Spacing during the Early Stages of Fertility Transition: The Experience of German Village Populations in the 18th and 19th Centuries.” *Demography* 24 (2): 143–62.
- Korting, C., C. Lieberman, J. Matsudaira, Z. Pei, and Y. Shen. 2023. “Visual Inference and Graphical Representation in Regression Discontinuity Designs.” *Quarterly Journal of Economics* 138 (3): 1977–2019.
- La Ferrara, E., and A. Milazzo. 2017. “Customary Norms, Inheritance, and Human Capital: Evidence from a Reform of the Matrilineal System in Ghana.” *American Economic Journal: Applied Economics* 9 (4): 166–85.
- Lataste, L., L. Claveau, C. Pionnier, A. J. Ducom, and M. Bouloiseau. 1901. *Archives parlementaires de 1787 à 1860. Recueil complet des débats législatifs et politiques des chambres françaises*. Volume 59. Paris: Société d’imprimerie et Librairie administratives et des chemins de fer.
- Le Bris, D. 2019. “Testing Legal Origins Theory Within France: Customary Laws Versus Roman Code.” *Journal of Comparative Economics* 47 (1): 1–30.
- Leeper, G. W., and C. Uren, Nicholas. 1993. *Soil Science: An Introduction*. Melbourne: Melbourne University Press.
- Le Play, F. 1875. *L’Organisation de la famille selon le vrai modèle signalé par l’histoire de toutes les races et de tous les temps*. Tours: A. Mame et Fils.
- Martinelli, P., and D. Pellegrino. 2024. “Land Inequality and Long-Run Growth: Evidence from Italy.” *Quaderni di Storia Economica* 52.
- Melki, M., H. Rapoport, E. Spolaore, and R. Wacziarg. 2024. “Cultural Remittances and Modern Fertility.” NBER Working Paper 32990.

- Montalbo, A. 2023. "Landholding Inequality and the Consolidation of Democracy: Evidence from Nineteenth-Century France." *The Journal of Economic History* 83 (1): 203–41.
- Moriceau, J.-M. 2002. *Terres mouvantes*. Paris: Fayard.
- Murphy, T. E. 2015. "Old Habits Die Hard (Sometimes)." *Journal of Economic Growth* 20 (2): 177–222.
- Murtin, F. 2013. "Long-Term Determinants of the Demographic Transition, 1870–2000." *Review of Economics and Statistics* 95 (2): 617–31.
- Newell, W. H. 1973. "The Agricultural Revolution in Nineteenth-Century France." *The Journal of Economic History* 33 (4): 697–731.
- Nunn, N., and D. Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94 (1): 20–36.
- Perrin, F. 2022. "On the Origins of the Demographic Transition: Rethinking the European Marriage Pattern." *Cliometrica* 16:431–75.
- Piketty, T. 2011. "On the Long-Run Evolution of Inheritance: France 1820–2050." *The Quarterly Journal of Economics* 126 (3): 1071–131.
- Poumarède, J. 1972. *Les successions dans le sud-ouest de la France au Moyen Âge*. Paris: Presses universitaires de France.
- Renard, C. 1997. "Enquête Louis Henry. Bibliographie de l'enquête." INED documents de travail 61.
- Rosental, P.-A. 1991. "Pratiques successorales et fécondité. L'effet du Code civil." *Économie & prévision* 4–5 (100–1): 231–8.
- Rossi, P., and M. Godart. 2022. "The Old-Age Security Motive for Fertility: Evidence from the Extension of Social Pensions in Namibia." *American Economic Journal: Economic Policy* 14 (4): 1–33.
- Roy, S. 2015. "Empowering Women? Inheritance Rights, Female Education and Dowry Payments in India." *Journal of Development Economics* 114:233–51.
- Sagnac, P. 1903. "La division du sol pendant la Révolution et ses conséquences." *Revue d'histoire moderne et contemporaine* 5 (7): 457–70.
- Sah, R. K. 1991. "The Effects of Child Mortality Changes on Fertility Choice and Parental Welfare." *Journal of Political Economy* 99 (3): 582–606.
- Séguy, I. 2001. *La population de la France de 1670 à 1829. L'enquête Louis Henry et ses données*. Paris: INED.
- Shaffer, J. W. 1982. *Family and Farm. Agrarian Change and Household Organization in the Loire Valley 1500–1900*. Edited by L. A. Loubere. New York: State University of New York Press.
- Spolaore, E., and R. Wacziarg. 2022. "Fertility and Modernity." *Economic Journal* 132 (642): 796–833.
- Tudor, C. 2021. "Whose Modernity? Revolution and the Rights of Woman." Ph.D. diss., Princeton University.
- van de Walle, E. 1986. "La fécondité française au XIX^e siècle." *Communications* 44:35–45.
- Vogl, T. S. 2016. "Differential Fertility, Human Capital, and Development." *The Review of Economic Studies* 83 (1): 365–401.
- Voltaire. 1829. In *Oeuvres de Voltaire; 26–30, 32. Dictionnaire philosophique. T. 32*. Paris: Werdet et Lequien fils.
- Weir, D. R. 1984. "Fertility Transition in Rural France, 1740–1829." *The Journal of Economic History* 44 (2): 612–4.

- Yver, J. 1966. *Égalité entre héritiers et exclusion des enfants dotés. Essai de géographie coutumière*. Paris: Sirey.
- Zink, A. 1993. *L'héritier de la maison. Géographie coutumière du Sud-Ouest de la France sous l'Ancien Régime*. Paris: EHESS.



The atlas of local jurisdictions of Ancien Régime France

Victor Gay ^{a,*}, Paula E. Gobbi ^b, Marc Goñi ^c

^a Department of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole, 1 Esplanade de l'Université, 310080 Toulouse, France

^b Université Libre de Bruxelles (ECARES) and CEPR, Belgium

^c University of Bergen and CEPR, Norway

ARTICLE INFO

Article history:

Received 11 October 2023

Received in revised form

6 March 2024

Accepted 21 March 2024

Keywords:

Administrative boundary

Historical geographic information system

Jurisdiction

Bailliage

Institution

Ancien Régime

France

ABSTRACT

This article describes the construction and content of an atlas of local jurisdictions of Ancien Régime France: *bailliages*. *Bailliages* were at the center of the Ancien Régime's jurisdictional apparatus: they administered the ordinary royal justice, delineated the area of influence of heterogeneous customary laws, and served as electoral constituencies for the Estates General of 1614 and 1789. Yet, their territorial extent was relatively unknown to the royal authority, leading early scholars to assert the impossibility of mapping the geography of *bailliages*. Based on Armand Brette's *Atlas des bailliages et juridictions assimilées* published in 1904, we develop a historical geographic information system containing shapefiles and associated data files of *bailliage* courts at the time of the convocation of the Estates General of 1789. This new source has many potential applications, including mapping the different legal systems that coexisted in France, such as Roman law in *pays de droit écrit* and customary law in *pays de droit coutumier*, and studying elections to the Estates General of 1789.

© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The legal system of Ancien Régime France was inherently complex and fragmented.¹ A myriad of local justices enforced a heterogeneous set of rules that varied from parish to parish, effectively regulating social and economic relations.² At the core of this jurisdictional apparatus were *bailliages*.³ *Bailliage* courts administered the ordinary royal justice and had the authority to try civil and criminal cases of ordinary importance in first instance. They also had appellate jurisdiction over cases from inferior, subordinate courts. Beyond their legal attributions, *bailliages* were key administrative divisions: they delineated the areas of influence of

various local customary laws and were used as electoral constituencies for the Estates General of 1614 and 1789. Despite their importance, and much as other jurisdictions at the time, the territorial extent of *bailliages* was relatively unknown to the royal authority. In fact, the complexity of their limits led early scholars to assert the impossibility of mapping the geography of Ancien Régime *bailliage* jurisdictions.

This article describes the construction and content of a historical geographic information system (GIS) of *bailliages* in 1789: the *Atlas of Local Jurisdictions of Ancien Régime France*. This shapefile and associated data files are based on Armand Brette's *Atlas des bailliages et juridictions assimilées*.⁴ This source overcomes uncertainty over the territorial extent of *bailliages* by taking advantage of a unique opportunity to gain insight into their territorial organization: the convocation of the Estates General of 1789, which used

* Corresponding author. Department of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole, 1 Esplanade de l'Université, 310080 Toulouse, France.

E-mail address: victor.gay@tse-fr.eu (V. Gay).

¹ Jean-Pierre Royer, *Histoire de la justice en France* (Paris: Presses universitaires de France, 1995); Didier Catarina, *Atlas des justices et seigneuries de Languedoc 1667–1789* (Toulouse: Fédération Historique de Midi-Pyrénées, 2014); Béatrice Fourniel, 'La carte judiciaire d'Ancien Régime: un enchevêtrement de ressorts', in *Atlas historique: Auvergne, Bourbonnais, Velay*, ed. by Stéphane Gomis (Clermont Ferrand: Université Clermont Auvergne, 2017).

² *Les justices de village. Administration et justice locales de la fin du Moyen Âge à la Révolution*, ed. by François Brizay, Antoine Follain, and Véronique Sarrazin (Rennes: Presses universitaires de Rennes, 2003).

³ Jean-Michel Gorry 'Les bailliages royaux', in *Atlas archéologique de Touraine*, ed. by Elisabeth Zadora-Rio (Tours: FERACF, 2014).

⁴ Armand Brette, *Atlas des bailliages ou juridictions assimilées ayant formé unité électorale en 1789 dressé d'après les actes de la convocation conservés aux Archives nationales* (Paris: Imprimerie Nationale, 1904). A geographic information system (GIS) is a type of database that links each element of data to a spatial location. GIS data are typically stored in shapefile format, which is a geospatial vector data format that consists of a set of files that collectively store both geometric and attribute data for geographic features. Attribute data are stored in attribute tables in dbf format. See Ian N. Gregory and Paul S. Ell, *Historical GIS: Technologies, Methodologies and Scholarship* (Cambridge: Cambridge University Press, 2007).

these jurisdictions as electoral constituencies. In particular, the text of the convocation as well as the minutes of bailliage assemblies in the preparation of these Estates precisely document the territorial distribution of parishes among bailliaiges.⁵

Since the publication of Brette's atlas, there have been several attempts to make its content accessible to researchers using modern mapping standards. In their *Atlas de la Révolution française*, for instance, Marie-Vic Ozouf-Marignier, Daniel Nordman, and Alexandra Laclau published a simplified version of Brette's map by representing the boundaries of bailliaiges – without additional information, such as bailliaiges' names or types – on a paper map at a scale an order of magnitude below the precision of Brette's original maps.⁶ Later, several regional historical atlases published digital (that is, as PDF or JPEG) versions of bailliage maps using more precise scales but focusing on limited parts of the territory, such as Languedoc or Touraine.⁷ In contrast, our atlas covers the entire territory and goes beyond these paper or digital maps by disseminating Brette's atlas as a historical GIS along with a set of attribute tables containing comprehensive information on the toponymy and characteristics of each bailliage. The data we propose can therefore be used directly by researchers and is compatible with other research programs that use localized historical information on Ancien Régime France.⁸

Creating a historical GIS of bailliaiges based on Brette's atlas poses several methodological challenges. Despite its widespread use in historical geography, Brette's atlas is not without limitations. Most importantly, these maps do not represent the exact territorial limits of bailliaiges. Rather, they indicate whether a particular parish or town belonged to a particular bailliage. As a result, the most common method used to construct geographic information systems based on historical maps – the vectorization of boundaries – would produce substantial inaccuracies, especially in regions where jurisdictions are highly fragmented. We propose a novel methodology that overcomes these limitations by combining GIS techniques with the use of historical toponymies. We also design our atlas in both polygon and point form to mitigate the methodological issue of a polygon representation of intertwined jurisdictions. This method can be emulated in many cases when creating historical GISs based on historical maps.

Our work contributes to the historical geography of administrative jurisdictions through GIS methods, which have gained a central place in historical geography scholarship along with the

spatial turn in the humanities.⁹ Indeed, historical GISs have undergone significant developments in the last two decades for France (*TRF-GIS*, 1870–1940), Germany (*HGIS Germany*, 1820–1914), Great Britain (*Great Britain Historical GIS*, 1801–2001), the Netherlands (*NLGis*, 1812–1997), and the United States (*NHGIS*, 1790–2022).¹⁰ The use of GIS in humanities research has enabled historians to radically change the way they think about space.¹¹ In France, early scholarship in the late nineteenth and early twentieth centuries focused on spatial considerations related to the creation of départements in the context of the centennial of the Revolution.¹² This field then experienced renewed interest in the late twentieth century with the prospect of the bicentennial of the Revolution.¹³ Nevertheless, historical GISs have focused almost exclusively on nineteenth- and twentieth-century administrative constituencies, limiting the ability of historians and social scientists to analyze anterior historical spatial processes – some exceptions for pre-industrial periods include historical GISs for Ireland (*The Down Survey of Ireland*, 1656–8), the Netherlands (*Historical Atlas of the Low Countries*, 1350–1800), Portugal (*Atlas Cartografia Histórica*,

⁹ Gregory and Ell, *Historical GIS: The Spatial Humanities: GIS and the Future of Humanities Scholarship*, ed. by David Bodenhamer, John Corrigan, and Trevor Harris (Bloomington: Indiana University Press, 2010); *Toward Spatial Humanities: Historical GIS and Spatial History*, ed. by Ian N. Gregory and Alistair Geddes (Bloomington: Indiana University Press, 2014).

¹⁰ For France, see Jean-Luc Pinol, 'Les systèmes d'information géographique et la pratique de l'histoire', *Revue d'histoire moderne & contemporaine* 58 (2011) 111–26; Victor Gay, 'Mapping the Third Republic. A Geographic Information System of France (1870–1940)', *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 54 (2021) 189–207. For Germany, see Andreas Kunz, *Das Historische Informationssystem HGIS Germany und seine Module* (Mainz: HGIS-Germany, 2008). For Great Britain, see Humphrey Southall, 'Rebuilding the Great Britain Historical GIS, Part 1: Building an Indefinitely Scalable Statistical Database', *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 44 (2011) 149–59; Humphrey Southall, 'Rebuilding the Great Britain historical GIS, part 2: A Geo-Spatial Ontology of Administrative Units', *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 45 (2012) 119–34; Humphrey Southall, 'Rebuilding the Great Britain Historical GIS, Part 3: Integrating Qualitative Content for a Sense of Place', *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 47 (2014) 31–44. For the Netherlands, see Vyacheslav Tykhonov, Richard Zijdemann, and Jerry de Vries, *NLGIS-2: Website of Netherlands Geographic Information System* (Amsterdam: International Institute for Social History, 2014). For the United States, see Catherine A. Fitch and Steven Ruggles, 'Building the National Historical Geographic Information System', *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 36 (2003), 41–51; Steven Manson, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles, *IPUMS National Historical Geographic Information System: Version 18.0* (Minneapolis: IPUMS, 2023).

¹¹ Luís Espinha da Silveira, 'Geographic Information Systems and Historical Research: An Appraisal', *International Journal of Humanities and Arts Computing* 8 (2014) 28–45; Gregory and Ell, *Historical GIS*, pp. 1–19; Patricia Murrieta-Flores and Bruno Martins, 'The Geospatial Humanities: Past, Present and Future', *International Journal of Geographical Information Science* 33 (2019) 2424–9.

¹² See a review of this early scholarship in Nicolas Verdier and Marie-Vic Ozouf-Marignier, 'Circonscriptions et réseaux de voies: un angle mort de la géographie historique', *Études rurales* 188 (2011) 114–42.

¹³ *Paroisses et communes de France*, *Dictionnaire d'histoire administrative et démographique*, ed. by Jean-Pierre Bardet and Jacques Dupâquier (Paris: Éditions du CNRS, 1974–91); Guy Arbellot, Jean-Pierre Goubert, Jacques Mallet, and Yvette Palazot, *Carte des généralités, subdélégations et élections en France à la veille de la Révolution de 1789* (Paris: Éditions du CNRS, 1986); Marie-Vic Ozouf-Marignier, *La formation des départements. La représentation du territoire français à la fin du XVIII^e siècle* (Paris: Éditions de l'EHESS, 1989); *Atlas de la Révolution française*, ed. by Serge Bonin and Claude Langlois (Paris: Éditions de l'EHESS, 1987–2000); Anne Zink, *Pays ou circonscription. Les collectivités territoriales de la France du Sud-Ouest sous l'Ancien Régime* (Paris: Publications de la Sorbonne, 2000); Claude Motte, Isabelle Séguy, and Christian Thérèse, *Communes d'hier, communes d'aujourd'hui. Les communes de la France métropolitaine, 1801–2001. Dictionnaire d'histoire administrative* (Paris: INED, 2003); *Des paroisses de Touraine aux communes d'Indre-et-Loire. La formation des territoires*, ed. by Elisabeth Zadora-Rio (Tours: FÉRAFC, 2008); *Le canton. Un territoire du Quotidien?*, ed. by Yann Lagadec, Jean Le Bihan, and Jean-François Tanguy (Rennes: Presses universitaires de Rennes, 2009).

⁵ Parishes were the most granular level of administration in Ancien Régime France – there were nearly forty-four thousand parishes in 1789. They encompassed three notions that did not overlap everywhere: the ecclesiastical parish, the fiscal parish, and the inhabitants community. The notion of fiscal parish was central during the transition to *municipalités* in 1789–93 and later to communes. See Jean-Michel Gorry, 'Des communes de la Révolution aux communes actuelles', in *Des paroisses de Touraine aux communes d'Indre-et-Loire: la formation des territoires*, ed. by Elisabeth Zadora-Rio (Tours: FÉRAFC, 2008), pp. 121–83; Jean-Michel Gorry 'Des communes de 1790 aux communes actuelles', in *Atlas archéologique de Touraine*, ed. by Elisabeth Zadora-Rio (Tours: FÉRAFC, 2014).

⁶ Marie-Vic Ozouf-Marignier, Daniel Nordman, and Alexandra Laclau, 'Les bailliaiges électoraux en 1789', in *Atlas de la Révolution française. Le territoire* (2). *Les limites administratives*, ed. by Daniel Nordman and Marie-Vic Ozouf-Marignier (Paris: Éditions de l'EHESS, 1989) pp. 21–9.

⁷ Elie Pélaquier, *Atlas historique de la province de Languedoc* (Montpellier: C.R.I.S.E.S., 2009); Gorry 'Les bailliaiges royaux'.

⁸ As noted by Ian N. Gregory and Richard G. Healy in 'Historical GIS: Structuring, Mapping and Analysing Geographies of the Past', *Progress in Human Geography* 31 (2007) 638–53 (p. 639–40): 'Electronic databases are far more useful and useable than their paper counterparts, but ironically it is harder for them to gain academic standing and for their creators to receive due recognition'.

1758–2001), and colonial Spanish America (*HGIS de las Indias*, 1701–1808).¹⁴ We contribute to filling this gap by providing one of the first curated historical GIS of administrative boundaries in Ancien Régime France.¹⁵ It will further enable historical geographers to contextualize and explore the spatial dynamics of historical processes related to the complex jurisdictional landscape of Ancien Régime France.

The atlas we propose has many potential applications. First, it will empower historical geography and empirical research in the context of Ancien Régime France by providing an underlying GIS of reference for bailliages – a central jurisdiction in the regulation of Ancien Régime society. Second, because legal rules differed across bailliage jurisdictions, our historical GIS can be used to map the different legal systems that coexisted in France, such as Roman law in *pays de droit écrit* and customary law in *pays de droit coutumier*.¹⁶ More precisely, each bailliage can be associated with a set of customary laws, so the map of bailliages is a necessary preliminary step to the construction of a customary map.¹⁷ Customary boundaries then make it possible to study how different legal institutions emerge or affect societies. For instance, it has recently been used to map the spatial distribution of pre-Revolutionary inheritance rules and study how the unification of these rules in 1793 contributed to France's early fertility decline.¹⁸ Third, because bailliages were used as electoral constituencies for the Estates General of 1789, our resource will be valuable for studying the social, political, and economic conditions leading up to the Revolution.¹⁹ In particular, bailliages can be matched with the individuals elected to the Estates General – all listed in Brette (1894–1915) – and with the locations where the *cahiers de doléance* were produced, a source

commonly used to study the state of the country on the eve of the Revolution.²⁰ Fourth, because royal agents headed each bailliage, our data can be matched with individual-level information on royal offices, including the venality of their associated position from, for instance, the *Enquête Colbert* of 1665.²¹ Finally, this resource will help historical geographers, as well as political and social scientists, shed new light on the evolution of jurisdictions and the extent to which post-revolutionary administrative divisions were rooted in, or deviated from, the jurisdictions that prevailed during the Ancien Régime.²²

The shapefile and associated datafiles of our atlas of local jurisdictions of Ancien Régime France are disseminated in the open-access repository 'Bailliages in 1789 France' hosted on the Harvard Dataverse under the CC-BY 4.0 license.²³

The remainder of this article is organized as follows: we first discuss Armand Brette's *Atlas des bailliages* – including an overview of the legal system of Ancien Régime France, the context of the convocation of the Estates General of 1789, as well as Armand Brette's cartography – then provide details on our mapping methodology, and finally describe the content of the data of our historical GIS. We conclude the article with two concrete examples of how our atlas of jurisdictions can be used.

Armand Brette's *Atlas des bailliages*

The legal system of Ancien Régime France

The legal system of Ancien Régime France was organized in several layers (Appendix Fig. A1). At the top of the system were the parliaments, which functioned as sovereign courts with appellate jurisdiction over civil and criminal cases from inferior courts. From the mid-thirteenth century, the Parliament of Paris was the only parliament and had jurisdiction over the entire royal domain. The number of parliaments then gradually increased to thirteen by the time of the Revolution, covering the entire territory.

Below parliaments, the local jurisdictions of bailliages administered the ordinary royal justice – in the southern part of the realm, these courts were also known as *sénéchaussées*.²⁴ Bailliages had the authority to try civil and criminal cases in first instance for royal cases of ordinary importance.²⁵ They also had appellate jurisdiction over ordinary cases from inferior, subordinate courts – *prévôtés*, *châtellenies*, and other manorial courts that existed throughout the territory, as well as regional courts such as *vigueries* in Languedoc

¹⁴ For Ireland, see Mícheál Ó Siochrú, *The Down Survey of Ireland Project* (Dublin: Trinity College Dublin, 2013). For the Netherlands, see Rombert Stapel, 'Historical Atlas of the Low Countries: A GIS Dataset of Locality-Level Boundaries (1350–1800)', *Research Data Journal for the Humanities and Social Sciences* 8 (2023) 1–33. For Portugal, see Luís Nuno Espinha da Silveira, Daniel Ribeiro Alves, Nuno Lima, Ana Alcântara, *Atlas Cartografia Histórica* (Lisbon: Faculdade de Ciências Sociais e Humanas, 2001). For colonial Spanish America, see Werner Stangl, 'Digital Resources: HGIS de las Indias', *Oxford Research Encyclopedia of Latin American History* (Oxford: Oxford University Press, 2020).

¹⁵ Some notable exceptions are Arbellot, Goubert, Mallet, and Palazot, *Carte des généralités, subdélégations et élections*; *Atlas de la Révolution française*, ed. by Bonin and Langlois; Pélaquier, *Atlas historique de la province de Languedoc*; *Atlas archéologique de Touraine*, ed. by Elisabeth Zadora-Rio (Tours: FERACF, 2014); Catarina, *Atlas des justices et seigneuries de Languedoc*; *Atlas historique: Auvergne, Bourbonnais, Velay*, ed. by Stéphane Gomis (Clermont-Ferrand: Université Clermont Auvergne, 2017).

¹⁶ Béatrice Fourniel and Jacqueline Vendrand-Voyer, 'Une bigarrure de loix dans une même province', in *Atlas historique: Auvergne, Bourbonnais, Velay*, ed. by Stéphane Gomis (Clermont Ferrand: Université Clermont Auvergne, 2017); Henri Klimrath, *Études sur les coutumes* (Paris: Levraud, 1837); Jean Hilaire and François Terré, 'Un cliché énigmatique. Une interprétation de la carte de la France coutumière', in *La Vie du Droit* (Paris: Presses universitaires de France, 1994).

¹⁷ See Victor Gay, Paula E. Gobbi, and Marc Goñi, 'The Customary Atlas of Ancien Régime France', *Explorations in Economic History* 93 (2024), 101588, and the associated open-source shapefile in Victor Gay, Paula E. Gobbi, and Marc Goñi, *Customs in Ancien Régime France* (Harvard dataverse, 2024), available at <https://doi.org/10.7910/DVN/HYE209>. We develop this example in the last section of this article.

¹⁸ Victor Gay, Paula E. Gobbi, and Marc Goñi, 'Revolutionary Transition: Inheritance Change and Fertility Decline', CEPR Discussion Paper 18607 (available at <https://hal.science/hal-04285818>).

¹⁹ François Furet, 'The Monarchy and the Procedures for the Elections of 1789', *The Journal of Modern History* 60 (1988), 558–74; Malcolm Crook, 'The Persistence of the Ancien Régime in France: The Estates General of 1789 and the Origins of the Revolutionary Electoral System', *Parliaments, Estates and Representation* 13 (1993) 29–40; Malcolm Crook, *Elections in the French Revolution: An Apprenticeship in Democracy, 1789–1799* (Cambridge: Cambridge University Press, 1996).

²⁰ Beatrice F. Hyslop, *French Nationalism in 1789 according to the General Cahiers* (New York: Columbia University Press, 1934); Gilbert Shapiro and John Markoff, *Revolutionary Demands: A Content Analysis of the Cahiers de Doléances of 1789* (Stanford: Stanford University Press, 1988); Joseph Goy, 'Transmission successorale et paysannerie pendant la Révolution française: un grand malentendu', *Études rurales* 110–2 (1988) 45–56; Mara P. Squicciarini and Nico Voigtlander, 'Knowledge Elites and Modernization: Evidence from Revolutionary France', NBER Working Paper 22779 (2016).

²¹ See, for instance, Christophe Blanquie, *Une enquête de Colbert en 1665. La généralité de Bordeaux dans l'enquête sur les offices* (Paris: L'Harmattan, 2012).

²² Ozouf-Marignier, *La formation des départements: Des Paroisses de Touraine aux Communes d'Indre-et-Loire*, ed. by Elisabeth Zadora-Rio.

²³ It is available at the following address: <https://doi.org/10.7910/DVN/T8UXHK>. See Victor Gay, Paula E. Gobbi, and Marc Goñi, *Bailliages in 1789 France* (Harvard Dataverse, 2024).

²⁴ Between bailliages and parliaments was another legal institution: presidial courts (*présidiaux*). Created in 1551, these courts had appellate jurisdiction over cases from bailliage courts. There were initially about sixty presidial courts in the mid-sixteenth century, and as many as 101 by the mid-eighteenth century.

²⁵ Cases of "ordinary" importance were those involving property damages of up to 250 livres tournois from 1551, then up to 2000 from 1774. See Marcel Marion, 'À propos de la géographie judiciaire de la France sous l'Ancien Régime. La question du ressort des présidiaux', *Revue historique* 89 (1905) 80–8.

and *vicomtés* in Normandy.²⁶ Bailliages were created in the twelfth century on the basis of pre-existing feudal legal structures. They expanded significantly over time with the integration of new provinces into the kingdom: from forty-two in the fifteenth century, bailliages grew to over four hundred on the eve of the Revolution.²⁷

Although bailliages were an emanation of the royal authority, their territorial jurisdiction was not centrally planned. Instead, bailliages developed haphazardly as the monarchy consolidated and feudal strongholds were unified, broken up, and suppressed. It was only with the process of codification and reformation of customs between the mid-fourteenth and late fifteenth centuries that bailliages acquired a clear status of legal territorial division.²⁸

Bailliage agents benefited from the contemporary jurisdictional confusion. Because they had a financial stake in the territorial reach of their court, they were in a constant state of jurisdictional conflict with neighboring bailliages in a never-ending quest for expansion.²⁹ Moreover, because bailliages were subject to a heterogeneous set of customary laws, litigants often appealed to neighboring bailliage courts, playing them off against each other in the hope of obtaining a more favorable resolution of their case.³⁰ As a result, the territorial geography of bailliage jurisdictions was inherently fragmented. It exemplified the entanglement of jurisdictions and rights that characterized Ancien Régime institutions.³¹

The intricate nature of Ancien Régime bailliage jurisdictions led early scholars to claim that it was impossible to map their spatial distribution.³² In fact, the royal authority – lacking precise knowledge of the territorial extent of bailliages – never produced such a map. In contrast, local administrators had a fair understanding of their spatial configuration, although they lacked the cartographic tools to effectively communicate this information up the jurisdictional hierarchy. Instead, they relied on parish lists – which territorial limits were well known to local communities – to manage

their jurisdiction.³³ The relative ignorance of central authorities may have been purposeful on the part of local bailliage agents as they had incentives to conceal information about the actual territorial extent of their jurisdiction in order to expand it. Thus, the uncertainty surrounding the territorial reach of bailliages may have been the result of a lack of appropriate cartographic tools combined with strategic motives rather than a reflection of sheer ignorance.³⁴

Amid such complexity, the archival imprint left by the convocation of the Estates General of 1789 represents a unique opportunity to gain insight into the jurisdictional landscape of Ancien Régime France throughout its territory.

The convocation of the Estates General of 1789

The Estates General of May–June 1789 were convened by King Louis XVI in August 1788 to address France's ongoing financial and economic crisis. They consisted of a consultative assembly composed of nearly twelve hundred elected representatives of the clergy, the nobility, and the Third Estate.³⁵ In order to organize the electoral process, the royal authority had to define the constituencies within which these representatives would be elected. Since the convocation was a judicial act – the modalities of which were codified by the royal decree of January 24, 1789 – the Royal Council decided that these elections would take place within the constituencies of the royal justice: bailliages.³⁶

Upon the convocation of the Estates General, bailliage officers convened representatives of parishes in their jurisdiction to bailliage assemblies to form an electoral constituency and proceed with the elections. When they did not know the exact territory of their own bailliage, officers resorted to information available in secondary sources such as Expilly's or Bruzen de La Martinière's geographic dictionaries, or the text of their bailliage's customs.³⁷ Uncertainties about the territorial limits of bailliages were thus resolved upon the reunion of parishes during bailliage electoral assemblies.

While bailliage courts covered most of the territory, the royal justice did not operate everywhere. Several territories on the periphery of the kingdom had retained the power of justice within their realm, so that bailliage courts did not exist there. These territories thus had to rely on alternative jurisdictions of similar size to form electoral constituencies for the Estates General. For instance, the *généralité* of Strasbourg in Alsace formed three electoral constituencies through the reunion of its six *districts*, which were

²⁶ The jurisdiction of bailliage courts was codified by the edict of Crémieu on June 19, 1536. Other jurisdictions concerned specific domains: for instance, financial matters were administered by the *Cours des aides*, the *Chambre des comptes*, and the *Cours des monnaies*.

²⁷ Ozouf-Marignier, Nordman, and Laclau, 'Les Bailliages Électoraux en 1789'.

²⁸ Martine Grinberg, Simone Geoffroy-Poisson, and Alexandra Laclau, 'Rédaction des coutumes et territoires au XVI^e Siècle: Paris et Montfort-L'Amaury', *Revue d'histoire moderne & contemporaine* 59 (2012) 7–55; Marie Seong-Hak Kim, *Custom, Law, and Monarchy: A Legal History of Early Modern France* (Oxford: Oxford University Press, 2021).

²⁹ Bernard Guenée, 'La géographie administrative de la France à la fin du Moyen Âge: élections et bailliages', *Le Moyen Âge* 66 (1961) 293–323.

³⁰ Fourniel, 'La carte judiciaire d'Ancien Régime'; Kim, *Custom, Law, and Monarchy*, p. 62.

³¹ Although not directly related to the legal system, it is worth mentioning that ecclesiastical divisions (*diocèses*, *archidiaconés*, *archiprêtres*, *doyennés*) represented another layer of territorial complexity – although diocèses had been relatively stable since the fifth century. On the cartography of ecclesiastical divisions before the Revolution, see François de Dainville, *Cartes anciennes de l'Église de France. Historique, répertoire, guide d'usage* (Paris: J. Vrin, 1957) and Jacques Dubois, 'La carte des diocèses de France avant la Révolution', *Annales*, 20 (1965), 680–91. For further details on institutions of Ancien Régime France, see Marcel Marion, *Dictionnaire des institutions de la France aux XVII^e et XVIII^e siècles* (Paris: Picard, 1923); Denis Richet, *La France moderne: l'esprit des institutions* (Paris: Flammarion, 2019 [1973]); Roland Mousnier, *Les institutions de la France sous la monarchie absolue* (Paris: Presses universitaires de France, 1974); Bernard Barbiche, *Les institutions de la monarchie française à l'époque moderne, 16^e–18^e siècles* (Paris: Presses universitaires de France, 1999); Christophe Blanquie, *Les institutions de la France des Bourbons (1589–1789)* (Paris: Belin, 2003); Albert Rigaudière, *Histoire du droit et des institutions dans la France médiévale et moderne* (Paris: Economica, 2010); and Laurent Avezou, *Les institutions de la France Moderne, XVI^e–XVIII^e siècle* (Paris: Armand Colin, 2014).

³² Gustave Dupont-Ferrier, *Les officiers royaux des bailliages et sénéchaussées et les institutions municipales locales en France à la fin du moyen âge* (Paris: Emile Bouillon, 1902).

³³ Guenée, 'La géographie administrative de la France à la fin du Moyen Âge'; Léonard Dauphant, *Géographies. Ce qu'ils savaient de la France (1100–1600)* (Ceyzérieu: Champ Vallon, 2018).

³⁴ Nelly Girard d'Albissin, 'Propos sur la frontière', *Revue historique de droit français et étranger* 47 (1969) 390–407.

³⁵ For more details on the organization of these elections, which occurred from January to May 1789, see Armand Brette, *Recueil de documents relatifs à la convocation des États généraux de 1789* (Paris: Imprimerie Nationale, 1894–1915); Furet, 'The Monarchy and the Procedures for the Elections of 1789'; Crook, 'The Persistence of the Ancien Régime in France'; Crook, *Elections in the French Revolution*.

³⁶ The decree of January 24, 1789 stated that "the terms bailliage or sénéchaussée shall be understood to mean all the jurisdictions to which the knowledge of royal cases is attributed." A similar procedure had been used in the previous Estates General of 1614. Bailliages that had been electoral constituencies during those Estates General were qualified as "principal" in 1789, while those that had not been convened during those Estates were qualified as "secondary." This did not presume of any difference between these bailliages other than a formal rank.

³⁷ Jean-Joseph Expilly, *Dictionnaire géographique, historique et politique des Gaules et de la France* (Paris: Desaint et Saillant, 1762–1770); Antoine-Augustin Bruzen de La Martinière, *Le grand dictionnaire géographique et critique* (La Haye: P. Gosse, R. C. Alberts, P. de Hondt, 1726–1739); Charles-Antoine Bourdot de Richebourg, *Nouveau coutumier général, ou Corps des coutumes générales et particulières de France et des provinces connues sous le nom de Gaules* (Paris: Michel Brunet, 1724); Martine Grinberg, *Écrire les coutumes. Les droits seigneuriaux en France (XVI^e–XVIII^e siècle)* (Paris: Presses universitaires de France, 2006).

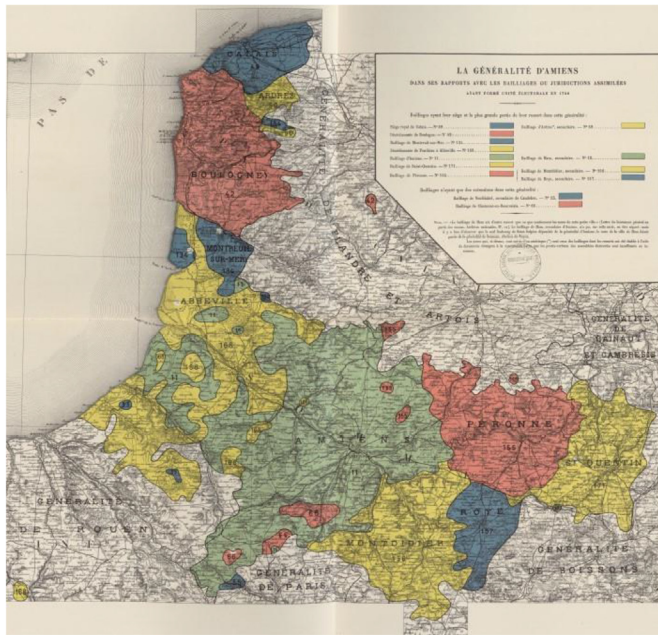


Fig. 1. Bailliages of the Généralité of Amiens. Source: Brette, *Atlas des bailliages ou juridictions assimilées*.

jurisdictions previously created by Alsace's provincial assemblies.³⁸ Furthermore, some territories in the généralité of Grenoble – though equipped with bailliages – chose alternative jurisdictions to form electoral constituencies. In particular, the Dauphiné Estates elected their representatives in January 1789 using special tribunals, called *élections*, as electoral constituencies.³⁹

A cartography of bailliages in 1789

Based on the text of the convocation and the minutes (*procès verbaux*) of bailliage electoral assemblies, Armand Brette designed an atlas of bailliage jurisdictions – the *Atlas des bailliages et juridictions assimilées ayant formé unité électorale en 1789*, published in 1904. Armand Brette produced this atlas as an appendix to his larger collection of documents about to the convocation of the Estates General of 1789, believing that maps would be more practical than lengthy tables.⁴⁰ This scholarly work was carried

out as he was part of the *Commission de la recherche et de publication des documents relatives à la vie économique de la Révolution*, a project promoted by Jean Jaurès and initiated in 1903 to anchor national unity and consolidate the Republic in the context of the centenary of the French Revolution and the ideological struggle for the separation of the Churches and the State, achieved in 1905.⁴¹

To construct this atlas, Armand Brette reported the forty thousand locations listed in the convocation and the minutes on a map of France, the *Carte du Dépôt de la guerre* at a scale of 1:320,000, which displays the names of all of France's *communes* circa 1852.⁴² The resulting atlas is composed of thirty-two large-scale maps, one for each généralité – we display an excerpt below (Fig. 1).⁴³ These maps show the territorial extent of each bailliage along with a nomenclature developed by Armand Brette and other information that we describe below.

In his introduction to the *Atlas des bailliages*, Armand Brette emphasized that his maps were not intended to represent the exact territorial limits of bailliages, but rather to indicate whether a particular parish or town belonged to a particular jurisdiction. In fact, Brette's atlas does not use cadastral information, so the colored shapes on the maps should only be interpreted insofar as they include a particular commune seat, indicated by a circle on the underlying *Carte du Dépôt de la guerre*. We explain below how our GIS deals with this limitation.

As the first attempt at a country-wide cartography of Ancien Régime bailliage jurisdictions, Brette's atlas was praised by contemporary scholars upon its publication.⁴⁴ Since then, it has been a seminal reference for historical geography research, whether to locate a specific bailliage or the bailliage to which some towns belonged, to confine the analysis to a specific, self-contained jurisdictional space, or to study the historical

³⁸ *Généralités* were the largest administrative divisions in Ancien Régime France and the basis for the administration of the territory under the authority of intendants. There were thirty-five généralités on the eve of the Revolution. See Jean-Michel Gorry 'L'administration générale et fiscale d'Ancien Régime: la généralité et les élections, in *Atlas archéologique de Touraine*, ed. by Elisabeth Zadora-Rio (Tours: FERACF, 2014). The three electoral constituencies of the généralité of Strasbourg were the result of the union of the districts of Colmar and Sélestat, Belfort and Huningue, and Haguenau and Wissembourg. In addition, the city of Strasbourg obtained two representatives for the Third Estate. Still in Alsace, the ten imperial cities of Haguenau, Colmar, Sélestat, Wissembourg, Landau, Obernai, Rosheim, Munster, Kaysersberg, and Turckheim obtained the right to form a single electoral constituency in Sélestat by the decree of February 7, 1789.

³⁹ In general, all three estates formed the same electoral constituency. The main exception to this rule was in Brittany, where the clergy used dioceses rather than bailliages as the basis for its electoral constituencies.

⁴⁰ Brette, *Recueil de documents relatifs à la convocation des États généraux de 1789*. A reduced version of this atlas was also published in a later opusculé entitled *Les limites et divisions territoriales de la France en 1789* (Paris: Cornély, 1907) in the form of a one-page map with a simplified geometry. This map is available in Gallica at the following address: <https://gallica.bnf.fr/ark:/12148/bpt6k81492f/f139>. For more details on Armand Brette's original scholarship in historical geography, see Alphonse Aulard, 'Armand Brette', *La Révolution française: revue d'histoire moderne et contemporaine* 62 (1912) 385–95.

⁴¹ Verdier and Ozouf-Marignier, 'Circonscriptions et réseaux de voies', pp. 118–9; Christine Peyrard, 'La création de la Commission ou l'œuvre de Jean Jaurès', in *Héritages de la Révolution française à la lumière de Jaurès*, ed. by Christine Peyrard and Michel Vovelle (Aix-en-Provence: Presses universitaires de Provence, 2002), pp. 19–42.

⁴² See Jean-Luc Arnaud, *La carte de France, histoire et techniques* (Marseille: Parenthèses, 2022). When the text of the convocation was insufficient and the minutes from the bailliage electoral assemblies were missing in the National Archives, Armand Brette used alternative sources, among which the minutes of the writing of customs in Bourdot de Richebourg's *Nouveau coutumier général*, various population counts such as Saugrain's *Dénombrement du royaume*, or Saugrain's, Expilly's, Simon's, or Boutiot's parish dictionaries. See Claude-Marin Saugrain, *Dénombrement du royaume par généralités, élections, paroisses et feux* (Paris: Saugrain, 1709); Claude-Marin Saugrain, *Dictionnaire universel de la France ancienne et moderne, et de la Nouvelle France* (Paris: Saugrain, 1726); Pierre-Guillaume Simon, *Dictionnaire des paroisses du ressort du Parlement de Paris* (Paris: Imprimerie du Parlement, 1776), and Théophile Boutiot and Émile Socard, *Dictionnaire topographique du département de l'Aube* (Paris: Imprimerie Nationale, 1874). The list of sources used by Brette are detailed for each bailliage in his atlas and available in the online repository.

⁴³ Two exceptions are the généralités of Toulouse and Perpignan, and those of Nancy and Metz, which are displayed together on the same maps in Brette's atlas. The atlas also includes a map of French colonies (Maurice, Saint-Domingue, Martinique, Guadeloupe, Pondichéry), which we do not include in our GIS.

⁴⁴ Philippe Sagnac, 'La France en 1789 et les États généraux. D'après les travaux de M. Armand Brette', *Revue d'histoire moderne & contemporaine* 6 (1904) 165–81; Marcel Marion, 'À Propos de la Géographie Judiciaire de la France sous l'Ancien Régime', Camille Bloch, 'Review of *Les limites et les divisions territoriales de la France en 1789*, by Armand Brette', *Bibliothèque de l'École des chartes* 68 (1907) 350–358.

geography of jurisdictions in Ancien Régime France and the origins of French *départements*.⁴⁵

Mapping methodology

To construct a GIS of bailliage jurisdictions in 1789, we propose a novel methodology that overcomes the limitations of our source map by combining GIS techniques and the use of historical toponomies. We proceed in four steps. First, we scan each of the thirty-two généralité-level maps of Brette's atlas at high resolution (that is, in TIFF format).⁴⁶ Second, we manually georeference these scans along an RGF93 projection. Next, we do not resort to the vectorization of bailliage boundaries, a method commonly used in the construction of geographic information systems based on historical maps.⁴⁷ This method would have led to substantial inaccuracies – especially in regions with highly fragmented jurisdictions – since colored areas on original maps should be interpreted as part of a jurisdiction only insofar as they include a particular commune seat. Instead, in a third step, we manually assign each commune to a specific bailliage. To that end, we overlay IGN's shapefile of current communes onto the georeferenced maps, revealing communes' seats and names as on Appendix Fig. A2.⁴⁸ This enables us to manually assign each commune polygon to a particular bailliage based on the comparison of names and seat locations across layers. In a fourth step, we dissolve commune polygons based on their bailliage attribute (Appendix Fig. A3).

The resulting shapefile of bailliaiges may be subject to two types of (limited) inaccuracies. First, the shapefile of contemporaneous communes does not exactly correspond to the base map of the *Dépôt de la guerre* of circa 1852 used by Armand Brette, as there were about 37,600 communes then versus 34,800 in 2021.⁴⁹

Despite their stability over time, the boundaries of some contemporaneous communes are the result of mergers of smaller communes between 1852 and 2021.⁵⁰ As a result, some of these communes may be split between two bailliaiges, making the resulting shapefile potentially inaccurate at its boundaries in some cases.⁵¹ For instance, the polygon corresponding to the bailliage of Ham includes the former parishes of Saint-Sulpice and Estouilly, as these were absorbed into the commune of Ham in 1964–5, although they historically belonged to the bailliage of Saint-Quentin. Second, the base map of the *Dépôt de la guerre* used by Armand Brette does not correspond exactly to the set of parish locations listed in the minutes of bailliage electoral assemblies in 1789, as there were nearly forty-four thousand parishes in 1789.⁵² As a result, some communes that were the result of mergers of smaller parishes or communes between 1789 and 1852 may have been similarly divided between two bailliaiges.

We offer users the ability to mitigate these potential sources of inaccuracy. Specifically, we project the coordinates of parishes at the time of the Revolution onto our shapefile of bailliaiges.⁵³ This enables us to generate a layer of parish points with bailliage attributes corrected for the aforementioned inaccuracies – for instance, we reclassify the parishes of Saint-Sulpice and Estouilly as part of the bailliage of Saint-Quentin. In addition, this alternative approach allows users to project these points back onto any commune- or parish-level shapefile and aggregate the results into bailliage-level polygons best suited for their purposes. This approach also provides the basis for alternative visual vocabularies that better represent the nature of territoriality as envisioned by contemporaries, for whom the mental construction of space was closer to a collection of places and networks than an aggregation of contiguous spatial units—we return to this issue in the final section of this article.⁵⁴

Data description

Our atlas of local jurisdictions of Ancien Régime France contains three shapefiles: a shapefile of the territory of France in 1789, a shapefile of the types of jurisdictions that formed electoral constituencies for the Estates General, and the main output of the atlas, the shapefile of bailliaiges in 1789. Each shapefile is available in two forms: a polygon form and a point form, where each point represents a parish at the time of the Revolution. The content of attribute tables is identical for the polygon- and point-form shapefiles,

⁴⁵ Philippe Sagnac, 'La composition des États généraux et de l'Assemblée nationale (1789). Étude statistique et sociale', *Revue historique* 206 (1951) 8–28; Louis Trenard, 'Prétentions et doléances de l'Ardenais. La réorganisation administrative du royaume', *Annales historiques de la Révolution française* 46 (1974) 187–215; Abel Poitrineau, 'Les assemblées primaires du bailliage de Salers en 1789', *Revue d'histoire moderne & contemporaine* 25 (1978) 419–42; John Markoff, 'The Social Geography of Rural Revolt at the Beginning of the French Revolution', *American Sociological Review* 50 (1985) 761–81; John Markoff, 'Contexts and Forms of Rural Revolt: France in 1789', *Journal of Conflict Resolution* 30 (1986) 253–89; John Markoff, 'Peasant Grievances and Peasant Insurrection: France in 1789', *The Journal of Modern History* 62 (1990) 446–76; Ted W. Margadant, *Urban Rivalries in the French Revolution* (Princeton: Princeton University Press, 1992); Squicciarini and Voigtländer, 'Knowledge Elites and Modernization'; Saumitra Jha and Steven Wilkinson, 'Revolutionary Contagion', Stanford University Graduate School of Business Research Paper 4084 (2023); Pierre Chaunu and Jean-Claude Gégot, 'Étude par sondage de la criminalité dans le bailliage de Falaise (XVIIe–XVIIIe siècle)', *Annales de Normandie* 16 (1966) 103–164; John A. Dickinson, 'L'activité judiciaire d'après la procédure civile. Le bailliage de Falaise, 1668–1790', *Revue d'histoire économique et sociale* 54 (1976) 145–68; Julius R. Ruff, *Crime, Justice and Public Order in Old Regime France* (London: Routledge, 1984); Antoine Follain, 'Les juridictions subalternes, sièges et ressorts des bailliaiges et vicomtés en Normandie sous l'Ancien Régime', *Annales de Normandie* 47 (1997), 211–26; Bernard Lepetit, 'Les temps de l'aménagement territorial. La formation des départements', *Les annales de la recherche urbaine* 43 (1989) 5–14.

⁴⁶ Each high-resolution scan is available in our "Bailliaiges in 1789 France" Data-verse. The text of the *Atlas des bailliaiges* is available from the numeric library of Sciences Po under the following Archival Resource Key (ARK) identifier: 46513/sc16qxl.

⁴⁷ See Gregory and Ell, *Historical GIS*, pp. 41–62; Jean-Luc Arnaud and Romain Suarez, 'The Map, from Analogue to Digital. Modes of Processing and Elements of a Method', *Humanités numériques* 7 (2023). The method of manual vectorization of boundaries and lines is, for instance, used in Julie Perret, Marco Gribaudo, and Marc Barthelemy, 'Roads and Cities of 18th Century France', *Scientific Data* 2 (2015) 150048, for France's roads in the late eighteenth century, or in Krzysztof Ostafin, Dominik Kaim, Tomasz Siwek, and Anna Mikla, 'Historical Dataset of Administrative Units with Socio-Economic Attributes for Austrian Silesia 1837–1910', *Scientific Data* 7 (2020) 208, for the administrative organization of Austrian Silesia in the nineteenth century.

⁴⁸ IGN, *ADMIN-EXPRESS édition mars 2021 par territoire France métropolitaine* (Paris: IGN, 2021).

⁴⁹ Gay, 'Mapping the Third Republic'.

⁵⁰ On the methodological issues of using current commune boundaries as the underlying framework, see Victor Gay, 'Mapping the Third Republic', p. 192; Stapel, 'Historical Atlas of the Low Countries', pp. 8–9.

⁵¹ We are constrained to use the commune shapefile of 2021 because commune-level shapefiles for 1852 do not exist at this time – this is among the endeavors of the ongoing COMMUNE HIS-DBD project (<https://anrcommunes.hypotheses.org/>), which is scheduled to make these shapefiles available in 2025.

⁵² The number of 43,792 parishes is based on the number of parish seats on Cassini's map of France (*Carte générale de la France*), which was surveyed between 1756 and 1789. See François de Dainville, 'La carte de Cassini et son intérêt géographique', *Bulletin de l'association des géographes français* 251 (1955) 138–47; Monique Pelletier, *La carte de Cassini. L'extraordinaire aventure de la carte de France* (Paris: Presses de l'École nationale des ponts-et-chaussées, 1990). The point coordinates of these parishes are available from Pascal Cristofoli, Marie-Christine Vouloir, Bertrand Duménieu, and Claude Motte, *Des chefs-lieux de Cassini aux communes de France (1756–1999)* (Paris: Laboratoire de démographie et d'histoire sociale [producer]; Paris: EHESS [distributor], 2021).

⁵³ We use the dataset of Cristofoli, Vouloir, Duménieu, and Motte, *Des Chefs-Lieux de Cassini aux Communes de France*.

⁵⁴ Merrick L. Berman, 'Boundaries or Networks in Historical GIS: Concepts of Measuring Space and Administrative Geography in Chinese History', *Historical Geography* 33 (2005) 118–33; Dauphant, *Géographies*; Mark Ravina, 'Algorithmic Maps and the Political Geography of Early-Modern Japan', *Journal of Cultural Analytics* 8 (2023) 1–32; Luca Scholz, 'Deceptive Contiguity: The Polygon in Spatial History', *Cartographica* 54 (2019) 206–16; Stapel, 'Historical Atlas of the Low Countries', p. 7.

except that the latter contains additional parish-level information: each parish's Cassini identifier, its name as it appears on the Cassini map and on the 1793 (*an 3*), 1801 (*an 8*), and 1999 censuses, and its geographic coordinates in RGF93 projection. Attribute tables associated with each shapefile are available in Stata data format (dta) and text delimited format (txt).⁵⁵

The territory of France in 1789

The first shapefile of our atlas represents the extent of Brette's atlas, that is, the territory of the Kingdom of France in 1789. This broadly corresponds to the current territory of mainland France, with three major exceptions: the Duchy of Savoy, the County of Nice, and the Comtat Venaissin.⁵⁶ Together, these territories represent only three percent of France's current territory and two percent of all parishes at the time of the Revolution ([Appendix Table A1](#)). Other exceptions include several small principalities (Montbéliard, Salm), independent cities (Avignon, Mulhouse), counties (Saar-Werden, Sault), and lordships (Montjoie, Mandeure, Bidache).⁵⁷ We display the shapefile of the territory of France in 1789 below ([Fig. 2](#)) – the corresponding shapefile in point form is displayed in [Appendix \(Fig. A4\)](#). Attribute tables associated with these shapefiles contain an identifier for each of these territorial entities, as given by Brette, along with their names (see [Table 1](#)).

Types of jurisdictions

Most of the electoral constituencies for the Estates General of 1789 correspond to the local jurisdictions of the royal justice: bailliages and sénéchaussées. This was the case for eighty-four percent of France's territory and parishes ([Appendix Table A2](#)). The remaining electoral constituencies correspond to alternative jurisdictions, generally because the royal justice did not operate there.⁵⁸ We map the spatial distribution of these jurisdictions below ([Fig. 3](#)). The attribute table associated with this shapefile contains a jurisdiction type identifier along with their names ([Table 1](#)).

Bailliages

The shapefile of bailliages in 1789 is displayed below ([Fig. 4](#)) – the corresponding shapefile in point form is displayed in [Appendix \(Fig. A5\)](#). It exhibits 435 jurisdictions. Two bailliages are not displayed on this map, although they are present in the data files

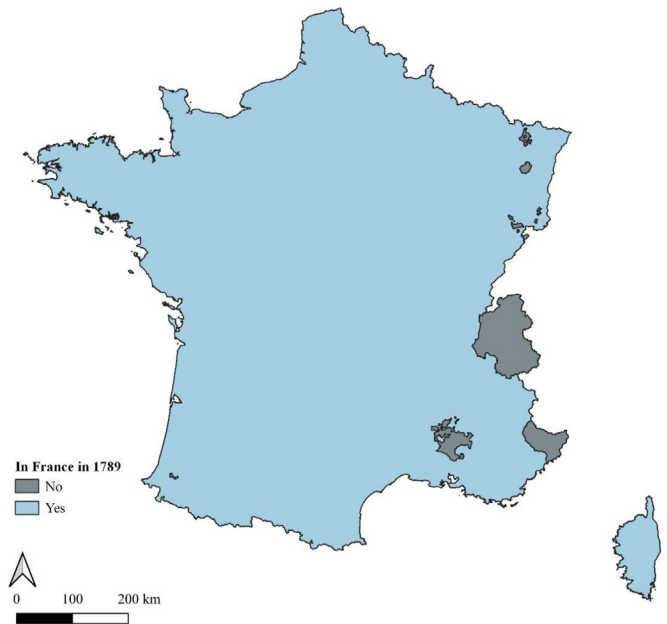


Fig. 2. Extent of the territory of France in 1789. Source: map produced with QGIS and IGN's ADMIN-EXPRESS édition mars 2021, available under a CC0 license.

associated with our atlas. The first is the bailliage of Sarrelouis, since its territory was ceded to Prussia by the Treaty of Paris of November 20, 1815. The second is the bailliage of Bassigny-Barrois, which did not have a territorial constituency for the Estates General of 1789.⁵⁹

The attribute table associated with this shapefile contains several variables that characterize each bailliage, including an identifier, a set of names, a rank, and a seat (*chef-lieu*). Below, we describe each of these variables. They are listed in [Table 1](#).⁶⁰

Bailliage identifiers

Each bailliage is characterized by an identifier according to Brette's nomenclature, under variable name *bailliage*.⁶¹ These identifiers range from 001 to 445, except for the jurisdictions of Dauphiné in the généralité of Grenoble, which combine a number and a letter from 442A to 442F.⁶²

Bailliage names

We provide bailliage names in several forms: a short form that corresponds to the most representative name of the bailliage

⁵⁵ These data files contain name strings in both uppercase and proper forms, while attribute tables (dbf files) contain name strings in uppercase only to avoid compatibility issues with accented letters in GIS processing software.

⁵⁶ Small parts of the northeastern territory of France were ceded to Prussia by the Second Treaty of Paris on November 20, 1815. This included the entire bailliage of Sarrelouis, part of the bailliage of Bouzonville in the généralité of Metz, and areas near Wissembourg and Landau concerning parts of the districts of Haguenau and Wissembourg, as well as the ten imperial cities in the intendance of Strasbourg. These areas are not displayed in our shapefiles.

⁵⁷ These territories were incorporated into the territory of France soon after the Revolution: 1790 for the lordship of Bidache, 1791 for the city of Avignon and the Comtat Venaissin, 1793 for the principalities of Montbéliard and of Salm, the county of Saar-Werden, and the lordship of Mandeure, and 1798 for the city of Mulhouse. The Duchy of Savoy and the County of Nice were only annexed in 1860. Moreover, Brette's atlas considers the island of Corsica to be part of the Kingdom of France, since the National Assembly issued a decree on November 30, 1789 stating that "the island of Corsica is part of the French empire." Finally, the enclaved foreign territory of the commune of Llivia, visible on Brette's atlas, is not included in our GIS, as it still belongs to Spain today.

⁵⁸ Brette distinguished eighteen other types of jurisdictions that formed electoral constituencies: *baronnies*, *comtés*, *districts*, *gouvernances*, *hautes justices*, *juridictions royales*, *pays*, *principautés*, *provinces*, *prévôtés*, *châtellenies*, *vicomtés*, *royaumes*, *sièges royaux*, *souverainetés*, *vigueries*, *villes*, and *élections*.

⁵⁹ Members of this electoral constituency elected one representative from the union of the bailliages of Bourmont and of Lamarche.

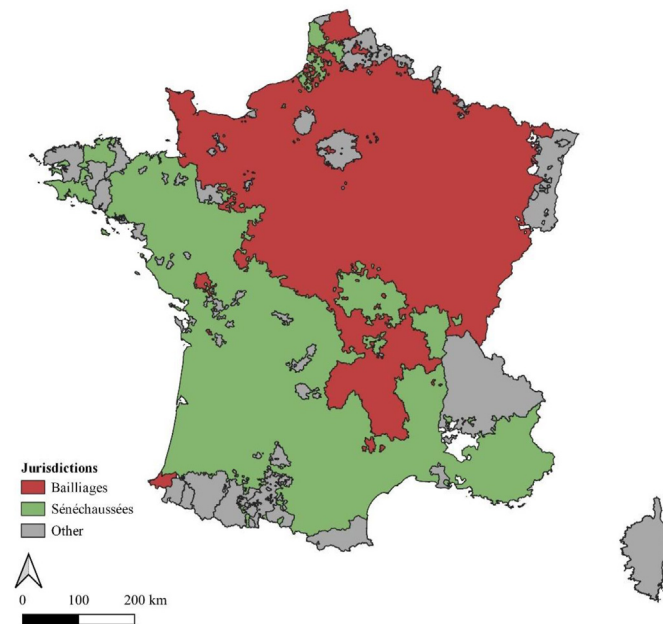
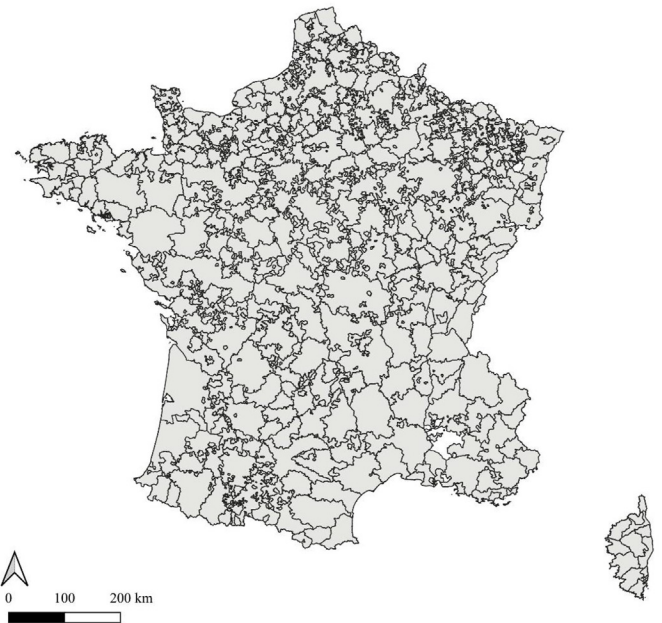
⁶⁰ The attribute table also contains the type of jurisdictions (*type_gen* and *type_det*) and the name of the généralité in which bailliages are located per Brette's classification (*generalite_name*). These thirty-three généralités are displayed in [Appendix Fig. A5](#). Note that seat locations are characterized by a set of identifiers, names, and geographic location (latitude and longitude).

⁶¹ We indicate variable names as they appear in the dta and txt files associated with the shapefiles. Corresponding variable names in attribute tables can be found in [Table 1](#). These differ due to technical constraints imposed by GIS processing software.

⁶² Brette's nomenclature of bailliages is not continuous for two reasons. First, four bailliages were erroneously included in the royal decree of January 24, 1798, convening the Estates General – the bailliages of Ezy (87), Pacy (91), Château-Renard (128), and Schambourg (371). These virtual bailliages did not actually correspond to a jurisdiction and did not exist at the time of the convocation. Second, the electoral constituencies specifically dedicated to the election of representatives of the Clergy in Brittany (dioceses), numbered 422–430, are not represented in our atlas.

Table 1Variables in the *Atlas of Local Jurisdictions* data files.

Shapefile/dataset	Variables		Description
	dta/txt	dbf	
BRETTE_FRANCE	territory	TER_ID	Territory identifier
	territory_name	TER_NS	Territory name (short, upper case)
	territory_name_prop		Territory name (short, proper case)
	territory_name_long	TER_NL	Territory name (long, upper case)
BRETTE_JURISDICTION	territory_name_long_prop		Territory name (long, proper case)
	type_gen	TYPG_ID	Type of jurisdiction identifier (general)
	type_gen_name	TYPG_N	Type of jurisdiction name (general, upper case)
	type_gen_name_prop		Type of jurisdiction name (general, proper case)
BRETTE_BAILLIAGES	type_det	TYPD_ID	Type of jurisdiction identifier (detailed)
	type_det_name	TYPD_N	Type of jurisdiction name (detailed, upper case)
	type_det_name_prop		Type of jurisdiction name (detailed, proper case)
	generalite_name	GEN_N	Généralité name (upper case)
	generalite_name_prop		Généralité name (proper case)
	bailliage	BAIL_ID	Bailliage or equivalent jurisdiction identifier
	bailliage_name	BAIL_NS	Bailliage or equivalent jurisdiction name (short, upper case)
	bailliage_name_prop		Bailliage or equivalent jurisdiction name (short, proper case)
	bailliage_name_long	BAIL_NL	Bailliage or equivalent jurisdiction name (long, upper case)
	bailliage_name_long_prop		Bailliage or equivalent jurisdiction name (long, proper case)
	bailliage_name_brette	BAIL_NB	Bailliage or equivalent jurisdiction original name (proper case)
	bailliage_rank	BAIL_RK	Bailliage or equivalent jurisdiction rank
	bailliage_sup	BAIL_SUP	Identifier of principal bailliage or equivalent jurisdiction
	type_gen	TYPG_ID	Type of jurisdiction identifier (general)
	type_det	TYPD_ID	Type of jurisdiction identifier (detailed)
	cl_flag	CL_FLG	Bailliage chef-lieu flag
	cl_noacass	CL_CASS	Bailliage chef-lieu cassini identifier
	cl_insee	CL_INSEE	Bailliage chef-lieu INSEE 2021 identifier
	cl_nom_cassini	CL_NC	Bailliage chef-lieu cassini name (proper case)
	cl_nom_1793	CL_N1793	Bailliage chef-lieu 1793 name (proper case)
	cl_nom_1801	CL_N1801	Bailliage chef-lieu 1801 name (proper case)
	cl_nom_1999	CL_N1999	Bailliage chef-lieu 1999 name (proper case)
	cl_nom_2021	CL_N2021	Bailliage chef-lieu 2021 name (proper case)
	cl_position_x	CL_X	Bailliage chef-lieu latitude in RGF93
	cl_position_y	CL_Y	Bailliage chef-lieu longitude in RGF93

**Fig. 3.** Types of jurisdictions that formed electoral constituencies in 1789. Source: map produced with QGIS and IGN's *ADMIN-EXPRESS édition mars 2021*, available under a CC0 license.**Fig. 4.** Bailliages in 1789. Source: map produced with QGIS and IGN's *ADMIN-EXPRESS édition mars 2021*, available under a CC0 license.

(*bailliage_name*), usually the name of the city where its seat was located – for instance, Amiens – a long form (*bailliage_name_long*) that also includes its jurisdiction type – for instance, Bailliage d'Amiens – and its original name (*bailliage_name_brette*) as written in Brette. In almost all cases (406 out of 437), long-form names are identical to original names.⁶³

In addition, when ambiguity would arise in short-form names, we qualify them with their jurisdiction type in parentheses. For instance, bailliaiges with identifiers 151 and 316 both have their seat in the commune of Mortagne, but the former refers to the *bailliage* of Mortagne while the latter refers to the *prévôté* of Mortagne. In this case, we modify these short-form names to Mortagne (*bailliage*) and Mortagne (*prévôté*).⁶⁴

Bailliage ranks

Electoral constituencies for the Estates General of 1614 were also based on bailliage jurisdictions. Bailliaiges that had been electoral constituencies during these Estates were qualified as “principal” in 1789, while those that had not been summoned were qualified as “secondary” and attached to a principal bailliage. We qualify the rank of each bailliage (*bailliage_rank*) and provide the principal bailliage to which secondary bailliaiges were attached (*bailliage_sup*).

Bailliage seats

Finally, we provide information on the location of the seat (*chef-lieu*) of each bailliage, that is, the location where the bailliage assembly occurred upon the convocation of the Estates General. In particular, we indicate each seat's Cassini and INSEE identifiers (*cl_noacass* and *cl_insee*), name on Cassini's map (*cl_name_cassini*), name in the censuses of 1793 (*cl_name_1793*), 1801 (*cl_name_1801*), 1999 (*cl_name_1999*), and 2021 (*cl_name_2021*), as well as the geographic coordinates in RGF93 projection (*cl_position_x* and *cl_position_y*). Almost all bailliage names correspond to the location of their seat, with a few exceptions that are flagged by the variable *cl_flag* in the attribute table.⁶⁵

Applications

To conclude this article, we provide two concrete examples of how our atlas of local jurisdictions of Ancien Régime France can be

used in practice: to construct a customary map and to develop alternative visual vocabularies for representing the relationship between space and jurisdictional power as envisioned by contemporaries.

A customary map of Ancien Régime France

As mentioned above, our historical GIS is a necessary preliminary step to mapping the different legal systems that coexisted before the Revolution.⁶⁶ Until their abolition and the advent of the Civil Code in 1804, nearly a hundred different customs regulated various dimensions of people's lives, from the legal status of individuals to marital rules, parental authority, inheritance, economic transactions, or punishment for crimes.⁶⁷ In 1454, Charles VII's ordinance of Montils-lès-Tours initiated a century-long process of codification of customs.⁶⁸ This process clearly delineated the bailliaiges within which each of the ninety-three customs would be enforced.⁶⁹ To construct a customary map of Ancien Régime France, we match each of the 435 bailliaiges in our atlas to its corresponding custom based on Armand Brette's *Recueil de documents relatifs à la convocation des États généraux de 1789*, which reports the customs that applied in about half of the realm's jurisdictions. For the other half, we resort to the original source used by Armand Brette: Bourdot de Richebourg's *Nouveau coutumier général*, which provides the original texts of most of the customs along with their associated jurisdictions. To cover the remaining jurisdictions missing from the *Nouveau coutumier général*, we use several secondary sources.⁷⁰ The resulting spatial distribution of general customs is provided in Fig. 5. This map can be used by future researchers to study the relationship between customary rules and a variety of economic, social, and demographic outcomes.

A network alternative to a 'jigsaw-puzzle' polygon representation

Our atlas is available along two formal representations of space: a polygon form, where each polygon represents a bailliage based on the aggregation of underlying 2021 commune-level polygons, and a

⁶⁶ This application is based on Victor Gay, Paula E. Gobbi, and Marc Goñi, 'The Customary Atlas of Ancien Régime France'.

⁶⁷ See, e.g., Émile Chénon, *Histoire générale du droit français public et privé des origines à 1815* (Paris: Recueil Sirey, 1926). While the northern part of the country was under customary law during the Ancien Régime (*pays de droit coutumier*), the southern part of the country was under written law (*pays de droit écrit*). However, some jurisdictions in written-law country also had a custom, such as the Basque Country, Provence, or Dauphiné. Likewise, some jurisdictions in Lorraine and Alsace also followed written law, although located in customary-law country. In practice, however, written law in these areas was often only complementary to customary law and applied only in the absence of a relevant customary rule, so that these areas may be classified as 'mixed' in terms of their legal system, as emphasized by Béatrice Fourniel and Jacqueline Ventrant-Voyer, 'Une bigarrure de loix dans une même province', in *Atlas historique: Auvergne, Bourbonnais, Velay*, ed. by Stéphane Gomis (Clermont-Ferrand: Université Clermont Auvergne, 2017). As a result, the traditional division of Ancien Régime France into a written-law and a customary-law country along a north-south divide is less clear than previously thought. On this point, see Jean Hilaire and François Terré, *La vie du droit: coutumes et droit écrit* (Paris: Presses universitaires de France, 1994).

⁶⁸ As explicitly stated in Charles VII's ordinance, this process would help customary law achieve greater certainty, stability, and uniformity, and thus, legitimacy. See, e.g., Martine Grinberg, *Écrire les Coutumes. Les Droits Seigneuriaux en France* (Paris: Presses universitaires de France, 2006).

⁶⁹ Our application refers to general customs, which are a set of rules enforced in a given bailliage. In contrast, local customs specified deviations from specific articles of their general customs. For a discussion of these distinctions, see Grinberg, *Écrire les Coutumes* and Gay, Gobbi, Goñi, 'The Customary Atlas of Ancien Régime France'.

⁷⁰ Among other sources, we use Anne Zink, *L'héritier de la maison. Géographie coutumière du sud-ouest de la France sous l'Ancien Régime* (Paris: EHESS, 1993) for the south-west and Laurence Joignon, *Coutumes, familles, successions et alliances en Lorraine: 1670–1900* (Paris: EHESS, 1989) for Lorraine.

⁶³ Long-form names differ from short-form names in thirty-one cases because we omit the name of the seat location when indicated in the original name. For instance, the long-form name of the bailliage with original name Bailliage de La Montagne, séant à Châtillon-sur-Seine is Bailliage de La Montagne, while its short name is La Montagne.

⁶⁴ These cases are relatively rare and further concern the following bailliaiges, which all have their seat in the same commune: Paris (ville et faubourgs) (144) and Paris (hors les murs) (145), Valenciennes (ville) (312) and Valenciennes (prévôté) (314), Metz (ville) (330) and Metz (bailliage) (331), and Arles (sénéchaussée) (384) and Arles (ville et territoire) (385).

⁶⁵ First, some bailliage names correspond to a region, especially in *pays d'Etat*. Pays d'Etat were regions that had retained a representative assembly of the three orders that negotiated tax collection processes with royal commissioners. See Mousnier, *Les institutions de la France sous la monarchie absolue*; Barbiche, *Les institutions de la monarchie française*. For instance, the bailliage of the Pays des Quatre-Vallées (023) had its seat in La Barthe-de-Neste. Second, by 1789, the seats of several bailliaiges had moved from their original locations without changing their names. For instance, the seat of the bailliage of Vouvant (167) had moved from Vouvant to La Châtaigneraie, a dozen kilometers away. Third, some bailliage seats were located outside the territorial jurisdiction of their bailliage. For instance, the bailliage of Andelat (095) had its seat in Murat, about 20 km from Andelat and outside the bailliage. Finally, bailliaiges that were divided between the inner and outer areas of a town had their seat in the same town. This was the case for the bailliaiges listed in footnote 64.

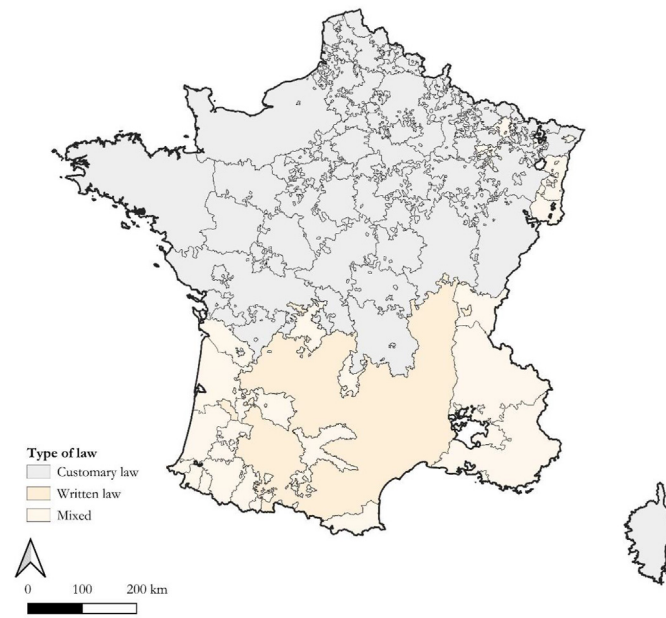


Fig. 5. General customary regions in 1789. Source: map produced with QGIS and IGN's *ADMIN-EXPRESS édition mars 2021*, available under a CC0 license, based on Gay, Gobbi, Goñi, 'The Customary Atlas of Ancien Régime France'.

point form, where each point represents a parish at the time of the Revolution. As noted above, a polygon-based approach poses a methodological challenge if a user wishes to represent early modern relationships between space and jurisdictional power as conceived by contemporaries.⁷¹ Indeed, polygon-based representations assume that the territorial reach of a given jurisdiction changed discretely at its boundaries, while in fact it was more of a continuous process. Jurisdictional conflicts on the outskirts of bailliages – due to the confrontation of expansionist bailliage agents – resulted in the fragmentation of their territories, as highlighted by the non-compactness of many bailliage shapes. Thus, the mental projection of political space by contemporaries resembled a collection of places rather than an aggregation of contiguous spatial units. This explains why, despite the availability of mapping technology, early modern political spaces were generally not represented as maps but as lists of places—a mode of representation at odds with the polygon-based nature of GIS software.⁷² More broadly, such 'jigsaw-puzzle view of discrete, bounded territories' might be considered anachronistic, reflecting our own preconceptions of what a political territory should be and obscuring our understanding of the relationship between space and power in Ancien Régime France.⁷³

The point-based layer avoids polygons altogether and can be the basis for alternative visual vocabularies to represent the nature of political territoriality as envisioned by contemporaries. In particular, we propose a network approach that represents the jurisdictional hierarchy between bailliage seats and the parishes within their reach, making the interpenetration of space and power

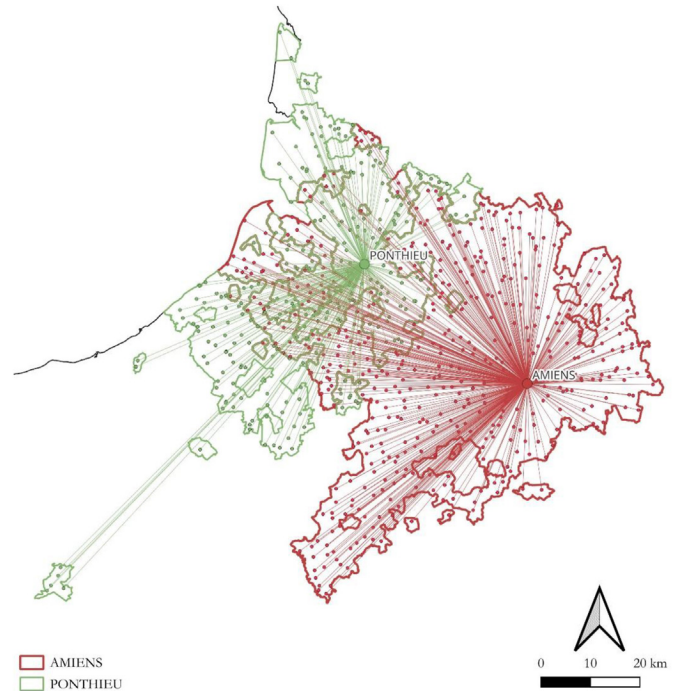


Fig. 6. Network representation of parishes of the bailliages of Ponthieu and Amiens. Source: map produced with QGIS, IGN's *ADMIN-EXPRESS édition mars 2021*, and Cristofoli, Vouloir, Duménieu, and Motte's *Des chefs-lieux de Cassini aux communes de France*, available under a CC0 license.

apparent. Specifically, this implies connecting each parish to its corresponding bailliage seat through a vector.⁷⁴ An example of this interpenetration for the bailliages of Ponthieu (séant à Abbeville) and Amiens is shown in Fig. 6. The figure covers about the same area as Fig. 1 and shows the corresponding polygons for reference. Our network approach can be easily generalized to the entire set of bailliages. This alternative representation reflects more accurately the way contemporaries conceived political territories: a set of intertwined hierarchies with overlapping areas of jurisdictional conflict.

Conclusion

This article describes the construction and content of a curated atlas of the local jurisdictions that were at the core of Ancien Régime's judicial apparatus: bailliages. These jurisdictions epitomize the interweaving of legal institutions that characterized Ancien Régime France. Their haphazard development alongside the consolidation of the monarchy and the dissolution of feudal strongholds was such that the exact territorial extent of these jurisdictions was not known with precision by the royal authority, which never mapped their spatial distribution. Based on Brette's *Atlas des bailliages* and the convocation of the Estates General of 1789, we fill this gap and provide an atlas of bailliages in shapefile form along with

⁷¹ Berman, 'Boundaries or Networks in Historical GIS'; Ravina, 'Algorithmic Maps and the Political Geography of Early-Modern Japan'; Scholz, 'Deceptive Contiguity'.

⁷² Patrick Boucheron, 'Représenter l'espace féodal: un défi à relever', *Espaces temps* 68–70 (1998), 59–66; Dauphant, *Géographies*.

⁷³ Martin M. Lewis and Karen E. Wigen, *The Myth of Continents. A Critique of Metageography* (Los Angeles: University of California Press, 1997), p. 12. More precisely, '[b]y a jigsaw-puzzle view of the world, we mean the expectation that a proper will always show a set of sharply bounded units that fit together with no overlap and no unclaimed territory.' Lewis and Wigen, *The Myth of Continents*, p. 11.

⁷⁴ This approach is similar to Berman's and Bol's approaches to Chinese historical geography. See Berman, 'Boundaries or Networks in Historical GIS'; Petter K. Bol, 'Creating a GIS for the History of China' in *Placing History: How Maps, Spatial Data, and GIS Are Changing Historical Scholarship*, ed. by Anne K. Knowles and Amy Hillier (New York: ESRI Press, 2008), pp. 27–59. An alternative approach – used by Ravina in the context of Tokugawa Japan – is to perform a Voronoi tessellation around each parish to reconstruct abstract pseudo-boundaries by aggregating the resulting Voronoi polygons. See Ravina, 'Algorithmic Maps and the Political Geography of Early-Modern Japan'.

relevant information at the bailliage level. Our database is available in the open-access repository 'Bailliages in 1789 France' hosted on the Harvard Dataverse.

Although GIS methods have reached a central place in historical geography scholarship and their importance to the humanities is growing, their application to the study of administrative jurisdictions in Ancien Régime France remains elusive. This has limited the ability of historians and social scientists to study the historical processes at play through the lens of spatial methods. By providing one of the first GIS for Ancien Régime France, our work will empower historical geographers to explore the spatial dynamics of this period, particularly those related to the functioning of a complex jurisdictional landscape and the Estates General of 1789. Compared to previous work on the geography of bailliages, our GIS covers the entire French territory and precisely maps the dramatic heterogeneity of bailliage shapes and sizes, as well as their numerous enclaves—sometimes dozens of kilometers from their bailliage seat. These advances will enable historical geographers to uncover the spatial intricacies of power and governance structures that contributed to the dislocation of France's Ancien Régime society and, ultimately, to the French Revolution.

Recognition of funders

Victor Gay gratefully acknowledge funding from the Agence Nationale de la Recherche (ANR) under grant ANR-17-EURE-0010 (Investissements d'Avenir program) and under grant ANR-20-CE38-0015 (ObARDI). The research was co-funded by the European Union (ERC, IDED, grant agreement 947654). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

CRediT authorship contribution statement

Victor Gay: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Paula E. Gobbi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Marc Goñi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Data availability

The data used in this article is available on the Harvard dataverse at <https://doi.org/10.7910/DVN/T8UXHK>.

Acknowledgement

We thank Gauthier Aubert, François Godicheau, Philippe Hamon, and Sébastien Poulanc for fruitful discussions.

APPENDIX

Appendix Figures

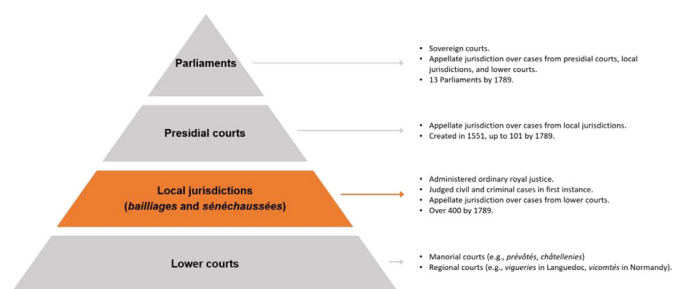


Fig. A1. Organization of Ancien Régime jurisdictions. Source: authors' own creation.

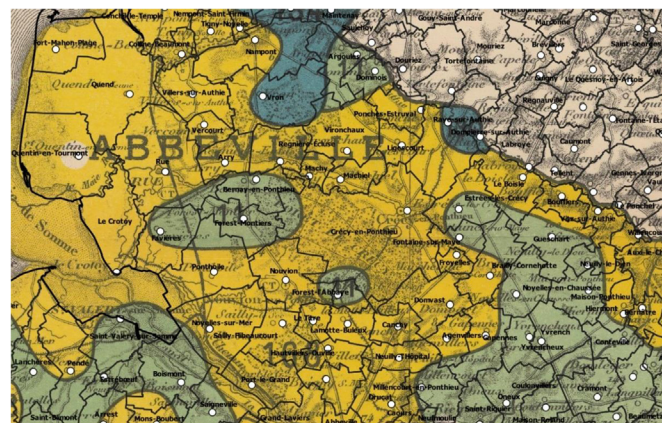


Fig. A2. Overlay between IGN's Communes shapefile layer and Brette's georeferenced Atlas. Produced with QGIS. Black lines indicate current communes boundaries and white dots these communes' seats. Source: Brette, *Atlas des bailliages* and IGN's ADMIN-EXPRESS édition mars 2021, available under a CCO license.

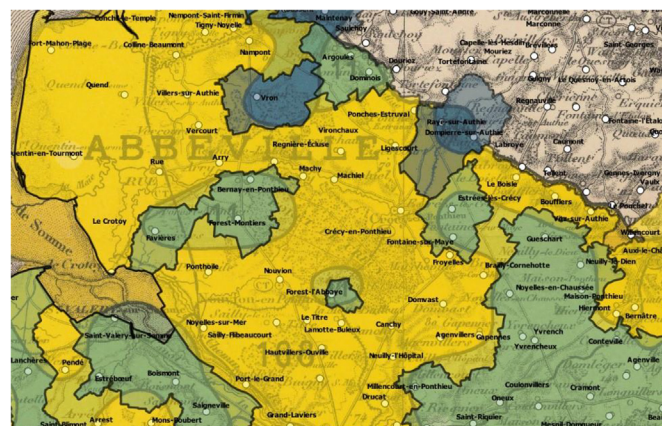


Fig. A3. Dissolved layer based on bailliages attributes. Produced with QGIS. Source: Brette, *Atlas des bailliages* and IGN's ADMIN-EXPRESS édition mars 2021, available under a CCO license.

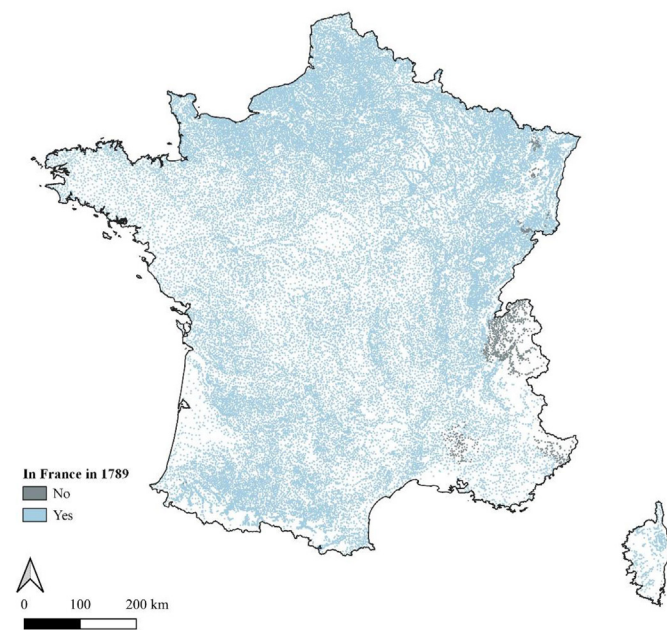


Fig. A4. Extent of France's territory in 1789. Source: map produced with QGIS, IGN's *ADMIN-EXPRESS édition mars 2021* and Cristofoli, Vouloir, Duménieu, and Motte's *Des chefs-lieux de Cassini aux communes de France*, available under a CC0 license.

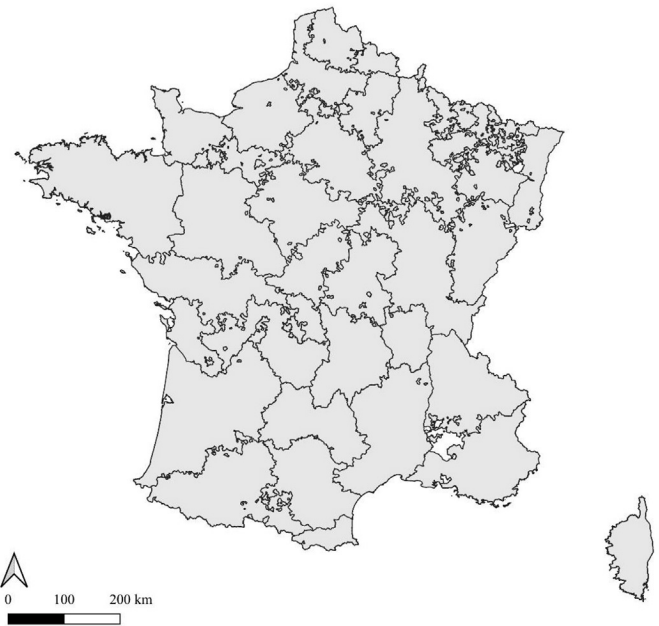


Fig. A6. Généralités in 1789. Source: map produced with QGIS and IGN's *ADMIN-EXPRESS édition mars 2021*, available under a CC0 license.

Appendix Tables

Table A1

Extent of the territory of France in 1789

Territory	Parishes		Area (km ²)	
	Total	%	Total	%
0 France	42,771	97.67	531,865	96.85
Outside of France	1021	2.33	17,291	3.15
I État d'Avignon	4	0.01	86	0.50
II Comtat Venaissin	105	0.24	1,992	11.52
III Souveraineté de Bidache et duché de Gramont	3	0.01	47	0.27
IV Seigneurie de Mandœuvre	1	0.00	15	0.09
V Principauté de Montbéliard	37	0.08	177	1.02
VI Ville de Mulhouse	5	0.01	38	0.22
VII Comté de Saar-Werden	36	0.08	292	1.69
VIII Principauté de Salm	23	0.05	268	1.55
IX Comté de Sault	9	0.02	339	1.96
X Seigneurie de Montjoie	15	0.03	108	0.62
XII Duché de Savoie	684	1.56	11,144	64.45
XIII Comté de Nice	99	0.23	2,784	16.10

Table A2

Types of jurisdictions that formed electoral constituencies in 1789

Jurisdiction	Parishes		Area	
	Total	%	Total	%
Bailliages	22,033	51.51	233,588	43.92
Sénéchaussées	13,837	32.35	215,300	40.48
Other	6,901	16.13	82,877	15.6

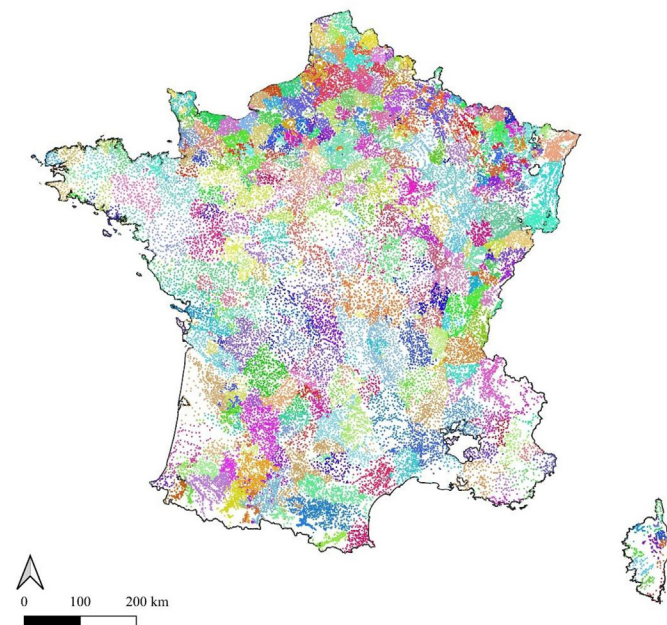


Fig. A5. Bailliages in 1789. Source: map produced with QGIS, IGN's *ADMIN-EXPRESS édition mars 2021* and Cristofoli, Vouloir, Duménieu, and Motte's *Des chefs-lieux de Cassini aux communes de France*, available under a CC0 license.



Contents lists available at ScienceDirect

Explorations in Economic History

journal homepage: www.elsevier.com/locate/eeh

Shorter Article

The Customary Atlas of Ancien Régime France[☆]Victor Gay^a, Paula E. Gobbi^{b,c,*}, Marc Goñi^{d,c}^a *Département of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France*^b *Université libre de Bruxelles (ECARES), Belgium*^c *CEPR, United Kingdom*^d *University of Bergen, Norway*

ARTICLE INFO

Dataset link: <https://doi.org/10.7910/DVN/HY E209>

JEL classification:

K00

N43

P48

Keywords:

Customary law

Custom

Institution

Legal origin

Ancien Régime

France

ABSTRACT

Customary law governed most European societies during the Middle Ages and early modern period. To better understand the roots of legal customs and their implications for long-run development, we introduce an atlas of customary regions of Ancien Régime France. We also describe the historical origins of French customs, their role as a source of law, and their legal content. We then offer some insights into the research possibilities opened by this database.

1. Introduction

A large literature argues that historical customs matter for long-run development (Acemoglu et al., 2005; Nunn, 2009, 2020). This consensus is based primarily on evidence drawing on pre-industrial Africa, in part because of the availability of comprehensive databases such as Murdock's (1957, 1967) Ethnographic Atlas and Standard Cross-Cultural Survey (Murdock and White, 1969).¹ In contrast, empirical evidence drawing on pre-industrial Europe remains scarce.² This gap is particularly problematic as major economic processes – the Industrial Revolution and the Demographic Transition – originated in Europe at a time when customs were still a dominant source of law. Yet, we still know relatively little about how customs interacted with these processes. A primary reason

[☆] Victor Gay gratefully acknowledge funding from the Agence Nationale de la Recherche (ANR), France under grant ANR-17-EURE-0010 (Investissements d'Avenir program) and under grant ANR-20-CE38-0015 (ObARDI). The research was co-funded by the European Union (ERC, IDED, grant agreement 947654). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

* Correspondence to: Avenue F.D. Roosevelt 42, B-1050 Brussels, Belgium.

E-mail addresses: victor.gay@tse-fr.eu (V. Gay), paula.eugenia.gobbi@ulb.be (P.E. Gobbi), marc.goni@uib.no (M. Goñi).

¹ See Lowes (2021, Table 1) for a list of studies in economics using these datasets to infer the causal relationship between ancestral factors and contemporary outcomes.

² An exception concerns inheritance customs in mid-nineteenth- and early-twentieth-century Baden-Württemberg. In particular, Hager and Hilbig (2019) examine how inheritance customs there affected long-run political and social inequality; Huning and Wahl (2021), regional development; Bartels et al. (2020), the distribution of income; Süß (2023), household formation and gender disparities. In contrast to this literature, we focus on customary rules over an entire polity and that predate the nineteenth century.

<https://doi.org/10.1016/j.eeh.2024.101588>

Received 19 December 2023; Received in revised form 8 March 2024; Accepted 10 March 2024

Available online 12 March 2024

0014-4983/© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

for this gap in knowledge is the unavailability of a “Murdock” counterpart for pre-industrial Europe.³ Indeed, the territorial extent of historical customary regions in Europe is generally unknown to us as they were eliminated during the nineteenth century as a result of the transition to modern states, the introduction of civil codes, and the development of alternative legal organizations (Padua-Schioppa, 1997). However, an accurate knowledge of their spatial distribution is crucial to properly study the role of customs for Europe’s long-run development.

In this article, we present the *Customary Atlas of Ancien Régime France*. This original atlas maps boundaries of the customary regions prevalent in France from the recording of customs in the mid-fifteenth century to the advent of the Civil Code in the early nineteenth century (Chénon, 1926, p. 7–9). In particular, building on Gay et al. (2023a) *Atlas of Local Jurisdictions of Ancien Régime France* and a variety of archival and secondary sources, we construct a shapefile of the 141 customary regions that existed in the early modern period and make it openly available for further research (Gay et al., 2023b).

Customs regulated many dimensions of people’s lives, from the legal status of individuals to marriage rules, parental authority, inheritance, economic transactions, or punishments for crimes. The case of France is paradigmatic of pre-industrial Europe, much of which was under customary law. In France, customs first developed among the tribes that inhabited its territory after the fall of the Western Roman Empire. The expansion of feudalism during the Middle Ages further reinforced the territoriality of customs, which evolved along the specificities of local contexts. Then, in the mid-fifteenth century, the monarchy initiated a century-long process of recording customs, which fixed both their content and their territoriality (Grinberg, 2006). Customs were eventually abolished during the French Revolution and replaced by the Civil Code in 1804.

The Customary Atlas we propose is the first of its kind for Europe as it covers an entire polity and is disseminated as a historical geographic information system (GIS). Prior to our work, legal historians had sketched (paper) maps of historical customs covering a few European regions, such as the German state of Baden–Württemberg (Krafft, 1930; Röhm, 1957), Romandy in French-speaking Switzerland (Poudret (1998), Belgium and northern France (Gilissen, 1958, 1979), (Joignon, 1989), or southwestern France (Zink, 1993). Related to our work, Klimrath (1843) proposed a map of general customs in Ancien Régime France, which we reproduce in Appendix Figure A.1. In contrast to this early attempt, our Customary Atlas maps both general and local customs, draws on a variety of archival and secondary sources, and uses GIS techniques, making it readily usable for further research.⁴

The Customary Atlas can be used to study the short- and long-term socioeconomic implications of different regulations of inheritance, marriage, parental authority, or economic activities. Given the substantial spatial variability in the territorial boundaries of customs resulting from their complex histories, the Customary Atlas will allow researchers to exploit local variation for causal identification. For instance, Gay et al. (2023c) use data from this atlas to study how inheritance customs affected fertility in late eighteenth-century France, finding that locations under egalitarian inheritance had lower fertility than those where a single heir received most of the inheritance. In addition, the Customary Atlas will also contribute to the law and economics literature that examines the impact of legal origins on financial and economic development outcomes (La Porta et al., 2008). Among other potential uses, it will help improve our understanding of the historical origins and consequences of the French civil law tradition and the legacies of customary and Roman written law (Le Bris, 2019).

The remainder of this article is organized as follows: Section 2 discusses the history of customary law in France, Section 3 explains the mapping methodology we use, Section 4 describes the data files of the Customary Atlas, and Section 5 presents various potential uses of this resource.

2. The history of customary law in France

Origins. Initially, customs encompassed a set of unwritten rules that regulated the lives of people in a given locality with an authority equal to that of formal laws (Gilissen, 1979, p. 27).⁵ During the Middle Ages and early modern period, most of Western Europe was under customary law (Gilissen, 1962b; Grinberg, 2006).⁶ In France, customs derived from the laws of the tribes that populated its territory after the fall of the Western Roman Empire—the Burgundians, the Visigoths, the Salian Franks, and the Ripuarian Franks. Originally, these were *personal* rather than *territorial* laws: they applied to a particular group of people rather than to the inhabitants of a particular locality.⁷ From the ninth century on, these laws took on a territorial dimension. This shift – driven by the increasing mixing of groups of different origins – made it difficult to distinguish between the appropriate laws that applied to particular individuals (Chénon, 1926, p. 128). The expansion of feudalism during the Middle Ages further reinforced the territoriality of these laws, as residents of a territory under a feudal lord had to abide by the laws that emanated from that ruler.

³ Giuliano and Nunn (2018) attempt to correct the under-representation of Europe in the Ethnographic Atlas by adding 17 ethnolinguistic groups from Europe. In contrast, the Customary Atlas delineates 141 historical customary boundaries within France, illustrating that customs also varied within ethnolinguistic groups in pre-industrial Europe.

⁴ General customs specified a set of rules that were enforced throughout a judicial district. In contrast, local customs specified deviations from specific rules of the general custom in particular locations. We provide further details on this distinction in Section 2.

⁵ Canonists required two elements for a custom to have the force of law: the repetition of the same behavior by people from time immemorial and popular belief in the binding force of the custom (Kim, 2021, p. 44). For instance, the custom of Normandy (1510) defined the criteria for establishing a custom as a combination of the consent of the people subject to the custom, the custom being observed and regarded as law, the frequency and habitual nature of this way of life, its long-standing notoriety, and its usefulness to the people of the locality (Grinberg, 2006, p. 127).

⁶ See Appendix Table B.1 for a list of historical customs in Western Europe along with their publication dates.

⁷ For instance, Ripuarian law stipulated that inhabitants of Ripuarian territory who belonged to other tribes should be judged according to the laws of their tribes of origin rather than according to Ripuarian law (*Lex Ripuaria*, tit. XXXI, art. 3–4, quoted in Chénon, 1926, p. 124).

The initial diversity of tribes that populated the territory of France contributed to the great heterogeneity of customs. Since northern regions were dominated by the Salian and Ripuarian Franks, the *Salic Law* and *Lex Ripuaria* became the basis for the development of customs there. In contrast, southern regions were dominated by the Burgundians and Visigoths, whose laws consisted of the *Lex romana Burgundionum*, *Lex Theodosii*, and *Corpus Juris Civilis*—all based on principles derived from Roman law. Thus legal rules there evolved into written law. As a result, France was soon divided into a customary-law country (*pays de droit coutumier*) in the north and a written-law country (*pays de droit écrit*) in the south, a partition that lasted until the Revolution.⁸

In addition to the north–south divide between customary and written law, a mosaic of customs emerged during the Middle Ages (Kim, 2021, p. 29–63). As customs shifted from personal to territorial, local variations in geography and the decentralized nature of feudalism contributed to the increasing fragmentation of customary rules across the territory (Chénou, 1926, p. 487).

Early-stage writing of customs. By the time feudalism had consolidated and the territory of the royal domain of France had stabilized, territorial jurisdictions (*districtus*) were regulated by customs to which inhabitants had to adhere. However, customs were not entirely static, and the line between customs as representing people's behavior and customs as regulating it was still blurred. In fact, customs could still evolve with local usage, especially since they were passed down orally from one generation to the next. Such uncertainty led to a proliferation of lawsuits and to cumbersome and time-consuming court rulings, as judges had to navigate the intricacies of gathering evidence, taking testimony, and analyzing historical precedents to establish the validity of a given custom in a locality—a lengthy and costly process known as *enquêtes par turbe* and formalized in 1270 by an ordinance of Louis XI (Kim, 2021, p. 41). Overall, this fed into demands for written documentation of customs as early as the late twelfth century (Chénou, 1926, p. 491).

The process of writing customs effectively reduced uncertainty about their content and territoriality. The earliest written customs were compiled in books (*coutumiers*) produced by legal practitioners who were primarily motivated by the practicality and usefulness of such texts for court rulings. They first appeared in early thirteenth-century Normandy with the *Grand coutumier de Normandie*, first in Latin (1200–45), then in French (1270–80). Other customs soon followed: the custom of Vermandois (c. 1253); the *Livre de justice et de plet* (1260–70) and the *Établissements de Saint Louis* (1272–3) for the customs of Orléanais, Paris, Touraine, Anjou, and Maine; the custom of Beauvaisis (1283); the custom of Champagne (1289) (Chénou, 1926, p. 553–7). Throughout the fourteenth century, customary texts continued to proliferate with, e.g., the custom of Brittany (1312–25), leading to a first compilation in the *Grand coutumier de France* in 1519 (Laboulaye and Dareste, 1868). However, these early manuscripts were not legal texts with the force of law, but rather reference material used by lawyers to facilitate litigation—although some *coutumiers* had a quasi-official nature, such as those of the customs of Normandy and Brittany (Kim, 2021, p. 43–4).

Official writing of customs. In April 1454, Charles VII's ordinance of Montils-lès-Tours launched the royal campaign to record the customs of the realm. This ordinance explicitly stated the kings' motives: by putting customs into writing, customary law would achieve greater certainty, stability, and uniformity, and thus legitimacy (Kim, 2021, p. 64–91). More broadly, his aim was to create a uniform body of customs throughout the realm to consolidate the authority of the monarchy over the mosaic of feudal powers in the aftermath of the Hundred Years' War (Grinberg, 1997).⁹

Five key elements initiated this codification process in the fifteenth and sixteenth centuries (Gilissen, 1962a, 1962b; Grinberg, 2006). First, written documents provided judicial authorities with a degree of certainty about the content and territoriality of customs. It also enabled these authorities to clearly distinguish customary law from rules emanating from other authorities, such as the Church. Second, the king was known as the “guardian of customs”, so their writing contributed to the consolidation of the monarchy. Third, the monarchy had long sought to unify the legal system within its realm, beginning with the royal domain.¹⁰ In fact, several smaller customs were absorbed into larger ones during the writing process. Fourth, Roman law was the only source of law taught in universities in the fifteenth and sixteenth centuries. Judges, who were increasingly educated in universities, were therefore relatively unaware of local customs as they were passed down orally (Gilissen, 1962b, p. 86). The writing of customs therefore helped to limit the expansion of Roman law into customary-law country. Finally, the writing of customs contributed to abolish the most unreasonable ones.¹¹ All these factors led to an official mandate that would make customary law resemble civil law. Upon their recording, customs ceased to evolve and became the pillar of monarchical law in France until the Revolution.

While in practice only a few customs were initially written and promulgated under Charles VII and Charles VIII, most were eventually codified after Louis XII reiterated the call for the recording of customs in his edict of March 1505 (Klimrath, 1837, p. 10). By and large, this codification process was completed by the end of the sixteenth century.

General and local customs. Although uniformity was a primary goal of the codification of customs, it was not fully achieved. In fact, this process led to the distinction between *general* and *local* customs. General customs provided a set of rules that were enforced throughout a judicial district. In contrast, local customs specified deviations from articles of their general custom in particular locations (Grinberg, 2006). This usually involved only a few articles, although some local customs were more comprehensive than

⁸ This division was formally established by Louis IX's ordinance of Vincennes in 1251, which separated the territory of *consuetudo gallicana* from that of *jus scriptum* (Hilaire and Terré, 1994, 101–3). Klimrath (1837) original map is displayed in Appendix Figure A.1 and the resulting division of France, in Appendix Figure A.2. See also Fig. 1 for a map based on data in the Customary Atlas. Note that Philip IV's ordinance of 1312 recognized Roman law as customary law where it was the source of law (Olivier-Martin, 1948, p. 122).

⁹ Appendix C provides further details on the process of recording customs.

¹⁰ An early attempt at unification is the *Lex Saxonum* of the late eighth century under the reign of Charlemagne.

¹¹ In fact, prior to the fourteenth century, the king – as guardian of customs – had intervened on several occasions to eradicate customs that were deemed either unjust, unreasonable, or perverse—a royal right that had been unchallenged since the eleventh century (Olivier-Martin, 1938).

others. The success of the harmonization process varied across the territory. In Normandy and Brittany, for instance, general customs covered large areas but included many local customs. In contrast, in the northeast, the territorial coverage of general customs was smaller but less subject to local deviations.

Most customs, either general or local, were revised in the late sixteenth century (Kim, 2021, p. 92–115).¹² However, their content was rarely changed, only the clarity of their writing in the French vernacular (Grinberg, 2006, p. 79–80). Thus, the content of most customs remained stable until the French Revolution. In 1804, the law of Ventôse 30, year XII (art. 7) repealed general and local customs relating to matters covered by the Civil Code (Gilissen, 1979, p. 250).¹³

The content of customs. The content of most customs resulting from the codification process described above was compiled in the *Nouveau coutumier général* (Bourdod de Richebourg, 1724). Despite their diversity, some common themes emerge among the vast number of customs: the legal status of individuals and goods, marriage rules, parental authority, illegitimacy, inheritance, economic transactions, and crimes. We provide further details on each of these common themes in Appendix D. In short, customs could differ in how they defined movable and non-movable goods, or in the rights of different types of individuals (nobles, commoners, minors, illegitimate children, etc.). Customary marriage rules differed with respect to the validity of marriages, the rights of husbands to their wives' assets, dowries, and punishments for adultery. Because mortality rates were high, customary rights of widows were thoroughly defined, with variations in their rights to their deceased husbands' assets. The timing and extent of parental authority were also regulated. In customs closer to Roman law, the strict authority of the *patria potestas* prevailed, while in others, parental authority was limited to the duty of raising children. Economic transactions related to taxes, prescriptions (claims), obligations, contracts, and financing (rents, loans, debts, etc.) were also regulated differently by different customs.

3. Mapping methodology

To help researchers better understand the roots and implications of customs, we propose a historical GIS of their spatial distribution in early modern France: the *Customary Atlas of Ancien Régime France*.

To construct the Customary Atlas, we begin with the geography of judicial districts in which customs were enforced: *bailliages*.¹⁴ We rely on Gay et al. (2023a) *Atlas of Local Jurisdictions of Ancien Régime France*, a shapefile of the 435 judicial districts that existed before the Revolution.¹⁵ We then match each judicial district to the custom that applied there. Our primary source is Brette's 1904b, 1915 *Recueil de documents relatifs à la convocation des États Généraux de 1789*, which reports the specific customs that applied in about half of the judicial districts of the realm. Unfortunately, Armand Brette was only able to cover half of the territory by the time of his death in 1912. For most of the rest of the territory, we resort to the original source used by Brette: the *Nouveau coutumier général* (Bourdod de Richebourg, 1724), which provides the original texts of most customs together with their associated judicial districts. To cover the few remaining districts, we use Zink (1993) for the Southwest, Joignon (1989) for Lorraine, and additional archival sources drawn from Gouron and Terrin (1975) seminal bibliography of customs. The output of this task is a correspondence table between the 435 judicial districts that existed in Ancien Régime France and the 141 customs included in the Customary Atlas.¹⁶ Finally, we aggregate judicial districts with the same custom into single customary regions. On average, a customary region includes three judicial districts, although some regions include many more, with up to 22 judicial districts for the custom of Normandy.¹⁷

We also record which judicial districts were under written law—a distinction that is important in the law and economics literature. As the historiography emphasizes, the increasing interest in the study of Roman law after the twelfth century led to the coexistence of both customary and written law in many regions (Olivier-Martin, 1948, p. 111). In particular, some judicial districts located in written-law country also had a custom, e.g., in the Basque country, Provence, or the Dauphiné (Poumarède, 1972; Zink, 1993). Likewise, some judicial districts in Lorraine and Alsace also followed written law, even though they were located in customary-law country (Ganghofer and Levesse, 1977; Joignon, 1989). In practice, however, written law in these areas was only complementary to customary law and only applied in the absence of a relevant customary rule. In Fig. 1, we display the distribution of written- and customary-law areas according to the Customary Atlas, along with areas where both systems were present, which we refer to as “mixed” (Fourniel and Vendrand-Voyer, 2017). As can be seen from this map, Klimrath (1843) approach was not entirely accurate, and the traditional partition of Ancien Régime France into a written-law and a customary-law country along a north–south divide is less clear-cut than previously thought (Hilaire and Terré, 1994, p. 157–84).

In addition, while Klimrath (1843) map focuses only on general customs, we consider both general and local customs. This is important for two reasons.¹⁸ First, a crucial feature of customary law is that it was inherently local. Thus, the boundaries of a given

¹² The official motivation for the revision of many published customs was that their original minutes (*procès-verbaux*) had been lost, making it necessary to resort again to the costly process of proving the existence of a disputed custom (Kim, 2021, p. 90).

¹³ Previous revolutionary laws had already abolished some parts of customary law. For instance, the law of 17 Nivôse, year II (January 6, 1794) had abolished customary inheritance laws.

¹⁴ The ordinance of Charles VIII of January 2, 1493, specified that these were the relevant jurisdictions for the enforcement of customary law (Isambert et al., 1825). These jurisdictions were also called *sénéchaussées* in the southern parts of the realm.

¹⁵ This work is based on Brette (1904a) *Atlas des bailliages et juridictions assimilées*, which draws on the minutes of bailliage electoral assemblies convened for the Estates General of 1789. In particular, see Figure 4 in Gay et al. (2023a).

¹⁶ The list of all sources used is available in Appendix Table B.2.

¹⁷ If we exclude local customs and instead attribute their general custom, customary regions include an average of four judicial districts, with up to 45 for the custom of Normandy.

¹⁸ In Appendix Figure A.3, we display the spatial distribution of customs when we disregard the distinction between local and general customs and focus only on general customs. We also provide this map as a shapefile because it may be useful to some users.

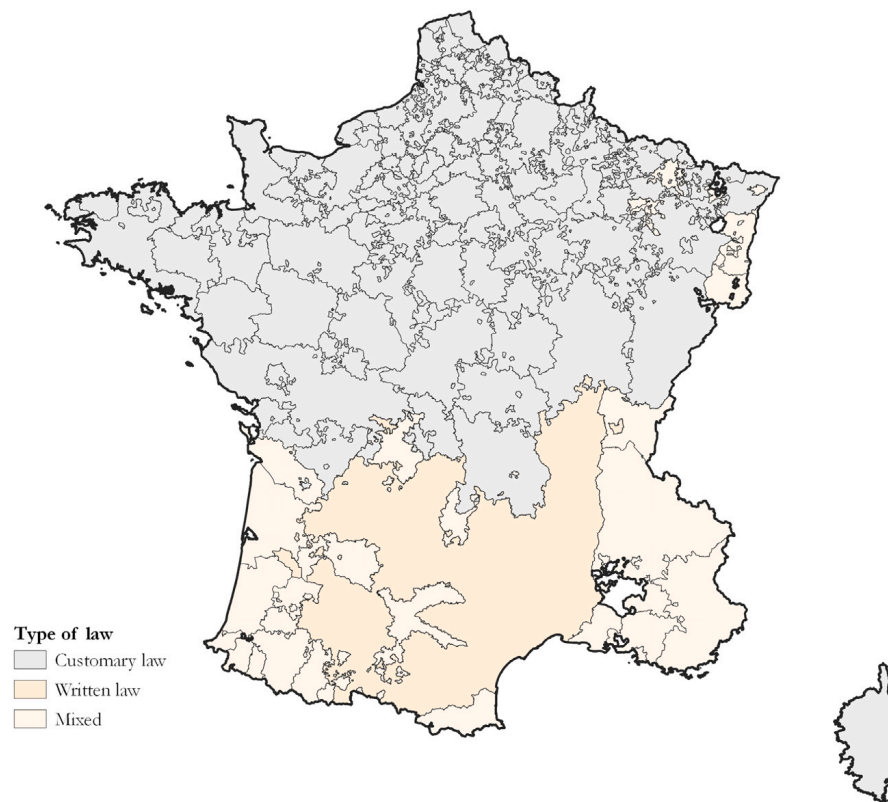


Fig. 1. Customary Regions in Ancien Régime France.

Notes. This figure displays the spatial distribution of customary regions in Ancien Régime France together with the division into customary-law and written-law country, along with mixed areas where both types of law co-existed. The corresponding shapefile is available from [Gay et al. \(2023b\)](#).

customary region ultimately depended on local customs. Second, some general customs contained several local customs, while others knew no such deviation—this was generally the case for general customs covering a relatively small area, such as the custom of Dourdan. As a result, the distinction between general and local customs was a matter of spatial and administrative organization rather than differences in the content of the custom.¹⁹ Thus, relying on judicial district boundaries has the advantage of mapping more customs than was previously done. Finally, the Customary Atlas includes local customs only insofar as their extent corresponds to at least one judicial district. This implies that some local customs – those smaller than a district, *e.g.*, those enforced only in a particular town – are not accounted for.²⁰ We leave the construction of an atlas of customs at the city level for future work.

Overall, the Customary Atlas advances on [Klimrath \(1843\)](#) in terms of precision, as it characterizes the territoriality of 141 customs rather than 52—for instance, [Klimrath \(1843\)](#) exhibits a single custom for the généralité of Normandy, while we include another 20 within that territory.²¹

4. Data description

4.1. Extent of the Atlas

The *Customary Atlas of Ancien Régime France* covers the territory of the Kingdom of France as of 1789. This corresponds to the current territory of mainland France, with three major exceptions: the Duchy of Savoy, the County of Nice, and the Comtat Venaissin. Other exceptions include several small principalities (Montbéliard, Salm), independent cities (Avignon, Mulhouse),

¹⁹ We do not make subjective decisions about which custom should be classified as general or local. Rather, we follow the classification given in the *Nouveau coutumier général* ([Bourdot de Richebourg, 1724](#)).

²⁰ We refer the interested reader to the bibliography of customs in [Gouron and Terrin \(1975\)](#) and to [Bourdot de Richebourg \(1724\) Nouveau coutumier général](#), which comprises four volumes of more than 1200 pages each and describes the content of more than 500 customs, both general and local.

²¹ In this respect, the Customary Atlas is close to [Voltaire \(1878, p. 229–30\)](#) well-known assessment that “[t]here are, it is said, one hundred and forty-four customs in France which have the force of law; these laws are almost all different. A man who travels in this country changes laws almost as often as he changes horse post”.

counties (Saar-Werden, Sault), and lordships (Montjoie, Mandeuze, Bidache) that were incorporated into France soon after the Revolution.²²

4.2. Customs' attributes

Each custom is characterized by several attributes: an identifier, a name, a year of publication, a type, and, for local customs, a general custom of reference.

Identifiers. We create a unique identifier for each custom. To account for the relationship between general and local customs, we use four-digit identifiers, with the leading two digits identifying general customs and the trailing two digits identifying local customs. Specifically, we sort the 94 general customs in our data alphabetically and assign them an identifier ranging from 01 to 94.²³ Similarly, we sort the 48 local customs in our data by general custom of reference and assign them an identifier ranging from 01 to 20—this identifier resets to 01 across general customs.²⁴ We then concatenate these two two-digit identifiers into four-digit identifiers, assigning the trailing digits 00 to general customs. For instance, the custom of Normandy – a general custom – has identifier 6400, while the custom of Alençon – a local custom within the general custom of Normandy – has identifier 6401. Finally, we assign identifier 0000 to written-law regions.

Names. Customs names are provided in both short and long forms. Long forms correspond to original the original names of customs found in archival sources. For instance, the long-form name of the general custom of Normandy is Coutumes du pays de Normandie, while that of the local custom of Alençon is Coutumes locales de la châtellenie d'Alençon. Their short names are Normandie and Alençon, respectively. The names of customs usually refer to either a region – a province or *pays*, such as Normandie, Poitou, or Touraine – a bailliage, or a set of bailliaiges that share a common custom. For instance, the custom of Péronne, Montdidier et Roye concerns the three bailliaiges of the same name.²⁵ We provide the names of customs in both proper and capitalized forms. Written-law regions have the name Droit écrit.

Year of publication. When available in archival or secondary sources, we also include the year of publication of the custom. This date corresponds to the most recent version of a custom, usually the revised version of the sixteenth century. For customs that were not revised, we provide the year of publication of their original version. Customs in our data were published on average in 1570, with a year of publication ranging from 1455 to 1788. This information is missing for 8 customs.

Types. The types of customs indicate whether they are general or local. Of the 141 customs in the Customary Atlas, 93 are general and 48 are local.

Reference custom. For local customs, we include their general custom of reference along with its identifier.

4.3. Data files

The Customary Atlas consists of two tabular datasets and three shapefiles. Their contents are detailed in Table 1. The first tabular dataset corresponds to the matching between judicial districts and customs. The second tabular dataset corresponds to the set of customs along with their characteristics described above. These datasets are available in both Stata data format (*dta*) and text delimited format (*txt*).

The three shapefiles are those of the division of France into written-law and customary-law country along with mixed cases, the spatial distribution of general and local customs, and the spatial distribution of general customs only. All three shapefiles use an RGF93 projection based on IGN (2021) shapefile of current communes.²⁶

We distribute shapefiles and associated data files of the Customary Atlas under the CC-BY 4.0 license in the “Customs of Ancien Régime France” repository hosted on the Harvard Dataverse at <https://doi.org/10.7910/DVN/HYE209> (Gay et al., 2023b).

²² We also include Corsica, as a decree stating that “the island of Corsica is part of the French empire” was passed on November 30, 1789. In addition, small portions of the northeastern territory of France were ceded to Prussia by the Second Treaty of Paris on November 20, 1815—this includes the entire judicial district of Sarrelouis, part of the district of Bouzonville, and areas around Wissembourg and Landau. These areas are not included in our shapefiles.

²³ While we have 93 general customs in our final custom-level dataset, we create identifiers for 94 general customs because the general custom of Artois (0500), while having its territorial core outside the Kingdom of France, had seven of its local customs within it, ranging from the local custom of AIRE (0501) to that of SAINT-OMER (0507).

²⁴ Local custom identifiers range up to 20 because that is the maximum number of local customs a general custom has—namely the custom of Normandy.

²⁵ To distinguish between the two general customs of Burgundy, we use the short name Bourgogne (Comté) for the custom of the County of Burgundy and the short name Bourgogne (Duché) for the custom of the Duchy of Burgundy.

²⁶ More specifically, we use the March 2021 edition of the ADMIN-EXPRESS. These base shapefiles are relatively high resolution, so we advise users who wish to import our shapefiles directly into statistical software such as R or Stata to first simplify their geometries, e.g., by using the commands *rmapshaper* or *simply*, respectively. Note also that while data files contain name strings in both upper and lower case, the attribute tables of shapefiles (*dbf* files) contain name strings in upper case only to avoid compatibility issues with accented letters in GIS software.

Table 1
Variables in the *Customary Atlas* Data Files.

Data file	Variable	Description
BAILLIAGES_CUSTOMS (tabular)	generalite_name	Généralité name (upper case)
	bailliage	Bailliage identifier
	bailliage_name	Bailliage name (upper case)
	written_law	Indicator for written or customary law
	custom	Custom identifier
CUSTOMS (tabular)	custom_name	Custom name (short, upper case)
	custom	Custom identifier
	custom_name	Custom name (short, upper case)
	custom_name_prop	Custom name (short, proper case)
	custom_name_long	Custom name (long, upper case)
	custom_name_long_prop	Custom name (long, proper case)
	publication	Publication year
	type	Type of custom (local or general)
	reference_custom	General custom of reference identifier
	reference_custom_name	General custom of reference name (short, upper case)
WRITTEN_LAW (shapefile)	reference_custom_name_prop	General custom of reference name (short, proper case)
	WRIT_LAW	Indicator for written law (yes; no; mixed)
CUSTOMS (shapefile)	CUST_ID	Custom identifier
CUSTOMS_GNL (shapefile)	CUST_NS	Custom name (short, upper case)
	CUST_G_ID	General custom identifier
	CUST_G_NS	General custom name (short, upper case)

5. Potential uses and avenues for future research

The *Customary Atlas of Ancien Régime France* opens up fruitful avenues of research. Similar to [Murdock \(1959\)](#) atlas of ethnic homeland boundaries in Africa, our atlas can be used to study how historical customs shaped economic development over the long run—in our case, in pre-industrial European societies. The Customary Atlas will also contribute to a better understanding of the origins of the French civil law tradition.

To date, the literature examining how historical customs shape economic outcomes has focused primarily on Africa, relying on information available in [Murdock \(1967\)](#) *Ethnographic Atlas* and the related Standard Cross-Cultural Survey ([Murdock and White, 1969](#)). The Customary Atlas we propose is a first step toward a broader atlas covering pre-industrial Europe as a whole. As such, it represents a critical stepping stone to understanding how legal institutions affected historical development and whether they played a role in the European Marriage Pattern, the Industrial Revolution, or the Demographic Transition—three processes that brought dramatic economic and demographic change and that have their origins in Europe. Moreover, given Europe's history of mass emigration to the New World, such an atlas can help to better understand the roots of historical migration patterns, whether different customs positively or negatively selected individuals into migration, and how legal institutions of host countries were shaped by those of migrants' countries of origin.

In particular, the Customary Atlas can be used to study how customary rules may have shaped individual behavior or the economy over the long term. It provides the relevant spatial units of analysis for research on the implications of historical variation in women's rights ([Doepke et al., 2022](#); [Hazan et al., 2019](#)), marriage rules ([Voena, 2015](#)), inheritance ([Bertocchi, 2006](#); [Curtis et al., 2023](#)), widowhood ([Lambert and Rossi, 2016](#); [Dillon and Voena, 2018](#)), parenting ([Doepke and Zilibotti, 2017](#)), illegitimacy, or the severity of criminal punishment.²⁷ We develop below two examples of particular interest to economists of potential applications of the data presented in this article: inheritance and legal origins.

5.1. Inheritance

The long-run consequences of customary inheritance rules is a topic of growing interest among economists ([Hager and Hilbig, 2019](#); [Huning and Wahl, 2021](#); [Bartels et al., 2020](#); [Süß, 2023](#); [Fontenay et al., 2023](#); [Gay et al., 2023c](#)). The case of Ancien Régime France is particularly interesting from this perspective, as customary rules were characterized by considerable heterogeneity, as shown in [Fig. 2](#). This map shows that there were (at least) ten different ways to regulate inheritance under customary law: areas under partible inheritance could follow strict partibility or partibility with option; areas under impartible inheritance could follow primogeniture, unigeniture, or ultimogeniture; and both partible and impartible areas could include or exclude women from inheritance.

Specifically, [Gay et al. \(2023c\)](#) use this map to study the effect of inheritance customs on fertility in late eighteenth-century France. They show that in a context where land was subject to indivisibility constraints, inheritance rules that divided land equally among offspring (partible inheritance) reduced the economic incentives to have children. This resulted in a gap of 0.7 children relative to historical customary regions where all land could be transferred to a single heir (impartible inheritance)—a gap that closed

²⁷ References and links to the original texts of all customs included in the Customary Atlas are provided in Appendix Table B.2.

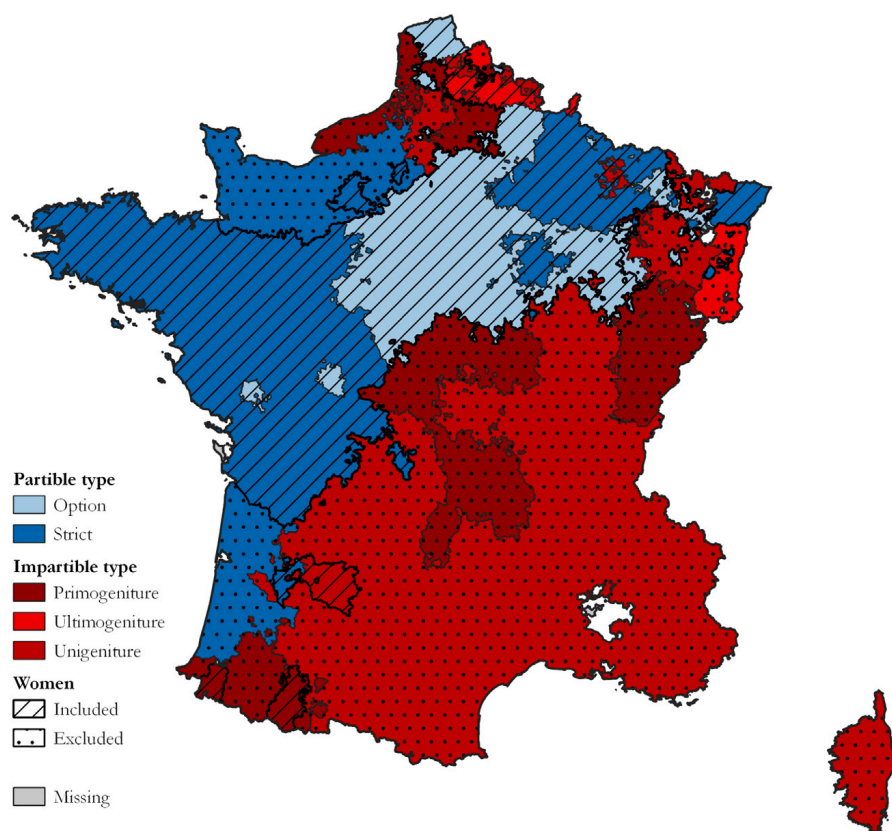


Fig. 2. Variation in Inheritance Customs in Ancien Régime France.

Notes. This figure displays the spatial distribution of inheritance customs in Ancien Régime France. Partible inheritance customs (blue) could be either strict or with option. Under strict partible inheritance, heirs had to share the inheritance equally, including all *intra-vivos* transfers. Under partible with option, heirs could opt out of the inheritance and instead keep any *intra-vivos* transfers they had received from the deceased. Impartible inheritance customs (red) could favor either the first born (primogeniture), the last born (ultimogeniture), or any of the offspring (unigeniture). Women could either be included (lines) or excluded from inheritance (dots). Shapefile based on Gay et al. (2023c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

shortly after the abolition of impartible inheritance customs during the French Revolution. In addition, using Murdock (1959) atlas of ethnic homeland boundaries in Africa along with Murdock (1967) Ethnographic Atlas, Fontenay et al. (2023) show that differences in fertility in contemporary sub-Saharan Africa across partible and impartible ancestral customs are of similar magnitude to those found in late eighteenth-century France by Gay et al. (2023c). These examples illustrate how the Customary Atlas feeds into similar research themes as Murdoch's Ethnographic Atlas, and how lessons from pre-industrial Europe can bridge past and present crucial issues, such as demographic transitions in developing countries.

5.2. Legal origins

The Customary Atlas will be a useful tool for research in law and economics on legal origins. Although customary law tradition covered most of pre-modern Europe – and is currently still a dominant source of law in sub-Saharan Africa – it has received far less attention in the literature than civil and common law traditions (La Porta et al., 2008). Indeed, the legacies of customary law for modern rules and regulations, as well as for economic outcomes, are relatively unknown. The Customary Atlas will contribute to fill this gap. In particular, a large body of work has analyzed modern commercial law and shown that French civil law – which derives from Roman law – is associated with weaker financial and economic development.²⁸ The reason is that, compared to English common law, it offers weaker protection to outside investors (Porta et al., 1998), regulates labor markets and entry more heavily (Djankov et al., 2002; Botero et al., 2004), increases procedural formalism, and reduces judicial independence (Djankov et al., 2003; Porta et al., 2004). Because the French civil law tradition was transplanted through conquest and colonization, these pervasive effects are visible throughout the world.

However, this vast literature has not yet reached a consensus on when and why the French and English legal traditions acquired their distinctive features. Two main theories have been proposed for the French case. The first theory holds that the French

²⁸ In France, Le Bris (2019) finds that regions with legal institutions historically closer to Roman (civil) law developed faster than those under customary law.

Revolution introduced extensive legal codes that deprived judges of their independence and legislative powers (Merryman and Pérez-Perdomo, 2018; Zweigert and Kötz, 1998; Klerman and Mahoney, 2007). The second theory holds that from the twelfth and thirteenth centuries, the French monarchy sought to exercise greater control over the legal system in order to unify the country in a struggle between the center and the periphery (Dawson, 1960; Berman, 1983). According to this view, Roman written law provided the backbone of such a system (Glaeser and Shleifer, 2002). By delineating the historical regions of customary law in Ancien Régime France, the Customary Atlas can shed new light on this legal struggle between the center and the periphery. Moreover, it can be used as an underlying frame of reference to precisely delineate the areas of influence of Roman written law and customary law on different legal domains. Thus, it will be useful to assess the extent to which Roman written law provided the backbone of the French legal tradition, or whether customary law also had important legacies on French law. Finally, since customs regulated many aspects of economic life in pre-industrial France, they can be codified in a manner similar to modern commercial law in order to quantitatively evaluate the influence of legal rules on economic and financial outcomes. In this respect, the Customary Atlas will be a useful tool for studying where French law first adopted its modern, regulation-oriented characteristics, as well as its economic consequences, taking advantage of within-country variation in the law.

To conclude, this article describes the construction and content of the *Customary Atlas of Ancien Régime France*. From the Middle Ages to the nineteenth century, most of Europe was characterized by a myriad of heterogeneous customary rules that regulated the lives of the population and many aspects of the economy. By providing the specific spatial boundaries of these customary regions for an entire polity, the Customary Atlas opens fruitful avenues of research for a better understanding of the roots and implications of historical legal institutions. It is also a first step in the construction of a broader European Customary Atlas.

CRedit authorship contribution statement

Victor Gay: Writing – review & editing, Writing – original draft, Funding acquisition, Data curation. **Paula E. Gobbi:** Writing – review & editing, Writing – original draft, Funding acquisition, Data curation. **Marc Goñi:** Writing – review & editing, Writing – original draft, Data curation.

Data availability

We disseminate all relevant data files under the CC-BY 4.0 license in the “Customs of Ancien Régime France” repository hosted on the Harvard Dataverse at <https://doi.org/10.7910/DVN/HYE209>.

Appendix A. Additional figures, tables, and details on the codification process of customs and their content

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.eeh.2024.101588>.

References

- Acemoglu, Daron, Johnson, Simon, Robinson, James A., 2005. Institutions as a fundamental cause of long-run growth. In: Aghion, Philippe, Durlauf, Steven N. (Eds.), *Handbook of Economic Growth*, Elsevier Academic Press, Amsterdam, pp. 385–472.
- Bartels, Charlotte, Jäger, Simon, Obergruber, Natalie, 2020. Long-term effects of equal sharing: Evidence from inheritance rules for land. NBER Working Paper 28230.
- Berman, Harold J., 1983. *Law and Revolution: The Formation of the Western Legal Tradition*. Harvard University Press, Cambridge.
- Bertocchi, Graziella, 2006. The law of primogeniture and the transition from landed aristocracy to industrial democracy. *J. Econ. Growth* 11 (1), 43–70.
- Botero, Juan C., Djankov, Simeon, Porta, Rafael La, de Silanes, Florencio López, Shleifer, Andrei, 2004. The regulation of labor. *Q. J. Econ.* 119 (4), 1339–1382.
- Brette, Armand, 1904a. *Atlas des bailliages ou juridictions assimilées ayant formé unité électorale en 1789 dressé d'après les actes de la convocation conservés aux Archives nationales*. Imprimerie nationale, Paris.
- Brette, Armand, 1904b. *Recueil de documents relatifs à la Convocation des états généraux de 1789. Tome troisième*. Imprimerie nationale, Paris.
- Brette, Armand, 1915. *Recueil de documents relatifs à la convocation des états généraux de 1789. Tome quatrième*. Imprimerie nationale, Paris.
- Chénou, Émile, 1926. *Histoire générale du droit français public et privé des origines à 1815. Volume 1. Recueil Sirey*, Paris.
- Curtis, Matthew, Gobbi, Paula E., Goñi, Marc, Haddad, Joanne, 2023. *Inheritance customs, the European marriage pattern and female empowerment*. Unpublished.
- Dawson, John P., 1960. *A History of Lay Judges*. Harvard University Press, Cambridge.
- Dillon, Brian, Voena, Alessandra, 2018. Widows' land rights and agricultural investment. *J. Dev. Econ.* 135, 449–460.
- Djankov, Simeon, La Porta, Rafael, López-de Silanes, Florencio, Shleifer, Andrei, 2002. The regulation of entry. *Q. J. Econ.* 117 (1), 1–37.
- Djankov, Simeon, La Porta, Rafael, López-de Silanes, Florencio, Shleifer, Andrei, 2003. Courts. *Q. J. Econ.* 118 (2), 453–517.
- Doepke, Matthias, Hannusch, Anne, Montenbruck, Laura, Tertilt, Michèle, 2022. The economics of women's rights. *J. Eur. Econom. Assoc.* 20 (6), 2271–2316.
- Doepke, Matthias, Zilibotti, Fabrizio, 2017. Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica* 85 (5), 1331–1371.
- Fontenay, Sébastien, Gobbi, Paula E., Goñi, Marc, 2023. Fertility in sub-Saharan Africa: The role of inheritance. ECARES Discussion Paper 2024–01.
- Fourniel, Béatrice, Vendrand-Voyer, Jacqueline, 2017. Une 'bigarrure de loix dans une même province'. In: Gomis, Stéphane (Ed.), *Atlas historique: Auvergne, Bourbonnais, Velay*. Université Clermont Auvergne, Clermont-Ferrand.
- Ganghofer, Roland, Levresse, Pierre, 1977. *Le droit romain en alsace du XIIe au XVIe siècle*. Giuffrè, Milan.
- Gay, Victor, Gobbi, Paula E., Goñi, Marc, 2023a. The atlas of local jurisdictions of Ancien Régime France. CEPR Discussion Paper 18608.
- Gay, Victor, Gobbi, Paula E., Goñi, Marc, 2023b. Customs in Ancien Régime France [database]. <http://dx.doi.org/10.7910/DVN/HYE209>, Harvard Dataverse.
- Gay, Victor, Gobbi, Paula E., Goñi, Marc, 2023c. Revolutionary transition: inheritance change and fertility decline. CEPR Discussion Paper 18607.
- Gilissen, John, 1958. *Une commune de l'agglomération bruxelloise: Uccle*. Éditions de l'Institut de sociologie, Bruxelles.
- Gilissen, John, 1962a. *La rédaction des coutumes dans le passé et dans le présent*. Éditions de l'Institut de sociologie, Bruxelles.
- Gilissen, John, 1962b. *Rapports généraux au Ve congrès international de droit comparé*. Centre inter-universitaire de droit comparé, Bruxelles.
- Gilissen, John, 1979. *Introduction historique au droit. Établissements Emile Bruylant*, Bruxelles,

- Giuliano, Paola, Nunn, Nathan, 2018. Ancestral characteristics of modern populations. *Econ. Hist. Dev. Reg.* 33 (1), 1–17.
- Glaeser, Edward L., Shleifer, Andrei, 2002. Legal origins. *Q. J. Econ.* 117 (4), 1193–1229.
- Gouron, André, Terrin, Odile, 1975. *Bibliographie des coutumes de France. Éditions antérieures à la Révolution*. Droz, Genève.
- Grinberg, Martine, 1997. La rédaction des coutumes et les droits seigneuriaux: nommer, classer, exclure. *Annales* 52 (5), 1017–1038.
- Grinberg, Martine, 2006. *Écrire les coutumes. Les droits seigneuriaux en France*. Presses universitaires de France, Paris.
- Hager, Anselm, Hilbig, Hanno, 2019. Do inheritance customs affect political and social inequality? *Am. J. Political Sci.* 63 (4), 758–773.
- Hazan, Moshe, Weiss, David, Zoabi, Hosny, 2019. Women's liberation as a financial innovation. *J. Finance* 74 (6), 2915–2956.
- Hilaire, Jean, Terré, François, 1994. *La vie du droit: coutumes et droit écrit*. Presses universitaires de France, Paris.
- Huning, Thilo R., Wahl, Fabian, 2021. The fetters of inheritance? Equal partition and regional economic development. *Eur. Econ. Rev.* 136, 103776.
- IGN, 2021. ADMIN-EXPRESS édition mars 2021 par territoire France métropolitaine [Database]. IGN, Paris, <https://geoservices.ign.fr/adminexpress>.
- Isambert, François-André, Jourdan, Athanase-Jean-Léger, Decrusy, Nicolas, 1825. *Recueil général des anciennes lois françaises depuis l'an 420 jusqu'à la révolution de 1789*. Tome XI. Belin-Leprieur, Paris.
- Joignon, Laurence, 1989. *Coutumes, familles, successions et alliances en Lorraine: 1670–1900* (Ph.D. thesis). Paris: EHESS.
- Kim, Marie Seong-Hak, 2021. Custom, Law, and Monarchy: A Legal History of Early Modern France. Oxford University Press, Oxford.
- Klerman, Daniel, Mahoney, Paul G., 2007. Legal origin? *J. Comp. Econ.* 35 (2), 278–293.
- Klimrath, Henri, 1837. *Études sur les coutumes*. Levraud, Paris.
- Klimrath, Henri, 1843. *L'histoire du droit français*. Tome second. Levraud, Paris.
- Krafft, Karl, 1930. *Anerbensitte und Anerbenrecht in Württemberg unter Besonderer Berücksichtigung cksichtigung von Württembergisch-Franken*. Kohlhammer, Stuttgart.
- La Porta, Rafael, López-de Silanes, Florencio, Shleifer, Andrei, 2008. The economic consequences of legal origins. *J. Econ. Lit.* 46 (2), 285–332.
- Laboulaye, Édouard, Dareste, Rodolphe, 1868. *Le grand coutumier de France*. Nouvelle édition. Auguste Durand and Pedone-Lauriet, Paris.
- Lambert, Sylvie, Rossi, Pauline, 2016. Sons as widowhood insurance: Evidence from Senegal. *J. Dev. Econ.* 120, 113–127.
- Le Bris, David, 2019. Testing legal origins theory within France: Customary laws versus roman code. *J. Comp. Econ.* 47 (1), 1–30.
- Lowes, Sara, 2021. In: Bisin, Alberto, Federico, Giovanni (Eds.), *The Handbook of Historical Economics*. Elsevier Academic Press, Amsterdam, pp. 147–177.
- Merryman, John, Pérez-Perdomo, Rogelio, 2018. *The Civil Law Tradition: An Introduction to the Legal Systems of Europe and Latin America*. Stanford University Press, Redwood City.
- Murdock, George Peter, 1957. World ethnographic sample. *Ethnology* 59 (4), 664–687.
- Murdock, George Peter, 1959. *Africa: Its Peoples and their Culture History*. McGraw-Hill Book Company, New York.
- Murdock, George Peter, 1967. *Ethnographic Atlas*. University of Pittsburgh Press, Pittsburgh.
- Murdock, George Peter, White, Douglas R., 1969. Standard cross-cultural sample. *Ehtnology* 8 (4), 329–369.
- Nunn, Nathan, 2009. The importance of history for economic development. *Annu. Rev. Econ.* 1 (1), 65–92.
- Nunn, Nathan, 2020. The historical roots of economic development. *Science* 367, 6485.
- Olivier-Martin, François, 1938. Le roi de France et les mauvaises coutumes au Moyen âge. *Z. der Savigny-Stiftung für Rechts Geschichte, Germanistische Abteilung* 58 (1), 108–137.
- Olivier-Martin, François, 1948. *Histoire du droit français des origines à la Révolution*. Domat Montchrestien, Paris.
- Padoa-Schioppa, Antonio (Ed.), 1997. *Legislation and Justice*. Clarendon Press; New York: Oxford University Press, Oxford.
- Porta, Rafael La, López-de-Silanes, Florencio, Pop-Eleches, Cristian, Shleifer, Andrei, 2004. Judicial checks and balances. *J. Polit. Econ.* 112 (2), 445–470.
- Porta, Rafael La, López-de-Silanes, Florencio, Shleifer, Andrei, Vishny, Robert W., 1998. Law and finance. *J. Polit. Econ.* 106 (6), 1113–1155.
- Poudret, Jean-François, 1998. *Coutumes et coutumier*. Staempfli éditions, Berne.
- Poumarède, Jacques, 1972. *Les successions dans le sud-ouest de la France au Moyen Âge*. Presses universitaires de France, Paris.
- Bourdot de Richebourg, Charles-Antoine, 1724. *Nouveau coutumier général ou corps des coutumes générales et particulières de France, et des provinces connues sous le nom des Gaules*. Michel Brunet, Paris.
- Röhm, Helmut, 1957. *Die Vererbung des Landwirtschaftlichen Grundeigentums in Baden-Württemberg*. Selbstverlag der Bundesanstalt für Landeskunde, Bonn.
- Süß, Karolin, 2023. Long-term effects of historical inheritance customs on household formation and gender disparities. *Ruhr Economic Paper* 1038.
- Voena, Alessandra, 2015. Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *Amer. Econ. Rev.* 105 (8), 2295–2332.
- Voltaire, 1878. *Coutumes*. In: *Dictionnaire philosophique*. Garnier, Paris.
- Zink, Anne, 1993. *L'héritier de la maison. Géographie coutumière du sud-ouest de la France sous l'Ancien Régime*. EHESS, Paris.
- Zweigert, Konrad, Kötz, Hein, 1998. *An Introduction to Comparative Law*. Oxford University Press; New York: Clarendon Press, Oxford.

1.3 Methodological Contributions

1.3.1 Reproducibility

This section provides the following replication article:

- Replication article: Gay, Victor. 2023a. “Culture: An Empirical Investigation of Beliefs, Work, and Fertility. A Verification and Reproduction of Fernández and Fogli (2009).” *Journal of Comments and Replications in Economics*, 2: 1–15.⁶

1.3.2 Data Curation and Data Articles

This section provides the following book chapter and data article:

- Book chapter: Gay, Victor. (Gay, 2025d). “Un data paper en SHS: pourquoi, pour qui, comment?” [A data paper in the social sciences: why, for whom, how?]. In *Publier, partager, réutiliser les données de la recherche: les data papers et leurs enjeux*. Edited by Christine Kosmopoulos and Joachim Schöpfel, 66–74. Lille, France: Presses universitaires du Septentrion.
- Data article: Gay, Victor. 2021. “Mapping the Third Republic. A Geographic Information System of France (1870–1940).” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 54(4): 189–207.⁷

⁶The appendix to this article is available at <https://doi.org/10.18718/81781.28>. The replication package associated with this article is available at <https://doi.org/10.15456/j1.2022252.0530923662>.

⁷The appendix to this article is available at <https://doi.org/10.1080/01615440.2021.1937421>. The original dataverse associated with this article is available at <https://dataverse.harvard.edu/dataverse/TRF-GIS>.

Culture: An Empirical Investigation of Beliefs, Work, and Fertility

A Verification and Reproduction of Fernández and Fogli (*American Economic Journal: Macroeconomics*, 2009)

Victor Gay*

Journal of Comments and Replications in Economics, Volume 2, 2023-2, DOI: [10.18718/81781.28](https://doi.org/10.18718/81781.28)

JEL: J13, J16, J22, J24, Z13

Keywords: Culture, Female Labor Force Participation, Fertility, Replication

Data Availability: The data and Stata code to reproduce the results of this replication can be downloaded at JCRE's data archive (DOI: [10.15456/j1.2022252.0530923662](https://doi.org/10.15456/j1.2022252.0530923662)).

Please Cite As: Gay, Victor (2023). Culture: An Empirical Investigation of Beliefs, Work, and Fertility: A Verification and Reproduction of Fernández and Fogli (2009). *Journal of Comments and Replications in Economics*, Vol 2(2023-2). DOI: [10.18718/81781.28](https://doi.org/10.18718/81781.28)

Abstract

In this article, I perform a verification and a reproduction of the main results in Fernández and Fogli (2009), which estimates the role of culture in explaining the labor and fertility decisions of second generation immigrant women to the United States in 1970. While I am able to verify Fernández and Fogli's (2009) main results as well as their robustness relative to both labor and fertility decisions, I am unable to reproduce them relative to labor decisions in alternative samples drawn from the same underlying population.

*Department of Social and Behavioral Sciences, Toulouse School of Economics and Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, France, victor.gay@tse-fr.eu.

Acknowledgements: I gratefully acknowledge funding from the ANR under grant ANR-17-EURE-0010 (*Investissements d'Avenir* program).

Received September 09, 2022; Revised September 02, 2023; Accepted September 04, 2023; Published October 16, 2023.

©Author(s) 2023. Licensed under the Creative Common License - Attribution 4.0 International (CC BY 4.0).

1 Introduction

Despite dramatic improvements over the past half-century, women's economic conditions relative to men's still exhibit profound disparities across countries, in particular regarding their labor market involvement (Olivetti and Petrongolo, 2016; Klasen, 2019). To account for such persistent gender gaps, economists have increasingly appealed to cultural explanations to complement more traditional economic ones (Fernández, 2007; Giuliano, 2021)¹. When doing so, a common empirical strategy to distinguish the role of cultural factors consists in using an "epidemiological approach" (Fernández, 2011)². This strategy identifies cultural effects by comparing behaviors among individuals with different cultural origins but who are embedded within the same institutional environment, thereby facing similar external incentives when making decisions. To proxy for culture, this approach uses past aggregate outcomes of individuals' places of origin, as only the cultural component of these variables should exhibit some explanatory power in this setting. Fernández and Fogli (2009, *American Economic Journal: Macroeconomics*, henceforth FF) was among the first studies to systematically apply this method³. Therein, the role of culture is highlighted in explaining the labor and fertility decisions of second generation immigrant women in the United States in 1970.

FF's empirical approach has been highly influential: this article is credited with 509 citations as of September 2023⁴. This represents more than twice as many citations to any article published in the same (inaugural) issue of AEJ: Macroeconomics or in inaugural issues of all three other AEJ journals. FF's impact per this citations metric also fares well compared to articles published in the concurrent issue of the *American Economic Review*, as only 3 of its 22 articles have received more citations than FF to date (Table A.1)⁵. Moreover, FF has dramatically influenced empirical methods in cultural economics: among its 509 citations, 138 studies have applied FF's empirical approach in a variety of contexts in order to provide a cultural explanation to variations in economic and demographic behaviors (Table A.2)⁶.

Despite its status of seminal study, there has been no attempt to replicate FF's results⁷. This article fills this gap by conducting a replication of FF's main result, i.e., that women from countries with traditionally higher rates of female labor force participation (FLFP) are more likely to be in the

¹Traditional economic explanations that account for gender gaps in labor force participation include, among others, economic development and the structural transformation along with changes in female education and fertility (Goldin, 1990; Gaddis and Klasen, 2014; Ngai and Petrongolo, 2017), medical progress in maternal health (Albanesi and Olivetti, 2016), declining prices of labor-saving consumer durable goods (Greenwood, Seshadri and Yorukoglu, 2005), and the increasing availability of oral contraceptives (Goldin and Katz, 2002; Bailey, 2006).

²Here, culture is understood as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation" (Guiso, Sapienza and Zingales, 2006, p. 23, cited in Giuliano, 2021).

³Earlier uses of comparable methods include Blau (1992), Carroll, Rhee and Rhee (1994), Antecol (2000), and Fortin (2005). The NBER Working Paper version of FF was published in 2005.

⁴This citation count is based on data from Clarivate Web of Science—Google Scholar credits FF with 1,696 citations as of September 2023. Further including journal articles citing FF's NBER Working paper version increases the Clarivate Web of Science citation count by an additional 34 citations.

⁵Considering annual rather than total citations, it is clear that starting from 2016 FF has been in the same category as these top-three AER articles in terms of citations (Figure B.1).

⁶Though widely used, FF's epidemiological approach is not without criticism, as selection into migration might bias the cultural effects identified through this method (Beblo, Gorges and Markowsky, 2020a; 2020b).

⁷The data and reproduction code were not made available by the authors. The webpage of the article (<https://www.aeaweb.org/articles?id=10.1257/mac.1.1.146>) and of both authors (<https://sites.google.com/site/raquelfernandezsite> and <https://sites.google.com/site/alessandrafoglisite>) were accessed in September 2023.

labor force and that those from countries with traditionally lower total fertility rates (TFR) have relatively less children. More specifically, I replicate results in FF's Table 2, which reports coefficients from estimating the following equation on a sample of second-generation immigrant women to the United States aged 30 to 40 in 1970:

$$Z_{isj} = \beta_0 + \beta_1' \mathbf{X}_i + \beta_2 \tilde{Z}_j + f_s + \varepsilon_{isj}, \quad (1)$$

where Z_{isj} is the work or fertility decision of woman i residing in the Standard Metropolitan Statistical Area (SMSA) s and of ancestry j ⁸. $\mathbf{X}_i$ includes a set of individual characteristics, f_s is a set of SMSA of residence fixed effects, and $\tilde{Z}_j$, a proxy for culture—past values of FLFP or TFR for i 's country of ancestry j . Standard errors are clustered at the country-of-ancestry level.

Throughout this replication attempt, I strictly follow Clemens's (2017, p. 327) definition of the nature of a replication test:

A *replication* test estimates parameters drawn from the same sampling distribution as those in the original study. [...] A replication test can take two forms: A *verification* test means ensuring that the exact statistical analysis reported in the original paper gives materially the same results reported in the paper, either using the original data set or remeasuring with identical methods the same traits of the same sample of subjects. [...] A *reproduction* test means resampling precisely the same population but otherwise using identical methods to the original study.

I first attempt to *verify* estimates reported in FF's Table 2 by constructing the same regression sample and estimating Equation 1 based on the guidelines provided in FF's original article. Then, I attempt to *reproduce* estimates reported in FF's Table 2 by resampling precisely the same population but otherwise using identical methods. In particular, while FF's analysis relies on the Metro sample of the US census of 1970, I use two alternative samples that are drawn from the same underlying population: the State and the Neighborhood samples of the US census of 1970. Overall, while I am able to verify estimates reported in FF's Table 2 as well as their robustness relative to both labor and fertility outcomes, I am unable to reproduce these estimates relative the labor outcome in alternative samples drawn from the same underlying population.

In the remainder of this article, I describe the construction of the dataset used in this replication exercise (Section 2) then perform the verification (Section 3) and reproduction (Section 5) tests of the estimates reported in FF's Table 2.

2 Data

In this section, I describe the procedures I implement to reconstruct the regression sample of FF's Table 2 (Section 2.1), its analysis variables (Section 2.2), and its cultural proxy variables (Section 2.3). I further describe the construction of alternative samples drawn from the same underlying population (Section 2.1). To assess the accuracy of my procedures relative to FF's original

⁸FF's Table 2 (p. 157) is reproduced from the original article in Table A.3.

dataset, I leverage summary statistics at the country and individual levels reported in FF's Tables 1 and A.1⁹.

2.1 Regression Sample

The 1970 Metro Sample of the US Census FF's Table 2 uses the 1 percent 1970 Form 2 Metro Sample of the US Census, which can be retrieved from IPUMS USA (Ruggles et al., 2021)¹⁰. It is a 1-in-100 random sample drawn from the 15 percent random sample of the population that was given Form 2 and in which the smallest identifiable geographic units are SMSAs (Bureau of the Census, 1972, p. 194–195).

Sample selection procedures The regression sample of the analysis in FF's Table 2 includes married women aged 30 to 40 residing in non-farming households, who hold non-agricultural occupations, were born in an identified US state, and whose fathers' were born in an identified country outside the United States¹¹. Respondents with fathers born in Russia, centrally planned economies, or in countries with less than 15 respondents are excluded¹².

Country-level summary statistics reported in FF's Table 1 provide the opportunity to assess the accuracy of my sample selection procedures. The original and verification samples are nearly identical: while the original regression sample contains 6,774 observations, the verification sample contains 6,768 observations (Table A.8). The 6 missing observations are from Italy (4), Germany (1), and the Philippines (1). I was unable to find the reason for these missing observations.

2.2 Analysis Variables

Outcome variables FF's Table 2 reports coefficients from estimating Equation 1 on two outcomes: the number of hours worked in the previous week and the number of children ever born to a woman. While the number of children is precisely reported in the original census data for up to 12 children, hours worked are reported in intervals. FF computes a measure of time worked by assigning the midpoint of each of these intervals (Table A.9)¹³.

A comparison with country-level summary statistics reported in FF's Table 1 reveals that both outcome variables in the original and verification samples display identical means by country of origin, except for countries for which observations are missing compared to the original sample—though differences are marginal (Table A.11). Turning to individual-level summary statistics reported in FF's Table A.1 similarly shows little differences across the original and verification samples (Table A.12).

⁹These tables are reproduced from the original article in Tables A.4 and A.5.

¹⁰The specification of the data extract from IPUMS USA is detailed in Table A.6.

¹¹Sample selection procedures are thoroughly described in FF's pages 152–154 and footnotes 18, 19, 22, and 23.

¹²The specific sample selection procedures applied to generate the regression sample of FF's Table 2 are detailed in Table A.7.

¹³The highest category for the hours worked variable is 60+ hours. For this category, I assign the value 66, as it is the maximum value reported in FF's Table A1. I also construct a measure of weeks worked last year in the same way to make comparisons of summary statistics across the original and verification samples.

Control variables Regressions in FF's Table 2 control for a set of individual characteristics: all specifications include respondents' age and age squared, some include their educational attainment, and the "full specification" further controls for their husbands' age, educational attainment, and total income. While FF applies no transformation to census data for respondents' age and their husbands' total income besides expressing it in tens of thousands of dollars, four indicator variables are constructed to measure educational attainment: below high school (omitted from regressions), high school degree, some college, and at least a college degree (Table A.10). The same transformations are applied to capture husbands' educational attainment. Moreover, FF creates ten age-range indicators to control for husbands' age. FF provides no indication regarding the size of these ranges, so I create ten equally-sized bins of 8.6 years from 14 to 100—the minimum and maximum of husbands' ages in the data.

As with outcome variables, comparing individual-level summary statistics of control variables across the original and verification samples reveals little differences (Table A.12).

2.3 Cultural Proxy Variables

To proxy for culture, FF uses past aggregate outcomes of respondents' countries of ancestry. Because the census does not provide the country of birth of respondents' mothers when both their parents were born outside the United States, FF uses their fathers' birthplace to assign country-of-ancestry culture.

Female labor force participation in 1950 To capture the cultural determinants of women's working behavior, FF uses country-of-ancestry FLFP in 1950 from the International Labour Organization (ILO). In particular, notes below FF's Table 1 specify that this variable is from "ILO, Economically Active Population, 1950–2010, (Geneva, 1997)" along with the following bibliographical reference: "International Labour Office. 1988. Current International Recommendations on Labour Statistics. Geneva: International Labour Organization."

These references are not entirely accurate. Going back to the original source, FF's data for FLFP in 1950 are from Table 4 of ILO's *Economically Active Population, 1950–2010, Vol. I, Asia* (1996, p. 39–203), *Vol. III, Latin America and the Caribbean* (1997a, p. 27–131), and *Vol. IV, Northern America - Europe - Oceania* (1997b, p. 41–211), entitled "Population and Economically Active Population by Sex and Age Group, 1950–2010." Moreover, while FF claims to be using "the rate of the economically active population for women over 10 years of age," a close comparison of the original ILO data to those in FF's Table 1 reveals that FF is actually using FLFP rates calculated relative to the total female population and not relative to the population of women over 10 years old.

Comparing both approaches reveals important differences (Table A.13). On average, FLFP rates relative to the total female population are 6 percentage points lower than those relative to the population of women over 10 years old, with differences ranging from 2 to 19 percentage points. FLFP rates relative to the population of women over 10 years old further exhibits more dispersion as its standard deviation across countries is 15 percentage points, while it is limited to 12 percentage points for FLFP rates relative to the total female population.

In the verification and reproduction tests, I consider the FLFP rate relative to the population of women over 10 years old as the appropriate measure since it is the measure claimed to be used by FF. I will however show how this inaccuracy affects the results when using FF's original FLFP values as per FF's Table 1.

Total fertility rate in 1950 To capture the cultural determinants of women's fertility, FF uses country-of-ancestry TFR in 1950 from the United Nations (UN). In particular, notes below FF's Table 1 specify that this variable is from the "United Nations *Demographic Yearbook* 1997, Historical supplement table 4" with no further indication in the bibliography.

The appropriate reference is UN's *Demographic Yearbook* 1997, Historical Supplement (1999), Table 4 (entitled "Selected Derived Measures of Natality: 1948–1997"), column *Total Fertility Rate*. However, it is unclear which years FF selected, as TFR data for 1950 is not available for all countries of ancestry present in the regression sample—the closest year for which TFR data is available across all countries is 1953¹⁴.

As a result, the 1950 TFR values (or the closest year to 1950) only corresponds to the values in FF's Table 1 for 6 out of 25 countries. Because the 1953 TFR values are available for all countries in the regression sample, I use these values in the verification attempt. This also ensures that the measure is defined for the same year. Reassuringly, absolute differences between the 1953 TFR values and FF's original values remain moderate, as their means are 3.48 for the former and 3.66 for the latter, with a country-wise average absolute difference of 0.20 (Table A.14). The difference is nonetheless substantial for the Philippines, since FF reports a value of 7.29 while the original value corresponds to 3.14. I will show how these differences affect the results when using FF's original TFR values as per FF's Table 1.

2.4 Alternative Samples from the 1970 US Census

Two alternative 1-in-100 samples were drawn from the same 15 percent sample of the population as the 1 percent 1970 Form 2 Metro Sample used in FF's Table 2: the 1 percent 1970 Form 2 State Sample and the 1 percent 1970 Form 2 Neighborhood Sample. The selection process of these samples was such that they are mutually exclusive and are representative of the same underlying population (Bureau of the Census, 1972, p. 197–198). They can therefore be used independently as well as combined for a reproduction test.

Sample selection and variables transformation procedures applied to these samples are identical to those applied to the Metro sample. The inspection of country and individual-level summary statistics reveals that characteristics of observations across these regression samples are nearly identical (Table A.15). Comparing the distributions of the main variables of the analysis along their CDFs directly similarly confirms that they are not different (Figures B.2 and B.3)¹⁵.

¹⁴It also remains unclear how the values for Germany are computed, since it was then split between the FRG and the GDR. I take the average TFR value of both countries in 1955, since this is the first year for which TFR data is available for the GDR.

¹⁵More specifically, I compare the distribution of the treatment (FLFP in 1950 and TFR in 1953) and outcome (hours worked and number of children) variables of interest pair-wise across the Metro, State, and Neighborhood samples based on Goldman and Kaplan's (2018) procedure and implemented through Kaplan's (2019) `distcomp` Stata command. These

3 Verification

Baseline results In this section, I verify the results reported in FF's Table 2 by estimating Equation 1 on the sample of second-generation immigrant women to the United States described in Section 2. The baseline verification test uses as cultural proxy variables the FLFP rates of women over 10 years old in 1950 and the TFR in 1953. I report the original FF estimates of interest $\hat{\beta}_2$ in Panel A of Table 1 along with the verification estimates in Panel B. In Panel C, I report estimates when using the cultural proxy values reported in FF's Table 1¹⁶.

Verification estimates are relatively close to those reported in FF's Table 2. In FF's preferred specification (Columns 2 and 5), the verification coefficient on FLFP is 0.059 (s.e. of 0.014) relative to an original coefficient of 0.072 (s.e. of 0.015). The verification coefficient on TFR is 0.228 (s.e. of 0.040) relative to an original coefficient of 0.219 (s.e. of 0.041). Verification coefficients that use FF cultural proxy variables in Panel C are nearly identical to those in FF's Table 2. This suggests that the discrepancies between the verification and original FF estimates are entirely driven by cultural proxy variables. Overall, the verification of FF's Table 2 is successful.

Robustness FF claims that these results are robust to changes in sample criteria and alternative variables as cultural proxies, though the article does not provide statistical output to support this assertion. In particular, FF claims that these results are robust to including all women independently of their marital status, including Russia or independently excluding China, Mexico, and Italy, and changing the sample to women that aged 40–50. FF further claims that these results are robust to using the following alternative cultural proxy variables: the percentage of the workforce that is female in 1960, the labor force participation of women aged 30–34 in 1950, and 1960 FLFP and TFR values.

To assess the robustness of results in FF's Table 2, I run FF's preferred specification for both outcomes of interest using the above alternative sample selection criteria and cultural proxy variables. Results are reported in Tables 2 and 3. Verification estimates are robust to all alternative cultural proxies and sample selection criteria, except when restricting the sample to women aged 40–50 for the hours worked outcome, and when excluding Mexico for the children outcome. Nevertheless, verification estimates can be considered as generally robust.

4 Reproduction

In this section, I reproduce the results reported in FF's Table 2 across alternative samples drawn from the same underlying population: the State and the Neighborhood samples of the 1970 US census. Because the Metro, State, and Neighborhood samples are mutually exclusive, I further combine them to create a Pooled sample of the 1970 US census, representing a 3-in-100 sample.

These samples differ along one dimension: while the smallest identifiable geographic units in the Metro sample are SMSAs, those in the State sample are states and those in the Neighborhood

tests all fail to reject the null that these distributions are different at the 1, 5, or 10 percent levels.

¹⁶Tables A.16 and A.17 reproduce the entire FF's Table 2 corresponding to Panels B and C of Table 1, respectively.

Table 1: Verification of FF Table 2.

Dependent variable	Hours worked			Children		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Original FF estimates						
Female	0.047***	0.072***	0.053***			–0.010
LFP 1950	(0.012)	(0.015)	(0.016)			(0.008)
TFR 1950			–0.225**	0.250***	0.219***	0.194***
			(0.103)	(0.056)	(0.041)	(0.051)
Controls	No	Yes	Yes	No	Yes	Yes
Observations	6, 774	6, 774	6, 774	6, 774	6, 774	6, 774
B. Verification estimates (verification cultural proxies)						
Female	0.040***	0.059***	0.043***			–0.009
LFP 1950	(0.010)	(0.014)	(0.013)			(0.006)
TFR 1953			–0.291**	0.266***	0.228***	0.205***
			(0.105)	(0.051)	(0.040)	(0.047)
Controls	No	Yes	Yes	No	Yes	Yes
Observations	6, 768	6, 768	6, 768	6, 768	6, 768	6, 768
C. Verification estimates (FF cultural proxies)						
Female	0.047***	0.072***	0.052***			–0.010
LFP 1950	(0.012)	(0.015)	(0.015)			(0.008)
TFR 1950			–0.238**	0.250***	0.215***	0.190***
			(0.105)	(0.056)	(0.041)	(0.050)
Controls	No	Yes	Yes	No	Yes	Yes
Observations	6, 768	6, 768	6, 768	6, 768	6, 768	6, 768

Notes: This table reproduces the estimates of interest of FF Table 2. Panel A reports the original FF estimates, Panel B, verification estimates when using verification cultural proxies, and Panel C, verification estimates when using FF cultural proxies according to FF Table 1. All specifications include respondents' age and age squared, their husbands' age-range indicators, and SMSA fixed effects. *Controls* include education indicators for both respondents and their husbands. Robust standard errors in parentheses account for clustering at country level.

*** Significant at the 1 percent level. ** Significant at the 5 percent level.

sample are census regions. Still, the Metro sample contains state information and all three samples contain census-region information. Therefore, the reproduction proceeds by estimating Equation 1

Table 2: Robustness of Verification Estimates to Alternative Sample Restrictions.

Sample	Baseline	All marital statuses	Include Russia	Exclude China	Exclude Mexico	Exclude Italy	Aged 40–50
A. Dependent variable is hours worked							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female LFP 1950	0.059*** (0.014)	0.046*** (0.012)	0.057*** (0.010)	0.068*** (0.013)	0.046*** (0.015)	0.057*** (0.016)	0.028 (0.029)
Observations	6,768	8,280	7,559	6,715	5,929	4,863	10,732
B. Dependent variable is children							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female TFR 1953	0.228*** (0.040)	0.210*** (0.036)	0.236*** (0.038)	0.236*** (0.040)	0.129 (0.091)	0.195*** (0.033)	0.291*** (0.070)
Observations	6,768	8,280	7,559	6,715	5,929	4,863	10,732

Notes: This table reproduces estimates from the full specification of FF Table 2 across alternative sample restrictions: the baseline sample in Column (1), including women of all marital statuses (together with marital status fixed effects but without husband controls) in Column (2), including Russia in Column (3), excluding China in Column (4), Mexico in Column (5), and Italy in Column (6), and on the sample of women aged 40–50 in Column (7). All specifications include respondents' age and age squared, their husbands' age-range indicators, SMSA fixed effects, and education indicators for both respondents and their husbands. Robust standard errors in parentheses account for clustering at country level.

*** Significant at the 1 percent level.

with alternative residence fixed effects, depending on their availability in the sample. A proper comparison of estimates obtained on all three samples can hence only be operated across specifications that use census-region fixed effects.

I report results in Table 4. First, reproduction estimates on the FF (Metro) sample using census-region instead of SMSA fixed effects renders coefficients stronger for the hours worked outcome and does not affect much those for the children outcome (Column 1). This suggests that a specification using census-region fixed effects can be reasonably used as a benchmark. When using the same specification with census-region fixed effects on the State and the Neighborhood samples (Columns 4 and 6), estimates hold for the children outcome—although they are slightly weaker—but are insignificant and closer to zero for the hours worked outcome. Pooling all three samples generates reproduction estimates that constitute a weighted average of sample-specific estimates (Column 7). Results are similar when the original FF cultural proxy variables are used instead (Table A.18).

Table 3: Robustness of Verification Estimates to Alternative Cultural Proxy Variables.

Proxy	Baseline	FF proxies	FLFP or TFR 1960	Share female 1960	FLFP 1950 30–34
A. Dependent variable is hours worked					
	(1)	(2)	(3)	(4)	(5)
Female LFP	0.059*** (0.014)	0.072*** (0.015)	0.064*** (0.015)	0.088*** (0.017)	0.045*** (0.012)
Observations	6,768	6,768	6,768	6,768	6,768
B. Dependent variable is children					
	(1)	(2)	(3)		
Female TFR	0.228*** (0.040)	0.215*** (0.041)	0.262*** (0.045)		
Observations	6,768	6,768	6,768		

Notes: This table reproduces estimates from the full specification of FF Table 2 using alternative cultural proxies: the baseline verification proxies in Column (1), the proxies from FF Table 1 in Column (2), proxies evaluated in 1960 in Column (3), the percentage of the workforce that is female in 1960 in Column (4), and the labor force participation of women aged 30–34 in 1950 in Column (5). All specifications include respondents' age and age squared, their husbands' age-range indicators, SMSA fixed effects, and education indicators for both respondents and their husbands. Robust standard errors in parentheses account for clustering at country level.

*** Significant at the 1 percent level.

To rationalize this unsuccessful reproduction on the hours worked outcome, I first inspect whether it can be explained by a different composition of the effective sample relative to respondents' countries of ancestry. In particular, I compute for each sample the share of the total residual variance by country of ancestry (Table A.19). I find little differences across samples: in all three samples, respondents from Mexico, China, Japan, and Germany are the bigger contributors to building the estimate of interest (Aronow and Samii, 2016). Then, I construct residual variance plots of residual hours worked against residual FLFP in 1950 across all samples based on the specification of Columns 1, 4, 6, and 7 of Table 4 (Figure B.4). Again, I find no clear outlier across samples. These analyses suggest that the underlying composition of the different reproduction samples cannot explain the discrepancy found in the resulting estimates.

Next, to explore whether this failure to reproduce FF's estimates for the hours worked outcome is due to "unprobable" draws of the State and Neighborhood samples, I combine the three census samples and draw 1,000 different random samples that each represent 1-in-100 samples from the

Table 4: Reproduction of FF Table 2 Across Census Samples.

1970 1% Form 2 Sample	A. Dependent variable is hours worked						
	Metro			State		Neighb.	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female LFP 1950	0.076*** (0.020)	0.052*** (0.015)	0.059*** (0.014)	0.031 (0.026)	0.019 (0.018)	0.027 (0.025)	0.045** (0.019)
Residence FE Observations	Region 6, 768	State 6, 768	SMSA 6, 768	Region 6, 694	State 6, 694	Region 6, 804	Region 20, 266
1970 1% Form 2 Sample	B. Dependent variable is children						
	Metro			State		Neighb.	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
TFR 1953	0.217*** (0.040)	0.218*** (0.040)	0.228*** (0.040)	0.158*** (0.043)	0.174*** (0.043)	0.165*** (0.035)	0.181*** (0.038)
Residence FE Observations	Region 6, 768	State 6, 768	SMSA 6, 768	Region 6, 694	State 6, 694	Region 6, 804	Region 20, 266

Notes: This table reproduces estimates from the full specifications of FF Table 2 across census extracts: the 1970 1% Form 2 Metro sample in Columns (1)–(3), the 1970 1% Form 2 State sample in Columns (4)–(5), the 1970 1% Form 2 Neighborhood sample in Column (6), and the combination of all three extracts in Column (7). All specifications include respondents' age and age squared, their husbands' age-range indicators, and education indicators for both respondents and their husbands. Robust standard errors in parentheses account for clustering at country level.

*** Significant at the 1 percent level. ** Significant at the 5 percent level.

1970 US census (6,700–800 observations). In a bootstrapping approach, I then run Equation 1 with census-region instead of SMSA fixed effects on each of these random samples for the hours worked outcome. I plot the resulting coefficients on the FLFP variable in Figure 1 and report summary statistics in Table A.20. They suggest that the original FF Metro sample is rather unusual compared to the State and Neighborhood samples. Indeed, estimates based on the Metro sample are in the 95th percentile of the distribution of estimates and only 53 percent of estimates are significant the 10 percent level. Results are similar when using the original FF proxy variable instead (Figure B.5)¹⁷.

¹⁷The same conclusion applies when using the children outcome, although the magnitude of resulting coefficients is large enough that they are all different from zero at conventional significance levels (Figure B.6).

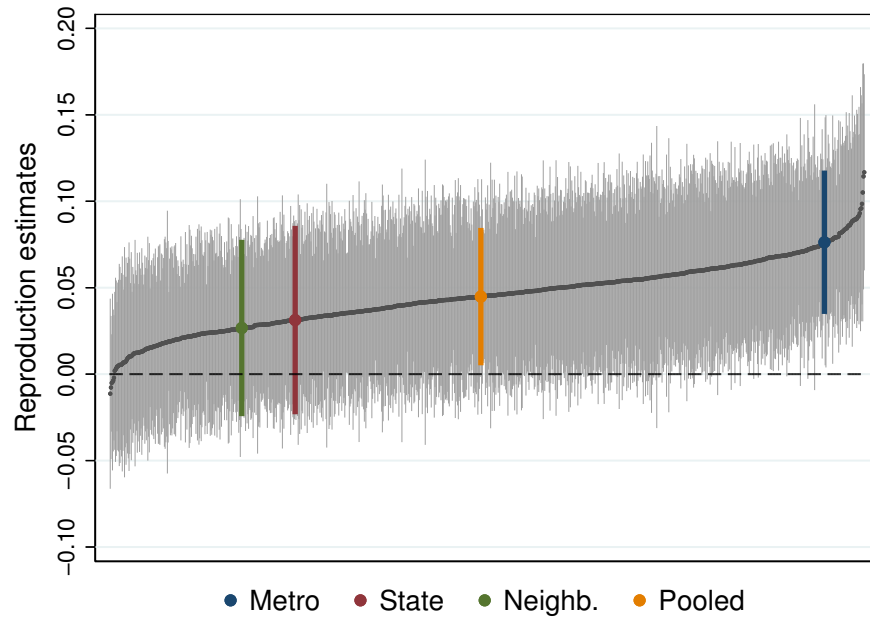


Figure 1: Reproduction Estimates of the FLFP Variable on 1,000 Random Samples. This figure plots coefficients on the FLFP verification variable from estimating Equation 1 on the hours worked outcome with census-region instead of SMSA fixed along with 95 percent confidence intervals on 1,000 different random samples representing 1-in-100 samples from the 1970 US census, from a sample pooling the Metro, State, and Neighborhood samples of the 1970 US census. It also highlights coefficients obtained when using the original Metro, State, Neighborhood, or pooled samples.

5 Conclusion

In this article, I perform a replication Fernández and Fogli's (2009) main results. While I am able to verify Fernández and Fogli's (2009) estimates and their robustness, results relative to the hours worked outcome cannot be reproduced in alternative samples drawn from the same underlying population. Extensions to other samples and meta-analytic approaches are therefore advisable to assess the validity and generalizability of the findings in this seminal study.

References

- Albanesi, S., & Olivetti, C. (2016).** “Gender Roles and Medical Progress.” *Journal of Political Economy*, 124(3), 650–695. DOI: [10.1086/686035](https://doi.org/10.1086/686035).
- Antecol, H. (2000).** “An Examination of Cross-Country Differences in the Gender Gap in Labor Force Participation Rates.” *Labour Economics*, 7(4), 409–426. DOI: [10.1016/S0927-5371\(00\)00007-5](https://doi.org/10.1016/S0927-5371(00)00007-5).
- Aronow, P. M., & Samii, C. (2016).** “Does Regression Produce Representative Estimates of Causal Effects?” *American Journal of Political Science*, 60(1), 250–267. DOI: [10.1111/ajps.12185](https://doi.org/10.1111/ajps.12185).
- Bailey, M. J. (2006).** “More Power to the Pill: The Impact of Contraceptive Freedom on Women’s Life Cycle Labor Supply.” *The Quarterly Journal of Economics*, 121(1), 289–320. DOI: [10.1093/qje/121.1.289](https://doi.org/10.1093/qje/121.1.289).
- Beblo, M., Görges, L., & Markowsky, E. (2020a).** “Gender Equality as a Confounder in the Epidemiological Approach.” *Economics Bulletin*, 40(2), 1292–1299. PDF: <http://www.access.econ.com/Pubs/EB/2020/Volume40/EB-20-V40-I2-P110.pdf>.
- Beblo, M., Görges, L., & Markowsky, E. (2020b).** “Gender Matters in Language and Economic Behaviour: Can we Measure a Causal Cognition Effect of Speaking?” *Labour Economics*, 65, 101850. DOI: [10.1016/j.labeco.2020.101850](https://doi.org/10.1016/j.labeco.2020.101850).
- Blau, F. D. (1992).** “The Fertility of Immigrant Women: Evidence from High-Fertility Source Countries.” In: George J. Borjas and Richard B. Freeman (Eds.), *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, 93–134. University of Chicago Press. URL: <http://www.nber.org/chapters/c6907..>
- Bureau of the Census. U.S. Department of Commerce (1972).** “Public Use Samples of Basic Records from the 1970 Census: Description and Technical Documentation.” Washington, DC: US Government Printing Office. PDF: https://usa.ipums.org/usa/resources/codebooks/1970_PUMS_codebook.pdf.
- Carroll, C. D., Rhee, B.-K., & Rhee, C. (1994).** “Are There Cultural Effects on Saving? Some Cross-Sectional Evidence.” *The Quarterly Journal of Economics*, 109(3), 685–699. DOI: [10.2307/2118418](https://doi.org/10.2307/2118418).
- Clemens, M. A. (2017).** “The Meaning of Failed Replications: A Review and Proposal.” *Journal of Economic Surveys*, 31(1), 326–342. DOI: [10.1111/joes.12139](https://doi.org/10.1111/joes.12139).
- Fernández, R. (2007).** “Women, Work, and Culture.” *Journal of the European Economic Association*, 5(2–3), 305–332. DOI: [10.1162/jeea.2007.5.2-3.305](https://doi.org/10.1162/jeea.2007.5.2-3.305).
- Fernández, R. (2011).** “Does Culture Matter?” In: Jess Benhabib, Alberto Bisin and Matthew O. Jackson (Eds.), *Handbook of Social Economics*, 481–510. North-Holland. DOI: [10.1016/B978-](https://doi.org/10.1016/B978-0-444-53861-1_0011)

0-444-53187-2.00011-5.

Fernández, R., & Fogli, A. (2005). “Culture: An Empirical Investigation of Beliefs, Work, and Fertility.” National Bureau of Economic Research Working Paper, 11268. URL: <https://www.nber.org/papers/w11268>.

Fernández, R., & Fogli, A. (2009). “Culture: An Empirical Investigation of Beliefs, Work, and Fertility.” *American Economic Journal: Macroeconomics*, 1(1), 146–177. DOI: [10.1257/mac.1.1.146](https://doi.org/10.1257/mac.1.1.146).

Fortin, N. M. (2005). “Gender Role Attitudes and the Labour-market Outcomes of Women across OECD Countries.” *Oxford Review of Economic Policy*, 21(3), 416–438. DOI: [10.1093/oxrep/gr1024](https://doi.org/10.1093/oxrep/gr1024).

Gaddis, I., & Klasen, S. (2014). “Economic Development, Structural Change, and Women’s Labor Force Participation: A Reexamination of the Feminization U Hypothesis.” *Journal of Population Economics*, 27, 639–681. DOI: [10.1007/s00148-013-0488-2](https://doi.org/10.1007/s00148-013-0488-2).

Giuliano, P. (2021). “Gender and Culture.” *Oxford Review of Economic Policy*, 36(4), 944–961. DOI: [10.1093/oxrep/graa044](https://doi.org/10.1093/oxrep/graa044).

Goldin, C. (1990). “Understanding the Gender Gap: An Economic History of American Women.” Oxford University Press.

Goldin, C., & Katz, L. F. (2002). “The Power of the Pill: Oral Contraceptives and Women’s Career and Marriage Decisions.” *Journal of Political Economy*, 110(4), 730–770. DOI: [10.1086/340778](https://doi.org/10.1086/340778).

Goldman, M., & Kaplan, D. M. (2018). “Comparing Distributions by Multiple Testing Across Quantiles or CDF Values.” *Journal of Econometrics*, 206(1), 143–166. DOI: [10.1016/j.jeconom.2018.04.003](https://doi.org/10.1016/j.jeconom.2018.04.003).

Greenwood, J., Seshadri, A., & Yorukoglu, M. (2005). “Engines of Liberation.” *The Review of Economic Studies*, 72(1), 109–133. DOI: [10.1111/0034-6527.00326](https://doi.org/10.1111/0034-6527.00326).

Guiso, L., Sapienza, P., & Zingales, L. (2006). “Does Culture Affect Economic Outcomes?” *Journal of Economic Perspectives*, 20(2), 23–48. DOI: [10.1257/jep.20.2.23](https://doi.org/10.1257/jep.20.2.23).

International Labour Office, Bureau of Statistics (1996). “Economically Active Population, 1950–2010, Vol. I, Asia.” STAT working paper, International Labour Organization. PDF: https://www.ilo.org/public/libdoc/ilo/1996/96B09_462_e_f_s.pdf.

International Labour Office, Bureau of Statistics (1997a). “Economically Active Population, 1950–2010, Vol. III, Latin America and the Caribbean.” STAT working paper, International Labour Organization. PDF: https://www.ilo.org/public/libdoc/ilo/1997/97B09_299_e_f_s.pdf.

International Labour Office, Bureau of Statistics (1997b). “Economically Active Population, 1950–2010, Vol. IV, Northern America - Europe - Oceania.” STAT working paper, International Labour Organization. PDF: https://www.ilo.org/public/libdoc/ilo/1997/97B09_300_e_f_s.pdf.

Kaplan, D. M. (2019). “distcomp: Comparing Distributions.” *The Stata Journal*, 19(4), 832–848. DOI: [10.1177/1536867X19893626](https://doi.org/10.1177/1536867X19893626).

Klasen, S. (2019). “What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries?” *The World Bank Research Observer*, 34(2), 161–197. DOI: [10.1093/wbro/lkz005](https://doi.org/10.1093/wbro/lkz005).

Ngai, R. L., & Petrongolo, B. (2017). “Gender Gaps and the Rise of the Service Economy.” *American Economic Journal: Macroeconomics*, 9(4), 1–44. DOI: [10.1257/mac.20150253](https://doi.org/10.1257/mac.20150253).

Olivetti, C., & Petrongolo, B. (2016). “The Evolution of Gender Gaps in Industrialized Countries.” *Annual Review of Economics*, 8(1), 405–434. DOI: [10.1146/annurev-economics-080614-115329](https://doi.org/10.1146/annurev-economics-080614-115329).

Ruggles, S., Flood, S., Foster, S., Goeken, R., Pacas, J., Schouweiler, M., & Sobek, M. (2021). “IPUMS USA: VERSION 11.0” [dataset]. Minneapolis, MN: IPUMS. DOI: [10.18128/D010.V11.0](https://doi.org/10.18128/D010.V11.0).

United Nations, Department of Economic and Social Affairs (1999). “1997 Demographic Yearbook, Historical Supplement.” United Nations. URL: <https://unstats.un.org/unsd/demographic/products/dyb/dybhists.htm>.

Un *data paper* en SHS : pourquoi, pour qui, comment ?

Victor Gay

Introduction

1. Destinée aux chercheurs en sciences humaines et sociales souhaitant se lancer dans l'écriture d'un *data paper*, ce chapitre propose un retour d'expérience prenant appui sur la production récente d'un *data paper* (Gay, 2021) et aborde les enjeux auxquels un auteur de *data paper* est souvent confronté. Il se penche tout d'abord sur les raisons méthodologiques pour lesquelles l'écriture d'un *data paper* peut être utile en sciences humaines et sociales. Il aborde ensuite le public auquel s'adresse ce nouveau format éditorial. Enfin, il propose quelques éléments pratiques afin d'aider les chercheurs à écrire leur *data paper*. Tout au long de ce chapitre, je prendrai appui sur mon expérience issue de la rédaction récente d'un *data paper* : « Mapping the Third Republic. A Geographic Information System of France (1870-1940) » (Gay, 2021). Ce *data paper* décrit un système d'information géographique de la France de la Troisième République – la base TRF-GIS¹. Cette base de données met à disposition nomenclatures et *shapefiles* annuels correspondant aux circonscriptions

administratives de France métropolitaine de 1870 à 1940. Elle décrit les circonscriptions administratives générales (départements, arrondissements, cantons) ainsi que les circonscriptions militaires, judiciaires, pénitentiaires, électorales, académiques, ecclésiastiques et les inspections du travail. Elle met aussi à disposition des nomenclatures annuelles établissant une correspondance entre chaque commune contemporaine et les circonscriptions auxquelles elle appartenait².

Un *data paper* en SHS : pourquoi ?

2. Pourquoi écrire un *data paper* ? Alors que les sciences humaines et sociales (SHS) connaissent un tournant quantitatif depuis une dizaine d'années, la valeur scientifique de la production de données reste peu reconnue : les comités d'évaluations privilégient encore l'article de recherche traditionnel tandis que les pairs réutilisent volontiers les données créées par d'autres sans pour autant leur accorder de citation. Ce manque de reconnaissance semble pourtant incompatible avec le travail chronophage que demandent la documentation du processus de production des données ainsi que leur mise en conformité avec les principes FAIR – deux éléments nécessaires à la reproductibilité des travaux de recherche. Dans ce contexte, le *data paper* constitue

1. TRF-GIS signifie *Third Republic France Geographic Information System*.

2. Le *data paper* est librement disponible sur HAL à l'adresse suivante : <https://hal.archives-ouvertes.fr/hal-02951461>. Les données sont accessibles sur le Harvard Dataverse à l'adresse suivante : <https://dataverse.harvard.edu/dataverse/TRF-GIS>.

un outil qui peut permettre aux producteurs de données de faire reconnaître leur contribution scientifique en rendant leurs données facilement citables, mais aussi en améliorant la pertinence ainsi que le périmètre de la réutilisation de leurs données.

3. L'usage des données a pris une place prépondérante dans la recherche en sciences humaines et sociales au cours des dix dernières années, en partie grâce à la production sans précédent de statistiques portant sur les faits sociaux et leur disponibilité via les catalogues de données en ligne tels que Progodo-Adisp³. C'est par exemple le cas en sociologie, une discipline au cœur des SHS : une analyse des 400 articles parus entre 2000 et 2020 dans la *Revue française de sociologie* révèle une nette tendance vers le quantitatif, si bien que depuis une décennie, plus de la moitié des articles dans cette revue contient au moins une table ou un graphique présentant des statistiques (figure 1). C'est aussi le cas en histoire, comme en témoigne le récent numéro des *Annales. Histoire, sciences sociales* consacré à l'histoire quantitative (Karila-Cohen *et al.*, 2018), ainsi que le succès de l'ouvrage *Méthodes quantitatives pour l'historien* (Lemerrier et Zalc, 2008), réédité en version anglaise il y a peu (Lemerrier et Zalc, 2020). De même, ce tournant quantitatif concerne les milieux anglo-saxons, aussi bien en histoire sociale qu'en histoire économique, où elle est plus ancienne : alors que la proportion d'articles avec au moins une

table ou un graphique présentant des statistiques restait stable autour de 90 % dans les principales revues en histoire économique entre 2005 et 2020 (Cioni *et al.*, 2021, 24), celle-ci est passée de 5 à 13 % dans l'*American History Review* sur la même période (Ruggles, 2021, 14)⁴.

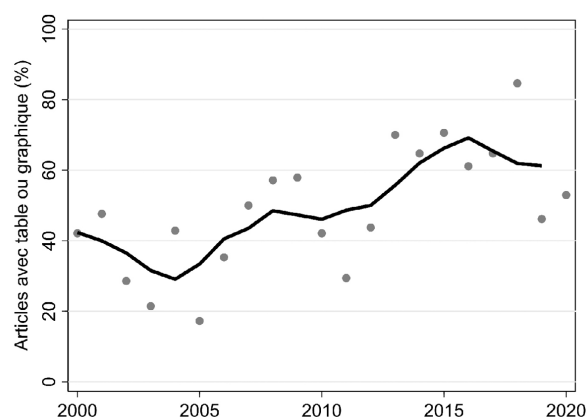


Figure 1. Pourcentage d'articles parus dans la *Revue française de sociologie* contenant au moins une table ou un graphique présentant des statistiques (2000-2020).

La courbe représente une moyenne mobile sur cinq ans. Les articles ont été récoltés par l'auteur via les portails Cairn.info pour les volumes 44 à 62 (2004-2021) et Persée pour les volumes 39 à 43 (1998-2002).

Les éditoriaux, notes, critiques, débats, articles traduits, commentaires et les articles *in memoriam* n'ont pas été retenus.

3. Le catalogue PROGEDO-ADISP diffuse enquêtes et bases de données issues de la statistique publique française. Il est accessible à l'adresse suivante : <http://www.progedo-adisp.fr>.

4. Les journaux considérés par Cioni *et al.* (2021) sont l'*Economic History Review*, le *Journal of Economic History*, *Explorations in Economic History*, la *European Review of Economic History* et *Cliometrica*.

4. C'est dans ce contexte de renouveau quantitatif que la crise de la reproductibilité, partant de la psychologie et de la médecine, a rattrapé les sciences sociales, en commençant par l'économie (Maniadis, Tufano et List, 2017) : à cause des divers biais affectant le processus de publication – manipulation des valeurs p, faible puissance statistique, biais de confirmation des auteurs et relecteurs, etc. – une majorité des résultats de recherche publiés constitueraient en réalité des « faux positifs », autrement dit, une illusion statistique (Ioannidis, 2005). Plusieurs réponses ont été apportées pour pallier ce problème : les plans de gestion de données, les plans de pré-analyse, ou encore les méta-analyses (Maniadis et Tufano, 2017 ; Christensen et Miguel, 2018). Il semble cependant que la condition *sine qua non* de sortie de crise est la reproductibilité en tant que telle, c'est-à-dire la capacité à reproduire les résultats des travaux de recherche publiés. Tel n'est pas encore le cas : par exemple, Chang et Li (2022) montrent que seule la moitié d'un échantillon de 67 articles de macro-économie publiés dans des revues de renom sont reproductibles. La première étape qui découle de ce programme consiste donc en la mise en place des conditions pour la réutilisation (adéquate) des données de la recherche.

5. Comment y parvenir ? C'est ici qu'interviennent les principes FAIR, selon lesquels la reproductibilité requiert en premier lieu des données trouvables, accessibles, interopérables et réutilisables (Wilkinson *et al.*, 2016). Le respect de ces principes est depuis quelques années déjà au cœur de la politique nationale de la recherche française au travers du plan national pour la science ouverte ainsi

que celle du financement de la recherche par l'ANR (CoSO 2019). Cependant, quel que soit le support des établissements et des infrastructures de recherche, le poids de la documentation et de la mise en conformité des données aux principes FAIR revient *in fine* en grande partie aux producteurs de données eux-mêmes, les chercheurs – qui sont par ailleurs déjà débordés par des tâches administratives qui ne cessent de s'accumuler (Ali et Rouch, 2013). Ce rôle est pourtant en contradiction avec les incitations auxquelles ils font face. En effet, la production de données n'est généralement pas valorisée par les comités d'évaluation, qui se fient encore bien plus aux publications traditionnelles dans des revues à comité de lecture (Gozlan, 2016). La production de données n'est pas plus valorisée par les pairs, qui en général ne citent pas les données qu'ils réutilisent. Par exemple, Robinson-García *et al.* (2015) montrent que seuls 18 % des données réutilisées par les articles en sciences sociales publiés entre mai et juin 2013 et disponibles dans *Web of Science* étaient citées explicitement dans leur bibliographie⁵.

6. Face à ces défis, les *data papers* offrent des perspectives intéressantes. En effet, en décrivant le processus de production des données dans un article publié dans une revue, ils offrent aux réutilisateurs de données un moyen simple pour citer les données, permettant par-là la reconnaissance scientifique du travail des producteurs

5. Une étude plus récente portant sur les pratiques de citation des données en biodiversité en 2019 constate que sur un échantillon aléatoire de 100 articles réutilisant de telles données, seuls 27 % les mentionnent explicitement dans leur bibliographie (Khan *et al.*, 2021).

de données par les comités d'évaluation, mais aussi par les pairs – au-delà, bien entendu, d'une utilisation à la fois plus adéquate grâce à la documentation et plus accessible grâce à la conformité aux principes FAIR⁶. De plus, dans la mesure où les *data papers* passent au crible du comité de lecture, ils génèrent des incitations pour les producteurs de données eux-mêmes à parfaire leurs données ainsi que leur description, ce qui *in fine* ne peut que favoriser leur diffusion et le périmètre de leur réutilisation (Walters, 2020). Dans ce sens, le *data paper* peut contribuer à résoudre le problème classique du passager clandestin qui caractérise la production de biens publics purs – ici, la donnée documentée et FAIR.

7. Le *data paper* « Mapping the Third Republic » (Gay 2021) répond à ces problématiques dans un contexte de recherche particulier, celui de la Troisième République (1870-1940). En effet, cette période historique est caractérisée par une production sans précédent de statistiques (Desrosières, 2010 [1993]) par des administrations opérant à des niveaux d'agrégation hétérogènes, dans un contexte de profonds changements socio-économiques tels que la Seconde Révolution industrielle (1870-1914) ou la Première Guerre mondiale (1914-1918). Ainsi, grâce aux nouveaux moyens de numérisation des archives statistiques, la recherche en histoire quantitative sur cette période est en forte expansion. L'analyse et la visualisation de données historiques géolocalisées requièrent cependant un cadre de référence ou un système

d'information géographique (SIG) commun. Celui-ci n'existant pas à ce jour pour la France de la Troisième République, chaque chercheur doit s'atteler individuellement à cette tâche. Cela implique une perte de temps considérable, des approximations à cause de la difficulté de l'entreprise, mais aussi un manque d'interopérabilité avec d'autres programmes de recherche et donc *in fine* un manque de reproductibilité. En effet, ces systèmes d'information sont rarement mis à disposition du public et plus rarement encore conformes aux principes FAIR⁷. La base de données TRF-GIS offre donc une solution à ces problèmes en proposant des données FAIR dont le processus de construction est documenté avec précision dans un *data paper*.

Un *data paper* en SHS : pour qui ?

8. Pour qui écrire un *data paper* ? Les utilisateurs potentiels des données décrites, bien sûr. Cependant, le lectorat d'un *data paper* est souvent bien plus large et interdisciplinaire qu'initialement envisagé par le producteur de données. En effet, les revues acceptant ce genre d'article restent rares en sciences humaines et sociales. À ma connaissance, il existe en 2021 deux revues dédiées aux *data papers* – des *data journals* – dans ce domaine : le *Journal of Open Humanities Data* et le *Research Data*
7. Une exception toutefois : les *shapefiles* de cantons pour 1884 et 1925 publiés par le LARHRA (2011) et construits selon la méthode de vectorisation manuelle de cartes historiques géoréférencées. Voir Gay (2021, 14-15) pour une analyse critique de cette méthode.

6. Même si les effets des *data papers* sur la réutilisation des données n'ont pas encore atteint leur plein potentiel, en tout cas en sciences dures (Jiao et Darch, 2020).

Journal for the Humanities and Social Sciences – bien que certains *data journals* publient en sciences sociales mais aussi dans d'autres disciplines, comme *Data in Brief* ou *F1000 Research*⁸. De plus, quelques revues en sciences humaines et sociales acceptent des *data papers* en plus d'articles de recherche traditionnels, comme *Cybergeo* pour la géographie, *Historical Methods: A Journal of Quantitative and Interdisciplinary History* pour l'histoire ou *Frontiers in Sociology* pour la sociologie. Enfin, un certain nombre de revues offrent la possibilité de publier des *data papers* dans leur section pour articles courts sans qu'ils explicitent la distinction de ce genre d'article avec un article de recherche traditionnel. Pour le domaine de l'histoire, c'est par exemple le cas pour *Explorations in Economic History* ou encore *Histoire & mesure*.

9. L'auteur de *data paper* doit donc composer avec une offre réduite qui implique mécaniquement un lectorat plus large que son propre champ de recherche. Une telle portée demande cependant une adaptation de l'écriture à une audience non spécialiste: un travail d'exposition du contexte qui évite le jargon disciplinaire, une explicitation des présupposés ayant présidé à la collecte de données et aux catégorisations, et une clarification des usages potentiels des données au-delà de son propre champ de recherche. Ces considérations ont aussi des implications sur la forme des données: on peut bien entendu mettre à disposition ses données dans le format

le plus courant dans sa discipline (spss, Stata ou sass) mais il est important de donner la possibilité à d'autres champs disciplinaires de s'en saisir en proposant aussi un format ouvert et universel comme le txt ou le csv. Le choix de l'entrepôt de données doit lui aussi être réfléchi: on peut ainsi se concentrer sur un entrepôt français comme Nakala ou Progedo-Adisp, ou bien viser un public plus international en utilisant le Harvard Dataverse, Figshare ou Zenodo – avec le risque de perdre en visibilité dans le paysage hexagonal, leur interface restant en anglais⁹.

10. Enfin, il convient de clairement délimiter les contours du *data paper* en tant que descripteur de données et non comme procédant de l'analyse ou de la généalogie de projet, surtout si celui-ci est publié dans une revue n'ayant pas de section dédiée aux *data papers*. En effet, ce format n'est pas encore ancré dans les usages et nombre d'articles dont le but principal est de décrire une base de données comportent aussi des analyses des données elles-mêmes, ce qui contribue à brouiller les pistes¹⁰.
11. Le *data paper* « Mapping the Third Republic » (Gay, 2021) propose nomenclatures et *shapefiles* annuels pour différentes administrations de la Troisième République (1870-1940). Dans la mesure où cette base de données

9. Il existe peu d'entrepôts spécifiquement dédiés aux sciences humaines et sociales, tels que Didomena (<https://didomena.ehess.fr/>), l'entrepôt de données de l'EHESS dont l'usage est limité à ses membres, ou encore Nakala (<https://nakala.fr/>), l'entrepôt développé par la TGIR Huma-Num.

10. Par exemple, à la fin d'un article décrivant une nouvelle base de données sur les esclaves émancipés de la ville du Cap en 1834, Ekama *et al.* (2021, 9) analysent les déterminants des prix des esclaves par une régression linéaire.

8. Une liste des revues publiant des *data papers* est proposée par Laurence Dedieu (Cيراد) à l'adresse suivante: <https://doi.org/10.18167/coopist/0057>.

permet de réaliser aussi bien de la cartographie de données historiques géolocalisées que d'appareiller des données issues de ces administrations pour réaliser des analyses statistiques, elle s'adresse à l'ensemble des sciences sociales avec une composante historique, et en premier lieu l'histoire économique et la démographie historique. C'est pourquoi mon choix s'est porté sur le journal *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, une revue relativement reconnue, avec une diffusion internationale large, et dont le lectorat reste interdisciplinaire en sciences sociales avec une forte composante historique et une dominante quantitative. *Explorations in Economic History* aurait été une autre candidate, mais son lectorat reste trop restreint à l'économie historique anglo-saxonne. *A contrario*, le lectorat de *Journal of Open Humanities Data* ou de *Research Data Journal for the Humanities and Social Sciences* est relativement moins tourné vers la cartographie ou l'analyse économétrique.

12. En termes de format, la base de données *TRF-GIS* est disponible dans le format dominant du champ de l'histoire économique: le *Stata data format* (dta). Mais pour permettre à d'autres disciplines de se saisir des données, j'ai aussi rendu celles-ci disponibles en format txt, facilement importable dans n'importe quel logiciel statistique ou cartographique. Enfin, je me suis porté sur l'entrepôt Harvard Dataverse. Cet entrepôt propose un espace de stockage gratuit conséquent pour les chercheurs individuels, des facilités de dépôts, une gestion automatique des métadonnées, ainsi qu'une

bonne ergonomie d'utilisation pour le téléchargement et la navigation. C'est aussi un entrepôt avec une très forte diffusion à l'international. Afin de ne pas manquer de toucher les communautés francophones, j'ai cependant pris soin de déposer les métadonnées de la base de données sur la plateforme data.gouv.fr¹¹.

13. Enfin, pour ce qui est de la forme du *data paper*, le passage du comité de lecture s'est passé sans encombre. Cependant, la remarque suivante de la part d'un relecteur anonyme montre bien le chemin qu'il reste à parcourir quant à ce type d'article: « La question est de déterminer si ce travail constitue un article. Je ne pense pas que ce soit le cas. Ainsi, je ne pense pas que la revue puisse accepter ce texte qui ne pose pas une problématique historique claire, mais qui n'est qu'une sorte de ligne directrice et la description d'un produit fini¹² ».

Un data paper en SHS : comment ?

14. Comment écrire un *data paper*, pour qu'il ne soit pas qu'un simple codebook, mais constitue une véritable clé d'accès pour la compréhension et la réutilisation des données décrites? Une tendance actuelle est la

11. Les métadonnées sur data.gouv.fr sont disponibles à l'adresse suivante: <https://www.data.gouv.fr/fr/datasets/systeme-dinformation-geographique-de-la-france-de-la-troisieme-republique-1870-1940>.

12. « The question is to determine if this work constitutes an article. I do not think so. Thus, I do not think that the review could accept this text which does not ask a clear historical problem, but is only a kind of guideline and a description of an end product. »

génération automatique de *data papers* à partir des métadonnées de la base de données (Schöpfel *et al.*, 2019). Bien que ce type de format reste bien moins chronophage qu'un *data paper* standard et permette une diffusion rapide via le moissonnage des catalogues de données, il ne semble pas permettre de réutilisation adéquate à cause du manque de description des données dont il fait preuve – cette méthode est donc à proscrire. La description des méthodes de construction et des choix de typologie doit être au cœur du *data paper*, chose que seul un humain est capable de réaliser.

15. Dans ce contexte, la première difficulté à laquelle est confronté un auteur de *data paper* lors de la rédaction est celle de la structuration du texte, car celle-ci ne correspond pas tout à fait à celle d'un article traditionnel. Un excellent modèle à suivre est celui proposé par le *data journal* en sciences dures *Scientific Data* du groupe *Nature*¹³. Bien que ce modèle s'adresse aux sciences dures, sa structure peut tout à fait s'adapter aux sciences humaines et sociales. La structure proposée commence par une première section (« Contexte et résumé ») qui décrit succinctement les données produites, leur contexte scientifique ainsi que leurs réutilisations potentielles. Une seconde section (« Méthodes ») décrit avec précision toutes les procédures utilisées dans le processus de

production des données afin que celui-ci soit reproductible. Ensuite, une section « Fichiers de données » décrit chaque jeu de données associé avec le *data paper*. Cela comprend variables, noms de fichiers, localisation, ainsi que formats et poids numérique. Une quatrième section (« Validation technique ») présente les analyses ou procédures ayant permis de confirmer la validité des données décrites (en sciences humaines et sociales, il peut s'agir d'une confrontation avec différentes sources ou avec des données comparables). Une cinquième section (« Notes d'usage ») permet à l'auteur de décrire plus en détail les procédures de réutilisation des données ainsi que de développer quelques exemples. Enfin, une dernière section (« Disponibilité du code ») permet le cas échéant de décrire les procédures d'accès au code de reproduction de la base de données.

16. Une seconde difficulté à laquelle est confronté l'auteur de *data paper* concerne la mise à disposition des données. En effet, contrairement à l'article de recherche classique où les fichiers de données et de reproduction ne sont possiblement requis que lors de la phase de publication proprement dite, ceux-ci sont ici requis lors de la soumission de l'article à la revue avec des procédures d'accès qui doivent être explicitées dans le texte. Dans la mesure où c'est aussi l'intégrité des données qui est évaluée par les relecteurs, ceux-ci doivent impérativement y avoir accès dès l'étape de relecture. Les données doivent donc déjà avoir un identifiant provisoire à ce stade, ce qui permet de rattacher les données au *data paper* et de créer un écosystème « données-*data paper* ». Heureusement,

13. Les instructions de soumission à *Scientific Data* sont disponibles à l'adresse suivante : <https://www.nature.com/sdata/publish/submission-guidelines>. D'autres modèles existent, comme celui proposé par *Data in Brief*, disponible à l'adresse suivante : <https://www.elsevier.com/journals/data-in-brief/2352-3409/guide-for-authors>. Plus généralement, Kim (2020) propose un tour d'horizon des formats de *data papers* proposés par un certain nombre de *data journals*.

la plupart des entrepôts de données (par exemple le Harvard Dataverse ou Figshare) offrent la possibilité aux producteurs de données de générer un lien privé accessible (anonymement) par l'intermédiaire d'un code qu'il faudra fournir au comité de lecture pour l'évaluation par les pairs. Une fois toutes les modifications effectuées, les données seront accessibles au public via un identifiant pérenne, par exemple un DOI.

17. La rédaction du *data paper* « Mapping the Third Republic » (Gay, 2021) suit rigoureusement le modèle proposé par *Scientific Data* décrit plus haut. En effet, après une introduction présentant le contexte scientifique dans lequel la base de données s'insère, le corps de l'article consiste en une section « Méthode » qui explique en détail non seulement la méthodologie technique de construction ainsi que ses limites, mais aussi les éléments institutionnels qui sous-tendent les variations temporelles de chaque administration au cours de la Troisième République. J'y aborde par exemple les changements territoriaux impliqués par la perte et le retour de l'Alsace-Lorraine, la réforme des arrondissements de 1926, les réformes militaires de 1873-1874, ou encore les différentes lois de redécoupage électoral (les éléments précis sont étayés par des tableaux situés dans un appendice électronique d'une centaine de pages). S'ensuivent une description des quinze jeux de données de la base *TRF-GIS* (variables, espace de stockage, formats, licence), puis une exposition de la validation technique de l'ensemble, sous deux formes : une confrontation des sources secondaires utilisées avec un ensemble de 175 sources primaires

(individuellement listées en appendice et disponibles en PDF dans l'entrepôt de données) et une validation de la méthode de construction des *shapefiles* par comparaison avec des données similaires (le *shapefile* des cantons de 1884 du LARHRA). L'article se termine par la description de l'emplacement du code et des données, ainsi qu'un exemple de réutilisation (la cartographie de l'abstention lors des élections législatives de 1914 au niveau des circonscriptions électorales).

Conclusion

18. La production d'un *data paper* demande un certain apprentissage dans la mesure où il s'agit d'un format nouveau en sciences humaines et sociales. Il est cependant en pleine expansion, et les chercheurs désireux de se lancer dans l'écriture d'un *data paper* ont à leur disposition de nombreux retours d'expérience à travers séminaires, journées spéciales, ou colloques dédiés (souvent mis en ligne, comme celui-ci). J'ai par exemple pu présenter le *data paper* décrivant la base *TRF-GIS* dans le cadre de la semaine DATA SHS 2020 organisée par la Plateforme universitaire de données de Toulouse (PUD-T) en décembre 2020, au webinar dédié aux *data papers* organisé par le Groupe de travail inter-réseaux « Atelier Données » du CNRS en février 2021, au séminaire « Des sources aux SIG : des outils pour la cartographie dans les humanités numériques » en mai 2021, ou encore au

séminaire « Histoire et numérique » du Centre d'histoire de Science Po (CHSP) en mai 2021¹⁴.

19. De nombreuses ressources sont aussi mises à disposition par différents acteurs institutionnels comme DoRANum et CoopIST (Cirad), qui proposent des dossiers sur le sujet, les URFIST ou l'INRAE, qui proposent des formations, ou encore les Plateformes Universitaires de Données (PUD) de PROGEDO, qui proposent des journées sur ce thème dans les MSHS lors de la semaine annuelle DATA SHS¹⁵.

Pour consulter les données mobilisées dans le chapitre, voir :

Harvard Dataverse : <https://dataverse.harvard.edu/dataverse/TRF-GIS>

PROGEDO Adisp : http://www.progedo-adisp.fr/enquetes_donhist.php

14. La vidéo de ma présentation dans le cadre du séminaire « Des sources aux SIG » est disponible à l'adresse suivante : <https://youtu.be/mBAIRdWR41k>, de la minute 2 à 45.

15. Les ressources de DoRANum sur les *data papers* sont disponibles à l'adresse suivante : <https://doranum.fr/data-paper-data-journal>. Celles de CoopIST du Cirad sont disponibles à l'adresse suivante : <https://doi.org/10.18167/coopist/0057>. Un exemple de support de cours proposé par l'INRAE est disponible à l'adresse suivante : <https://dx.doi.org/10.15454/1.478247389988942E12>. Voir aussi Reymonet (2017).



Mapping the Third Republic: A Geographic Information System of France (1870–1940)

Victor Gay 

Toulouse School of Economics (TSE) and Institute for Advanced Study in Toulouse (IAST), University of Toulouse Capitole, Toulouse, France

ABSTRACT

This article describes a comprehensive geographic information system of Third Republic France: the TRF-GIS. It provides annual nomenclatures and shapefiles of administrative constituencies of metropolitan France from 1870 to 1940, encompassing general administrative constituencies (départements, arrondissements, cantons) as well as the most significant special administrative constituencies: military, judicial and penitentiary, electoral, academic, labor inspection, and ecclesiastical constituencies. It further proposes annual nomenclatures at the contemporaneous commune level that map each municipality into its corresponding administrative framework along with its population count. The 901 nomenclatures, 830 shapefiles, and complete reproduction material of the TRF-GIS are available at <https://data-verse.harvard.edu/dataverse/TRF-GIS>.

KEYWORDS

Historical GIS; administrative boundaries; nomenclatures; toponyms; France; Third Republic

JEL CODES

N01; N43; N44; N93; N94; Y91

1. Introduction

For the seven decades between the collapse of the Second Empire in 1870 and the establishment of the Vichy Regime in 1940, France remained under a single political regime: the Third Republic. Despite the unprecedented stability of its institutions, France underwent dramatic socio-economic changes during this period. These were the result of several critical junctures, among which the Second Industrial Revolution and the modernization of France (1870–1914), the establishment of free, mandatory, and secular education (1881–1882), the Separation of the Churches and the State (1905), the First World War (1914–1918), the economic and political turmoils of the 1930s, and the defeat by Nazi Germany (1939–1940). Research programs aimed at understanding how these experiences shaped the French society increasingly rely on empirical evidence, in part due to the conjunction of two phenomena: the unprecedented production of administrative statistics under the Third Republic (Desrosières 1998 [1993]) and the recent revival of quantitative history (Ruggles 2021). Ongoing advances in optical character recognition (OCR) techniques applied to digitized archival

documents (e.g., Ostafin et al. 2020) suggest that this trend will likely intensify in the future.

But a fundamental element empirical researchers typically need in order to process, map, and analyze spatially localized historical data is an underlying geographic frame of reference—a geographic information system (GIS). And contrary to the United States (<https://nhgis.org>; Fitch and Ruggles 2003; Manson et al. 2020) and Great Britain (<https://visionofbritain.org.uk>; Southall 2011, 2012, 2014), a historical GIS of administrative constituencies of France has not been produced thus far.¹ This lack is especially problematic for quantitative research in the context of Third Republic France, as statistics then were oftentimes produced by administrations operating at heterogeneous and incompatible levels of aggregation. As a result, it currently falls onto individual researchers to complete this time-consuming, challenging, and yet, crucial task. Although département-level historical shapefiles are routinely produced for the needs of single studies, these are generally insufficiently documented and unavailable to the community. Furthermore, nomenclatures and shapefiles for more fine-grained administrative constituencies remain rare, and, when they exist, are hardly curated: they are

CONTACT Victor Gay  victor.gay@tse-fr.eu  Toulouse School of Economics (TSE) and Institute for Advanced Study in Toulouse (IAST), University of Toulouse Capitole, Toulouse, France.

 Supplemental materials for this article are available at <https://doi.org/10.1080/01615440.2021.1937421>

© 2021 Taylor & Francis Group, LLC

generally neither findable due to lacking metadata, nor accessible under appropriate licensing, nor interoperable with other GISs, nor reusable in machine-readable format (Ruggles 2018).² Overall, research endeavors in the context of Third Republic France are bound to be hindered by the lack of a shared frame of reference that would not only ease researcher's efforts by generating economies of scale, but also improve reproducibility and interoperability across scientific studies (Wilkinson et al. 2016; Christensen and Miguel 2018).

To alleviate these issues and empower research programs in this context, this article proposes a comprehensive geographic information system of Third Republic France: the TRF-GIS. This historical GIS provides annual nomenclatures along with shapefiles of administrative constituencies of metropolitan France (mainland France and Corsica) from 1870 to 1940—901 nomenclatures and 830 shapefiles in total. It encompasses general administrative constituencies (départements, arrondissements, cantons) as well as the most significant special administrative constituencies: military constituencies (military regions and subdivisions), judicial constituencies (courts of appeal and of first instance), penitentiary constituencies, electoral constituencies (circonscriptions), academic constituencies (academies), labor inspection constituencies, and ecclesiastical constituencies (dioceses). The TRF-GIS further proposes annual nomenclatures at the contemporaneous commune level that map each municipality into its corresponding administrative framework along with its population count.

The construction of TRF-GIS shapefiles uses a methodology that differs from the manual vectorization of georeferenced historical maps (as in, e.g., LARHRA 2011; Perret, Gribaudo, and Barthelemy 2015; Ostafin et al. 2020): it first reconstructs the administrative framework contemporaneous communes were embedded in, then aggregates these units into relevant constituencies using current communes shapefiles. Despite some limitations, this method offers substantial advantages: it provides more precise results than existing outputs, helps circumvent the scarcity of annual historical maps, and alleviates the substantial time and resource investments associated with manual vectorization methods. By providing a comprehensively-curated common frame of reference that encompasses all aspects of the French society, the TRF-GIS will help create the conditions for a better understanding of the dramatic socio-economic changes that France underwent during the seven decades that lasted the Third Republic.

The remainder of this article is organized as follows: Sec. 2 describes the construction methodology of the TRF-GIS, Sec. 3 details the data records available in the database, Sec. 4 provides elements of technical validation, and Sec. 5 contains supportive elements for end users, including a walk-through for map production.

2. Methods

This section describes the general construction methodology of the TRF-GIS (Sec. 2.1), its limitations (Sec. 2.2), and provides historical institutional details for each constituency described by the TRF-GIS (Sec. 2.3).

2.1. General methodology

The underlying structure of the TRF-GIS builds on three datasets that are available under open licensing: INSEE's *Code Officiel Géographique* (COG), the *Histoire Administrative des Communes* (HAC) database, and IGN's GEOFLA 2011 shapefile.³

2.1.0.1. INSEE's Code Officiel Géographique (COG).

INSEE's annual COGs provide the official nomenclature of communes in a given year (Lang 2016). They uniquely identify existing communes through a five-digit coding scheme. The first two digits of this coding scheme correspond to a commune's département while its last three digits correspond to a commune's identifier within the département, generally assigned in ascending alphabetical order. For instance, the commune of Allonne in the département of Oise (60) holds INSEE code 60009. This nomenclature was first established in 1943 and has remained stable ever since, barring changes in communes' territorial structures. To reflect these changes, the COG is updated annually. The construction of the TRF-GIS uses the COGs 2005 and 2011.⁴

2.1.0.2. The Histoire Administrative des Communes (HAC) database.

The HAC database provides, for each of the 41,410 communes that ever existed between 1801 and 2005, a unique record that contains the commune's dates of creation and subsequent modifications, the general administrative constituencies it ever belonged to, its INSEE codification, the evolution of its toponymy, and its population counts across censuses. These records are distributed through <http://cassini.ehess.fr> (Motte and Vouloir 2007). Constructed by the LaDéHiS and INED (Motte, Ségué, and Théré

2003), information in the HAC database relies on official administrative acts published in the *Bulletin des Lois* from 1801 to 1931 and in the *Journal Officiel* hereafter.⁵ Except for a few cases documented in the TRF-GIS source code, communes in the HAC database are effectively characterized in 2005 geography. Communes that no longer existed by 2005 are assigned the INSEE code of their absorbing commune. For instance, absorbed by Beauvais in 1943, the commune of Voisinlieu holds Beauvais' INSEE code 60057. Former and current communes can nevertheless be uniquely identified through the url query parameter of their individual record on <http://cassini.ehess.fr>, which I denominate their "Cassini code." It is assigned in ascending alphabetical order from 0 to 41,475 (with a few gaps). For instance, while Beauvais holds Cassini code 3332, Voisinlieu holds Cassini code 40965.⁶

2.1.0.3. IGN's GEOFLA® Communes Édition 2011 France Métropolitaine (GEOFLA) shapefile. IGN's GEOFLA 2011 shapefile provides the official representation of the 36,568 communes of metropolitan France in 2011 geography in polygon and vector forms.⁷ It is expressed through an RGF93 Lambert-93 projection system and derived from the geometry of IGN's BD CARTO with a precision of 1:1,000,000.⁸

2.1.1. TRF-GIS shapefiles

Annual TRF-GIS shapefiles use the GEOFLA 2011 as their underlying frame, to the exclusion of the territories that did not belong to France during certain years (Appendix Figure A.2): Alsace-Lorraine (1871–1918) and the communes of La Brigue and Tende (1870–1940). As a result, shapefiles that constitute the base frame of TRF-GIS shapefiles for 1870 and 1919–1940 represent 36,566 of the 36,568 existing communes of 2011, and 34,932 for 1871–1918. Each of these communes is then matched to the general administrative constituencies they belonged to in the relevant year using information from the HAC database, which initial 2005 geography is converted into 2011 geography using the COGs 2005 and 2011—between these two dates, nine communes disappeared and seven were created.

While nearly all communes of 2011 existed in 1801, some were not created until after 1870. For these 737 communes, I assign the constituencies their parent commune belonged to prior to their creation. For instance, the commune of Voisinlieu was created from Allonne in 1930, a parent commune that belonged to the département of Oise between 1870

and 1929. Hence, Voisinlieu is assigned to the département of Oise in the base frames for the TRF-GIS shapefiles 1870–1929.

Commune-level polygons are then dissolved to create general administrative constituency-level polygons. The construction of special administrative constituency shapefiles relies on this initial characterization, as these constituencies were generally based on the general administrative framework constituted by départements, arrondissements, and cantons.

2.1.2. TRF-GIS nomenclatures

Annual TRF-GIS constituency-level nomenclatures use contemporaneous toponyms and codifications as provided by the HAC database for general administrative constituencies. For special administrative constituency nomenclatures, I rely directly on historical sources. Annual commune-level nomenclatures are composed only of those communes that existed in the relevant year. For instance, the commune of Voisinlieu only appears in the commune-level nomenclatures 1930–1940. Appendix Figure A.3 displays the evolution of the number of communes that compose annual commune-level nomenclatures. It ranges from 35,974 communes in 1871 to 38,038 in 1939–1940.

2.1.3. Municipal arrondissements of Paris, Lyon, and Marseille

The three building blocks of the TRF-GIS have different treatments of the 45 municipal arrondissements of Paris, Lyon, and Marseille: while the GEOFLA treats them as independent elements, explaining why it contains 36,610 distinct entities representing 36,568 communes, COGs and the HAC database do not. Because they have been stable since 1859, the TRF-GIS treats the 20 municipal arrondissements of Paris as distinct entities. However, it treats the communes of Lyon and Marseille as single entities because their municipal arrondissements were modified several times since 1870, making it challenging to match these to the GEOFLA.

2.1.4. Timeline

The TRF-GIS encompasses the period of the Third Republic. This political regime was proclaimed following the fall of the Second Empire on September 4, 1870, and later dissolved on July 10, 1940, after the capitulation against Nazi Germany and the advent of the Vichy Regime. Although the Third Republic began in 1870, the general administrative constituencies that prevailed during this regime were only effective by May 1871 and the signature of the Frankfurt treaty,

which settled the aftermath of the Franco-Prussian War and the annexation of Alsace-Lorraine by the newly unified German Empire. For interoperability with research programs covering the Second Empire, the TRF-GIS begins in 1870 and therefore encompasses constituencies prevailing by the end of that regime.

2.2. Limitations

The TRF-GIS exhibits two limitations. First, its shapefiles rely on delineations of communes in 2011 geography through the GEOFLA 2011. I use this dataset for quality and reproducibility purposes, as it is the earliest geography for which an IGN-produced commune-level shapefile is accessible under open licensing. Moreover, annual shapefiles of commune

boundaries for 1870–1940 have not been produced thus far.⁹ This needs not invalidate the TRF-GIS, for two reasons. First, commune boundaries did not change much since the beginning of the Third Republic: out of the 36,568 communes represented in the GEOFLA 2011, less than seven percent ever underwent territorial modifications (creation, suppression, territorial transfer) between 1870 and 2011 (Figure 1). Second, TRF-GIS shapefiles represent constituencies that are for the most part defined at higher levels of aggregation than communes, thereby limiting potential inaccuracies at their boundaries. Still, this might represent an issue for the larger urban centers as some increased in size over time, absorbing peripheral parcels and hamlets. For instance, the area of the Bois de Vincenne was not absorbed by Paris until 1929. TRF-GIS shapefiles are therefore better suited for regional and national extents.¹⁰

The second limitation of TRF-GIS shapefiles is due to some constituencies being at times defined across commune boundaries, generally in the vicinity of urban centers. For instance, the area around the city of Montluçon was divided between two cantons, Montluçon-Ouest and Montluçon-Est, with the commune of Montluçon being subdivided between these two cantons along the Cher river (Figure 2a). While this poses no issue for the construction of constituency-level nomenclatures, shapefiles and commune-level nomenclatures have communes as their base unit. I circumvent this problem by implementing INSEE's and IGN's common methodology when building their COGs and GEOFLAs and use the concept of pseudo-constituency: whenever part of a constituency is defined across commune boundaries, a pseudo-constituency is created for that specific part, which level of aggregation is the commune itself and

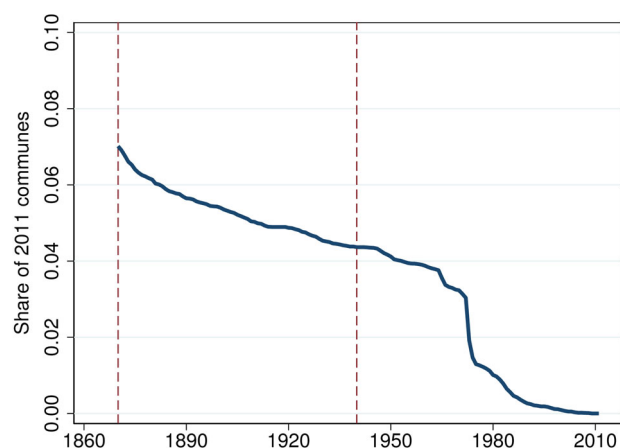


Figure 1. Share of communes of 2011 that underwent territorial modifications between 1870 and 2011. Red vertical dashed lines indicate the proclamation (1870) and dissolution (1940) of the Third Republic.

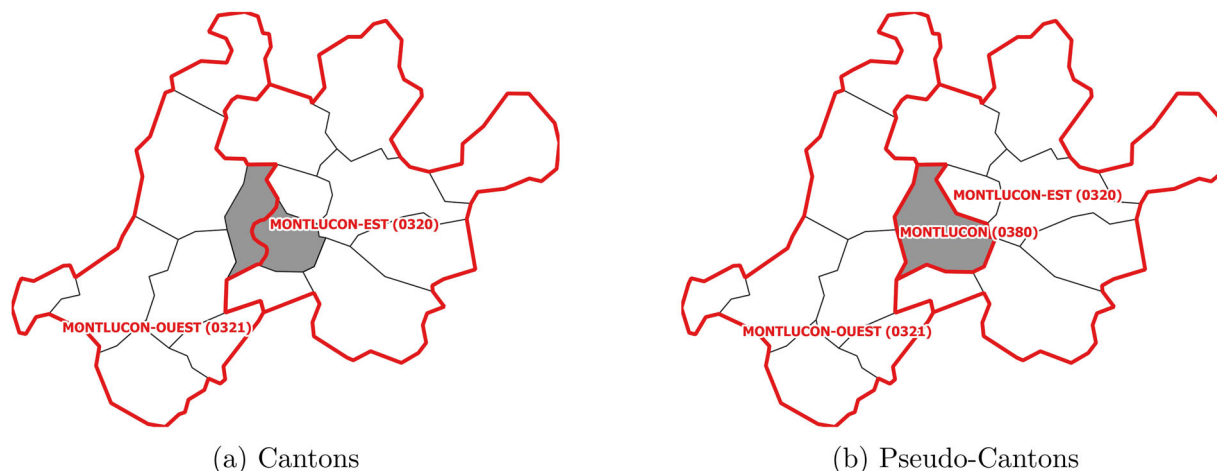


Figure 2. Cantons and pseudo-cantons of the Montluçon area (1870–1940). Red edges indicate cantons (a) and pseudo-cantons (b) boundaries. Black edges indicate commune boundaries. The shaded area corresponds to the commune of Montluçon.

which encompasses all constituencies defined within that commune. In the case of Montluçon's cantons, TRF-GIS shapefiles and commune-level nomenclatures assign communes that fall entirely into cantons to either cantons of Montluçon-Ouest (seven communes) and Montluçon-Est (eight communes), while the commune of Montluçon itself, being subdivided between these two cantons, is assigned to the pseudo-canton of Montluçon (Figure 2b).

2.3. Institutional details

2.3.1. General administrative constituencies

General administrative constituencies (départements, arrondissements, cantons) formed the basic territorial divisions through which the State deployed its administration. During the Third Republic, these constituencies were strictly nested: cantons belonged to a single arrondissement, and arrondissements, to a single département. In each département, a prefect represented the State and supervised its administration within the département from the prefecture. Similarly, a subprefect held office in each arrondissement at the subprefecture. While départements and arrondissements were both administrative units and territorial constituencies, cantons were territorial divisions with no administrative prerogative—except for local justices of the peace—that formed the basis of military and electoral constituencies. This organization had been in place since the National Constituent Assembly in 1789–1791 (Ozouf-Marignier 1989) and was later consolidated by the law of August 10, 1871, which stabilized this framework and made départements the central constituency for France's territorial administration (Gros 2014, 307–35).¹¹ It was only essentially altered by two events: the annexation of Alsace-Lorraine in 1871 and its recovery in 1919, and the arrondissements reform of 1926 (Masson 1984, 359–99).

2.3.1.1. The annexation of Alsace-Lorraine in 1871 and its recovery in 1919. Following the Franco-Prussian War, the Frankfurt treaty of May 10, 1871 (complemented by the convention of October 12, 1871) stipulated the transfer of Alsace-Lorraine to the German Empire. Annexed territories consisted of the département of Bas-Rhin, most of the département of Haut-Rhin, parts of the départements of Moselle and Meurthe, and a few communes of the département of Vosges—1,694 communes in total.¹² But boundaries of these territories did not overlap with pre-war limits of general administrative constituencies. As a result, it

was necessary to reconfigure these constituencies within the territories of annexed départements that had remained in France: the remaining territories of the département of Haut-Rhin were grouped into the Territoire-de-Belfort and the département of Meurthe was renamed Meurthe-et-Moselle, absorbing the remaining territories of the département of Moselle.¹³ Upon the official recovery of Alsace-Lorraine per the Versailles treaty of June 28, 1919, territories of this region kept their administrative structures defined in the aftermath of the Franco-Prussian War: the Territoire-de-Belfort and the département of Meurthe-et-Moselle remained unchanged and the territories formerly in the départements of Bas-Rhin, Haut-Rhin, and Moselle recovered their pre-war structures, barring changes that occurred under German rule between 1871 and 1918.¹⁴

2.3.1.2. The arrondissements reform of 1926. Due to the deterioration of France's budgetary situation after World War I and the relative uselessness of the arrondissement as administrative constituency, the Poincaré government suppressed 106 arrondissements per the decree-law of September 10, 1926, bringing their number to 279 (Verdier 2009). Although four were reestablished by 1940, the configuration of arrondissements defined in 1926 remained unchanged until the end of the Third Republic.¹⁵

Overall, except for these two events and the creation of five dozen cantons, general administrative constituencies remained relatively stable throughout the Third Republic (Ozouf-Marignier and Verdier 2009, 2013).¹⁶ Appendix Figure A4 displays the evolution of their number between 1870 and 1940: from 87 to 90 départements, from 279 to 385 arrondissements, and from 2,861 to 3,028 cantons. Figure 3 shows the geographic configuration of these administrative constituencies at four key dates: 1870, 1871, 1919, and 1926.

2.3.2. Special administrative constituencies

2.3.2.1. Military constituencies. The territorial organization of the military prevalent during most of the Third Republic was based on a series of reforms passed in 1873–1874 (Boulanger 2001, 15–37; Chanet 2006). At its core, the law of “general organization of the army” of July 24, 1873, ensured consistency between military recruitment and military command. It divided the territory into 18 military regions, each further divided into eight subdivisions per the decree of August 6, 1874. Moreover, per the decree of September 28, 1873, one army corps per military

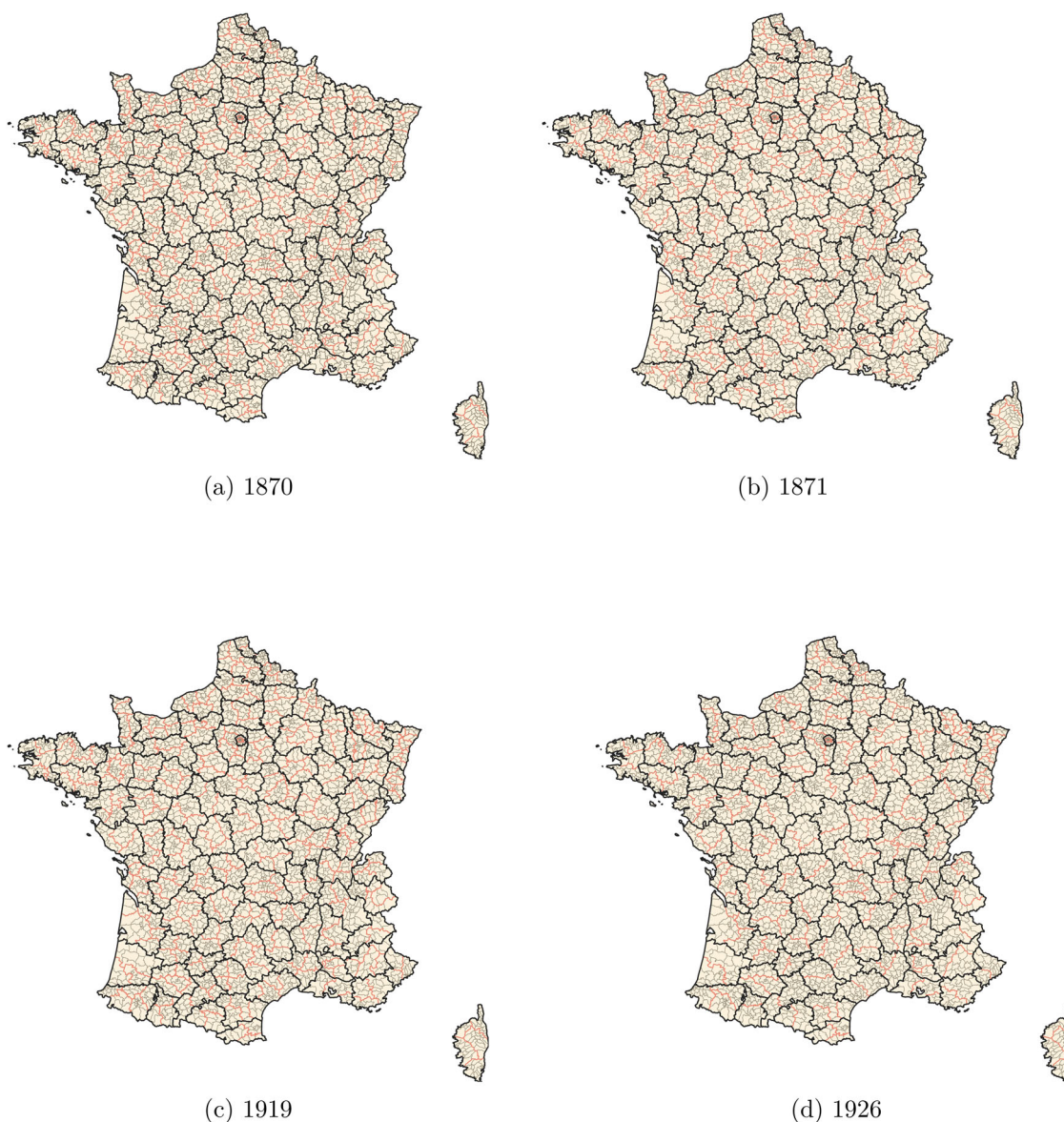


Figure 3. General administrative constituencies in 1870, 1871, 1919, and 1926. Black lines delineate départements, red lines, arrondissements, and gray lines, cantons.

region was created. Each army corps was composed of two infantry divisions of two brigades with two regiments—one infantry regiment per subdivision of military region. In each subdivision of military region was located a recruitment bureau, which managed recruitment and mobilization under the authority of the military region's command. Although they overlapped geographically, military regions were the relevant administrative structures for recruitment, while army corps were relevant for military command.

Except for the regions of Paris and Lyon, which had specific military governments, delineations of military regions and subdivisions followed canton and arrondissement boundaries so as to maintain balance in terms of population, transportation networks,

topography, and cultural homogeneity. This organization was not fundamentally altered until 1940, though it underwent a series of modifications. Until the early-1920s, the military adapted its territorial organization mostly along the north-eastern border in order to ensure efficient mobilization in case of military conflict with Germany.¹⁷ After its recovery in 1919, the military organization in Alsace-Lorraine followed the same logic. Modifications from the late 1920s onward were instead dominated by the willingness of the military command of nesting subdivisions of military regions into département boundaries, which was nearly achieved by the mid-1930s. At odds with this system, the military organization that prevailed before the reforms of 1873–1874 did not exhibit such

consistency between military recruitment and military command. It simply divided troops into six army corps across 22 divisions along département boundaries.¹⁸ Appendix Figure A.5 displays the evolution of the number of military constituencies between 1874 and 1940: from 18 to 20 military regions, and from 100 to 156 subdivisions of military regions.

The construction of annual constituency-level shapefiles for military corps and divisions (1870–1873) and military regions and subdivisions (1874–1940) is based on their characterization in terms of general administrative constituencies. While they overlapped département boundaries between 1870 and 1873, military constituencies generally followed cantons and arrondissement boundaries between 1874 and 1940. For this period, I first match each canton from the TRF-GIS canton-level nomenclature 1874 to its military region and subdivision following the decree of August 6, 1874. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and military constituencies.¹⁹

Annual constituency-level shapefiles for military regions and subdivisions (1874–1940) exhibit the same limitations as canton-level shapefiles, as some subdivisions of military regions were defined across canton and commune boundaries.²⁰ Following the same logic as for canton-level shapefiles, I create pseudo-military constituencies whenever relevant, mostly in the vicinity of the larger urban centers of Paris, Lyon, and Marseille. Appendix Figure A.6 displays the geographic configuration of military constituencies at four key dates: 1870, 1874, 1919, and 1940.

2.3.2.2. Judicial and penitentiary constituencies.

France's judicial organization remained broadly stable from the early nineteenth century to the end of the Third Republic (Farcy 1992, 29–47). Ordinary judicial institutions consisted of five nested jurisdictional levels. At the first level, justices of the peace (*justices de paix*) and simple police courts (*tribunaux de simple police*) had cantons as constituency and were headed by the same cantonal judge. Their attributions were limited to minor civil cases that could be resolved by conciliation and to minor criminal offenses. The second level was constituted by courts of first instance (*tribunaux de première instance*), the ordinary first-degree jurisdiction for civil cases, which also ruled on appeals to justices of the peace. In criminal matters, its criminal courts (*tribunaux correctionnel*) ruled on all offenses for which sentences were shorter than five years but exceeded those imposed by simple police

courts. Until the decree of September 3, 1926, courts of first instance had arrondissements as constituency—except for the département of Seine, which had only one, and the arrondissement of Puget-Théniers, which had none. These courts were located at the subprefecture, barring a dozen exceptions. In the spirit of the arrondissements reform of 1926, the judicial reform of the same year suppressed 228 courts of first instance. However, this reform soon failed and all but six courts were reinstated in 1930 per the law of August 22, 1929—three courts were later suppressed in 1931–1932. At the third level, assize courts (*cours d'assises*) were autonomous départemental jurisdictions that ruled on crimes. They were only in session for one trimester each year except for the assize court of Paris, which was permanent. The fourth jurisdictional level was constituted by courts of appeal (*cours d'appel*), which ruled on appeals to courts of first instance, both in civil and criminal matters. There were 26–28 courts of appeal, each composed of one (Bastia) to seven (Paris) départements. Their delineations remained stable throughout the Third Republic barring modifications induced by the annexation and recovery of Alsace-Lorraine.²¹ Finally, the higher jurisdiction was the Court of Cassation (*Cour de Cassation*), which reviewed other courts' rulings in last resort, both in civil and criminal matters. Appendix Figures A.7a and A.7b display the evolution of the number of judicial constituencies between 1870 and 1940: from 26 to 28 courts of appeal, and from 138 to 370 courts of first instance.

The construction of annual constituency-level shapefiles for courts of appeal is straightforward as these followed département boundaries. For courts of first instance, I rely on their delineations along canton boundaries—although courts of first instance were based on arrondissements until 1929, their delineations from 1930 to 1940 were based on the geography of arrondissements prior to the reform of 1926, which left cantons untouched. I first match each canton from the TRF-GIS canton-level nomenclature 1870 to its court of first instance. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and courts of first instance.²² Because justices of the peace, police courts, and assize courts had cantons and départements as constituencies, their geographies and nomenclatures follow that of these general administrative constituencies. Appendix Figure A.8 displays the geographic configuration of judicial constituencies at four key dates: 1870, 1871, 1926, and 1940.

Besides ordinary judicial institutions, special jurisdictions ruled over litigation specific to some professional groups, such as commercial courts (*tribunaux de commerce*) and labor courts (*conseils de prud'hommes*). Because these only existed in some cities and had heterogeneous constituencies, I do not provide nomenclatures and shapefiles for these special jurisdictions. Annual commune nomenclatures do, however, detail their locations.

Finally, linked to the judicial system, the penitentiary system was administered through regional constituencies that comprised several départements from 1871 onward—in 1870, penitentiary constituencies and départements were confounded. Their number decreased over time, from 45 in 1871–1887 to 16 in 1926–1940 (Appendix Figure A.7c).²³ Appendix Figure A.9 displays the geographic configuration of penitentiary constituencies at four key dates: 1871, 1888, 1909, and 1940.

2.3.2.3. Electoral constituencies. The Third Republic was a Parliamentary regime in which the National Assembly, France's Lower House, elected the head of the executive branch (together with the Senate from 1875 onward). Members of the National Assembly, the députés, were elected by male citizens aged 21 and older through a single-member system, with one député per electoral constituency, denominated *circonscription*—though the general elections of 1871, 1885, 1919, and 1924 used a block-list system. While block-list elections generally had département-level electoral constituencies, arrondissements formed the basis of electoral constituencies throughout the Third Republic (Gaudillère 1995; Marty 2013). Its electoral geography was based on the redistricting of 1875, which followed three principles that barely changed during the next seven decades: each arrondissement had at least one député, and therefore constituted at least one circonscription; each additional hundred thousand inhabitants entitled an arrondissement to an additional député; and a circonscription was modified only if its number of députés had to change due to population growth. There was only one major redistricting after 1875: that of 1927 after the return of the single-member system, which was broadly based on the geography of arrondissements prior to the reform of 1926. The redistricting of 1889, also after the return to the single-member system, was essentially identical to that of 1875. Still, the population rule for redistricting entailed marginal changes after each census, adjusting constituencies in 10–20 arrondissements every five years.²⁴ Appendix Figure A.10a displays the evolution of the number of circonscriptions during the Third

Republic: from 87 to 100 for block-list elections, and from 526 to 598 for single-member elections.

The construction of annual circonscription-level nomenclatures and shapefiles relies on two elements: TRF-GIS annual canton-level nomenclatures and Gaudillère's (1995) *Atlas Historique des Circonscriptions Électorales Françaises*, which provides maps of circonscriptions from 1815 to 1986 for each département, with cantons as underlying frame. Using Gaudillère's (1995) 800-page atlas, I manually match each contemporaneous canton to its electoral constituency for each redistricting year. I then expand these configurations to the following years until the next redistricting, taking into account territorial changes to cantons through TRF-GIS annual canton-level nomenclatures. I use the same methodology but at the level of communes (and of municipal arrondissements of Paris) for the départements of Seine and Seine-et-Oise as circonscriptions were generally defined at such levels in these two départements.

Annual circonscription-level shapefiles exhibit the same limitations as canton-level shapefiles, as some circonscriptions were defined across canton and commune boundaries due to the population rule for redistricting. For instance, the city of Montluçon was divided along canton boundaries into the circonscriptions of Montluçon-1 (canton of Montluçon-Est) and Montluçon-2 (canton of Montluçon-Ouest) during single-member elections. Following the same logic as with canton-level shapefiles, I create pseudo-circonscriptions whenever relevant. Appendix Figure A.11 displays the geographic configuration of electoral constituencies at four key general elections dates: 1876, 1889, 1919, and 1928.

2.3.2.4. Academic constituencies. The administration of public education during the Third Republic was operated through a system of regional academies that was defined in 1854 and that overlapped département boundaries. Since then, each academy was administered by a rector who supervised primary, secondary, and higher education within the academy. He was assisted by several academic inspectors—one per département, two or more in the populous départements of Nord and Seine. The territorial configuration of the 17 academies that existed by 1870 was seldom modified during the Third Republic.²⁵ Appendix Figure A.10b displays the evolution of the number of academies during the Third Republic, which remained at 16–17.

The construction of annual academy-level shapefiles is straightforward as academies were supra-

départemental constituencies that underwent little modifications. I first match each département from the TRF-GIS département-level nomenclature 1870 to its academy per the decrees of August 22nd, 1854, and June 13th, 1860. Then, for each year onward, I update this initial configuration with territorial modifications to both départements and academies.²⁶ Appendix Figure A.12a displays the geographic configuration of academic constituencies in 1919.

2.3.2.5. Labor inspection constituencies. The service of labor inspection of the Ministry of Labor was created by the law of May 19th, 1874, which objective was to regulate child labor. However, it was not until the law of December 13th, 1892 that a service of labor inspectors structured the territory in order to control the application of labor regulations (Chetcuti and Le Noël 1998). It was initially organized into 11 supra-départemental constituencies, each headed by one divisional inspector. Divisional inspectors were assisted by several départemental inspectors operating in territorial sections—92 in total. Labor inspection constituencies underwent little modifications after 1892.²⁷ Appendix Figure A.10c displays the evolution of the number of labor inspection constituencies during the Third Republic, which remained at 11–12.

To construct annual labor inspection-level shapefiles, I first match each canton from the TRF-GIS canton-level nomenclature 1892 to its labor inspection constituency per the law of December 13th, 1892—although these constituencies were supra-départemental, cantons of the département of Isère were initially split between the 10th and 11th constituency. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and labor inspection constituencies. Appendix Figure A.12b displays the geographic configuration of labor inspection constituencies in 1919.

2.3.2.6. Ecclesiastical constituencies. Despite the Separation of the Churches and the State in 1905, catholic ecclesiastical constituencies (dioceses) continued to play an important role in the French society and to produce statistics (Sorrel 2013). Inherited from the geography of Roman provinces, most dioceses were created between the second and the fifth centuries. Their delineations were substantially altered twice after the Revolution: by the Concordat of 1801, which broadly nested the diocesan geography into that of the newly created départements, and by the Papal bull *Paternae charitatis* of 1822, which restored some pre-revolutionary dioceses (Duquesnoy 2020). By 1870,

the metropolitan territory was divided into 86 dioceses, among which 17 archbishoprics (*archevêchés*) and 69 bishoprics (*évêchés*). Ruled by an archbishop, archbishoprics were metropolitan sees and headed the territories of several subordinate bishoprics. Some episcopal sees were directly subject to the Holy See—the dioceses of Metz and Strasbourg after 1919. Throughout the Third Republic, the territories of most dioceses followed département boundaries. Still, some followed arrondissements.²⁸ The diocese of Carcassonne (Aude) even had an enclave in the département of Ariège through the canton of Quérigut. The diocesan geography of the départements of Savoie and Haute-Savoie was more complex, as it divided their territories along commune boundaries (Deloche 2017).²⁹ Dioceses boundaries underwent little changes during the Third Republic.³⁰ Appendix Figure A.10d displays the evolution of the number of dioceses between 1870 and 1940, which remained at 84–87.

To construct annual diocese-level shapefiles, I first match each canton from the TRF-GIS canton-level nomenclature 1870 to its diocese per the Papal bull of 1822, taking into account the creation of the diocese of Laval in 1855. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and dioceses. I use the same methodology but at the level of communes for the départements of Savoie and Haute-Savoie. Appendix Figure A.12c displays the geographic configuration of dioceses in 1919, together with the extent of metropolitan bishoprics.

3. Data records

The TRF-GIS database is composed of two types of data files: annual constituency-level nomenclatures and annual constituency-level shapefiles.³¹ Annual constituency- and commune-level nomenclatures are available in Stata data format (.dta) as well as in text delimited format (.txt), with the main file-specific metadata replicated as notes within .dta files and along with separate codebooks for .txt files. Each annual constituency-level shapefile is composed of a (zipped) set of five files: a shape format file (.sph), a shape index format file (.shx), an attribute format file (.dbf), a projection description file (.prj), and a character encoding file (.cpg). These shapefiles use an RGF93 Lambert-93 projection, IGN's reference projection system. I further follow IGN's methodology and provide shapefiles both in polygon and vector forms. Shapefiles metadata are also

provided for each constituency in QGIS metadata format (.qmd). Furthermore, separate datasets of chefs-lieux coordinates are provided in .csv format—these coordinates correspond to the current locations of their city halls.

The TRF-GIS follows FAIR data management practices (Wilkinson et al. 2016). Data files are available in the TRF-GIS Dataverse, hosted by the Harvard Dataverse, under the CC-BY 4.0 license at <https://dataverse.harvard.edu/dataverse/TRF-GIS>. Therein, each constituency constitutes a separate dataset—15 datasets in total. Each dataset is attributed a unique and persistent Digital Object Identifier (DOI) for better findability and citability. Data files are curated through DataCite Metadata Schema 4.3 and use the Library of Congress Subject Heading (LCSH) controlled vocabulary for topic classification.

3.1. General administrative constituencies

3.1.1. Départements

Annual département-level nomenclatures are composed of 10 variables (Appendix Table A.24). They include départements' codes, names, and whether they were composed of enclaved territories and islands (e.g., the Enclave des Papes of the département of Vaucluse, enclaved in the territory of the département of Drôme). They also include information on départements' chefs-lieux (prefectures): their INSEE codes, names, whether they are current or former communes (e.g., the prefecture of the département of Ardennes was located in the former commune of Mézières), and their geographic coordinates. For départements codification, I give number 99 to the département of Meurthe in 1870. Codifications and toponyms are otherwise standard. Note that although the TRF-GIS treats the Territoire-de-Belfort as a département, it did not achieve this status officially until 1922. Annual département-level shapefiles contain the same variables as département-level nomenclatures.

3.1.2. Arrondissements

Annual arrondissement-level nomenclatures are composed of 15 variables (Appendix Table A.24). They include arrondissements' département code and name as well as arrondissements' codes and names. They also include information on arrondissements' chefs-lieux (subprefectures): their INSEE codes, names, whether they are current or former communes, whether they were common to other arrondissements (e.g., Strasbourg was the subprefecture of

the arrondissements of Strasbourg-Campagne and of Strasbourg-Ville), whether they were located outside of their arrondissement (e.g., Strasbourg was located outside of the arrondissement of Strasbourg-Campagne), and their geographic coordinates. Arrondissements one-digit codification uniquely identifies arrondissements within départements and follows the COG 2011. For arrondissements that no longer existed by then, I assign a higher number in ascending alphabetical order within each département and keep this codification constant so that arrondissements can be uniquely identified across time. Moreover, I keep the contemporaneous toponymy of arrondissements, which names generally were that of their subprefectures.³² Note that although the TRF-GIS treats municipal arrondissements of Paris and the Territoire-de-Belfort as arrondissements, these territories did not have this status officially. Annual arrondissement-level shapefiles contain the same variables as arrondissement-level nomenclatures.

3.1.3. Cantons

Annual canton-level nomenclatures are composed of 18 variables (Appendix Table A.24). They include cantons' département and arrondissement codes and names as well as cantons' codes, names, and communal composition. The later variable, which follows INSEE's COGs codification convention, indicates whether a canton was composed of entire communes and/or of fractions of communes (e.g., the canton of Montluçon-Est was composed of eight entire communes and a fraction of the commune of Montluçon). They also include information on cantons' chefs-lieux (*bureaux centralisateurs*): their INSEE codes, names, whether they are current or former communes, whether they were common to other cantons (e.g., Montluçon was the bureau centralisateur of both cantons of Montluçon-Est and Montluçon-Ouest), and their geographic coordinates. Cantons two-digit codification uniquely identifies cantons within départements and follows the COG 2011. For cantons that no longer existed by then, I use the codification in the COG 1954. For those that no longer existed by 1954, I assign a higher number in ascending alphabetical order within each département and keep this codification constant so that cantons can be uniquely identified across time. Moreover, I keep the contemporaneous toponymy of cantons, which names generally included that of their bureau centralisateur.³³ Note that although the TRF-GIS treats municipal arrondissements of Paris as cantons, these

territories did not have this status officially. Moreover, the communes of partially annexed cantons that had remained in France after the Franco-Prussian War were not re-assigned to functioning cantons until 1873. As a result, their chefs-lieux were (virtually) outside of France in 1871 and 1872.³⁴

Because the commune-level GEOFLA 2011 shapefile forms the base frame for annual canton-level shapefiles, and cantons were sometimes defined across commune boundaries, TRF-GIS shapefiles do not entirely match canton-level nomenclatures. Instead, they use the concept of pseudo-canton. Following INSEE's and IGN's codification conventions, pseudo-cantons are assigned codes greater than 80 in ascending alphabetical order within départements. Moreover, their toponymy corresponds to the common city infra-communal cantons are aggregated into. For instance, in the case of the Montluçon area displayed in Figure 2, canton-level shapefiles contain the pseudo-canton of Montluçon, which extent corresponds to the commune of Montluçon and which is given pseudo-canton code 80. During the Third Republic, 10–12 percent of cantons were composed of fractions of communes, which translated into 110–120 pseudo-cantons (Figure A.4c).³⁵ Annual canton-level shapefiles contain the same variables as canton-level nomenclatures, except that their base units are pseudo-cantons rather than cantons. To reflect this difference, variables prefix in these shapefiles is `pct` rather than `ct`.

3.2. Special administrative constituencies

3.2.1. Military constituencies (1870–1873)

Annual military corps-level nomenclatures are composed of 10 variables (Appendix Table A.25). They include military corps' codes, names, and composition (entire départements). They also include information on military corps' chefs-lieux: their INSEE codes, names, whether they are current or former communes, and their geographic coordinates. Codifications and toponyms are contemporaneous, with toponyms corresponding to military corps chefs-lieux. Annual military corps-level contain the same variables as military corps-level nomenclatures.

Annual military division-level nomenclatures are composed of 14 variables (Appendix Table A.25). They include military divisions' corps code and name as well as military divisions' codes, names, and composition (entire départements). They also include information on military divisions' chefs-lieux: their INSEE codes, names, whether they are current or

former communes, and their geographic coordinates. Codifications and toponyms are contemporaneous, with toponyms corresponding to military divisions chefs-lieux. Annual military divisions-level shapefiles contain the same variables as military divisions-level nomenclatures.

3.2.2. Military constituencies (1874–1940)

Annual military region-level nomenclatures are composed of 10 variables (Appendix Table A.25). They include military regions' codes, names, and composition. For instance, throughout the Third Republic, the 2nd military region (Amiens) was composed of fractions of arrondissements and of entire cantons. They also include information on military regions' chefs-lieux: their INSEE codes, names, whether they are current or former communes, and their geographic coordinates. Codifications and toponyms are contemporaneous, with toponyms corresponding to military regions chefs-lieux. Because the commune-level GEOFLA 2011 shapefile forms the base frame for military region-level shapefiles, and military regions were sometimes defined across cantons and commune boundaries, these shapefiles do not entirely correspond to military region-level nomenclatures. Instead, these shapefiles use the concept of pseudo-military region. Pseudo-military regions are assigned military region codes greater than 90 in ascending alphabetical order, and their names correspond to their territorial composition. For instance, the city of Lyon was divided between the military regions of Grenoble and Besançon from 1874 to 1921. This pseudo-military region is assigned pseudo-region code 90 and name “Grenoble–Besançon.” Pseudo-military regions are not assigned a chef-lieu, so all variables with prefix `cl` are empty for actual pseudo-military regions.³⁶ Annual military region-level shapefiles contain the same variables as military region-level nomenclatures, except that their base units are pseudo-military regions rather than military regions. To reflect this difference, variables prefix in these shapefiles is `pmreg` rather than `mreg`. In addition, the variable `pmreg_type` indicates whether military regions were effectively pseudo-military regions.

Annual military subdivision-level nomenclatures are composed of 15 variables (Appendix Table A.25). They include military subdivisions' region code and name as well as military subdivisions' codes, names, and composition. They also include information on military subdivisions' chefs-lieux: their INSEE codes, names, whether they are current or former communes, whether they were common to other

subdivisions (e.g., Rouen was the chef-lieu of both military subdivisions of Rouen-Nord and Rouen-Sud), and their geographic coordinates. Codifications and toponyms are contemporaneous, with toponyms corresponding to military subdivisions chefs-lieux. Because the commune-level GEOFLA 2011 shapefile forms the base frame for military subdivision-level shapefiles, and military subdivisions were sometimes defined across cantons and commune boundaries, these shapefiles do not entirely correspond to military subdivision-level nomenclatures. Instead, these shapefiles use the concept of pseudo-military subdivision. Pseudo-military subdivisions are assigned military subdivision codes greater than 10 in ascending alphabetical order, and their names correspond to their territorial composition. For instance, the city of Rouen was divided between the military subdivisions of Rouen-Nord and Rouen-Sud from 1874 to 1929. This pseudo-military region is assigned code 11 and name “Rouen (Nord–Sud).” Pseudo-military subdivisions are not assigned a chef-lieu, so all variables with prefix `c1` are empty for actual pseudo-military subdivisions.³⁷ Annual military subdivision-level shapefiles contain the same variables as military subdivision-level nomenclatures, except that their base units are pseudo-military subdivisions rather than military subdivisions. To reflect this difference, variables prefix in these shapefiles is `pmsub` rather than `msub`. In addition, the variable `pmsub_type` indicates whether military subdivisions were effectively pseudo-military subdivisions.³⁸

3.2.3. Judicial and penitentiary constituencies

Annual court of appeal-level nomenclatures are composed of 9 variables (Appendix Table A.26). They include courts of appeal’s codes and names, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Toponyms are contemporaneous and correspond to courts of appeal chefs-lieux. Courts of appeal two-digit codification uniquely identifies these constituencies over time and is assigned in ascending alphabetical order. Annual court of appeal-level shapefiles contains the same variables as court of appeal-level nomenclatures.

Annual court of first instance-level nomenclatures are composed of 14 variables (Appendix Table A.26). They include their court of appeal’s code and name, as well as courts of first instances’ codes and names. They also include information on courts of first instances’ chefs-lieux: their INSEE codes, names, whether they were also a subprefecture, and their geographic coordinates. Toponyms are contemporaneous

and correspond to courts of first instance chefs-lieux. Courts of first instance two-digit codification uniquely identifies these constituencies over time within their court of appeal and is assigned in ascending alphabetical order. Annual court of first instance-level shapefiles contains the same variables as court of first instance-level nomenclatures.

Annual penitentiary constituency-level nomenclatures are composed of 9 variables (Appendix Table A.26). They include penitentiary constituencies’ codes and names, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Toponyms are contemporaneous and correspond to penitentiary constituencies’ chefs-lieux. Penitentiary constituencies two-digit codification uniquely identifies these constituencies over time and is based on the contemporaneous codification defined in the official act of 1871 reorganizing penitentiary constituencies—for 1870, it corresponds to départements codes; penitentiary constituencies created after 1871 are assigned codes that are ascending in the temporal order of their creation. Annual penitentiary constituency-level shapefiles contain the same variables as penitentiary constituency-level nomenclatures.

3.2.4. Electoral constituencies

Annual circonscription-level nomenclatures are composed of 6 variables (Appendix Table A.27). They include circonscriptions’ département code as well as circonscriptions’ codes, names, and composition. The later variable indicates whether a circonscription was composed of entire départements, arrondissements, cantons, or communes. Some codes specifically refer to the circonscriptions of Paris. Contrary to other constituencies, electoral constituencies are uniquely identified by their name, which corresponds to their arrondissement name, with the addition of a number for the larger urban centers, or, between 1928 and 1940, to the most important city of the constituency—sometimes two adjoined names, such as “Nantua–Gex.” For convenience, I provide a codification in ascending alphabetical order within each département, but it is not necessarily consistent over time due to the changing nature of circonscriptions’ boundaries. Likewise, urban circonscriptions which contemporaneous names were composed of the arrondissement name and a number, such as “Saint-Denis-3” in Seine-et-Oise, might not be consistent over time—this is generally the case in the more urban areas in which redistricting was frequent. Moreover, no chefs-lieux are indicated as electoral

circonscriptions were not administrative units, but only territorial constituencies—electoral operations were instead coordinated at cantons' bureaux centralisateurs. Each annual circonscription-level nomenclature further indicates three elements among its metadata: the election system (block-list or single-member), whether it was a general election year, and the relevant redistricting year. Annual circonscription-level shapefiles contain the same variables as circonscription-level nomenclatures, except that their base units are pseudo-circonscriptions rather than circonscriptions. To reflect this difference, variables prefix in these shapefiles is `pcirco` rather than `circo`. In addition, the variable `pcirco_type` indicates whether circonscriptions were effectively pseudo-circonscriptions.³⁹

3.2.5. Academic constituencies

Annual academy-level nomenclatures are composed of 8 variables (Appendix Table A.27). They include academies' codes and names, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Toponyms are contemporaneous and correspond to academies chefs-lieux. Academies two-digit codification uniquely identifies these constituencies over time and is assigned in ascending alphabetical order—except for the academy of Chambéry, created in 1860 and suppressed in 1919, which holds academy code 17. Annual academy-level shapefiles contain the same variables as academy-level nomenclatures.

3.2.6. Labor inspection constituencies (1892–1940)

Annual labor inspection-level nomenclatures are composed of 6 variables (Appendix Table A.27). They include labor inspections codes, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Labor inspection constituencies did not have a specific toponymy beyond their 2-digit codification, which was fixed over time. Annual labor inspection-level shapefiles contain the same variables as labor inspection-level nomenclatures.

3.2.7. Ecclesiastical constituencies

Annual diocese-level nomenclatures are composed of 14 variables (Appendix Table A.27). They include the archbishopric's code and name each diocese was attached to, its type (metropolitan see or exempt), as well as dioceses codes, names, and types (archbishoprics or bishoprics). They further include information on their chefs-lieux: their INSEE codes, names, and

geographic coordinates. The diocesan toponymy is contemporaneous and corresponds to customary dioceses names, which consisted in the locations of their chefs-lieux (episcopal sees)—except for the diocese of Aire, which episcopal see was transferred to Dax in 1933, and for the diocese of Tarentaise, which name referred the Tarentaise Valley. Different from their customary names, dioceses' formal names were composed of all prerevolutionary dioceses their bishops were entitled. For instance, by 1854, the bishop of Bayeux was entitled the former diocese of Lisieux. As a result, the formal name of the diocese of Bayeux became "Bayeux-Lisieux." These names were sometimes cumbersome: for instance, by 1877, the formal name of the diocese of Avignon was "Avignon-Apt-Cavaillon-Carpentras-Orange-Vaison." Annual diocese-level nomenclatures therefore use dioceses customary names.⁴⁰ Dioceses two-digit codification uniquely identifies these constituencies over time and is assigned in ascending alphabetical order—except for the archbishopric of the dioceses of Metz and Strasbourg after 1919, exempt and subject to the Holy See, which is assigned code 99. Annual diocese-level shapefiles contain the same variables as diocese-level nomenclatures.

3.3. Annual commune-level nomenclatures

Annual commune-level nomenclatures are composed of 69 variables (Appendix Table A.28). Each commune can be identified through its unique Cassini code from the HAC database as well as through its INSEE code in 2005 geography.⁴¹ Communes toponyms are contemporaneous. These nomenclatures include the names and codes of all the above general and special administrative constituencies each commune belonged to in a given year—pseudo-constituencies in the cases of cantons, military regions and subdivisions, and electoral circonscriptions. They also include indicator variables for whether a commune was the chef-lieu of a constituency or the siege of a special jurisdiction: a commercial or a labor court (Farcy 1992, 875–86). Appendix Table A.28 describes the 21 additional variables included in annual commune-level nomenclatures that are not present in other constituency-level nomenclatures.

To enable users to combine statistics produced by administrations at different and potentially ex-ante incompatible levels of aggregation, commune-level nomenclatures further provide population counts from the HAC database, which reports this

information based on population censuses.⁴² During the Third Republic, population censuses were carried out every five years—except in 1871, when it was postponed to 1872 due to the Franco-Prussian War, and in 1916, when it was canceled due to the First World War. Three variables relate to population counts: `pop`, the census population count, `ipop`, the interpolated population count, and `pop_flag`, a population count flag. The `pop` variable is missing outside of census years (extended missing value `.a`) and for a few communes during census years: when a commune was created during the same year as a census but after it was carried out (`.b`), when the census archive was lost (`.c`) or unreadable (`.d`), and when the commune was not enumerated (`.e`). Except in 1872, when 233 communes were not enumerated (nearly exclusively in the département of Hautes-Pyrénées) and the archives for another 482 communes were lost (exclusively in the département of Vosges), missing population counts during census years remain negligible: they always account for less than 15 communes—less than 3 communes in 9 out of all 13 censuses that were carried out during the Third Republic. Based on these raw population counts, the variable `ipop` provides interpolated population counts for all years and all communes present in annual commune-level nomenclatures. The variable `pop_flag` documents the imputation method used: none for census years (1), linear interpolation for inter-census years (2), linear interpolation for census years when the census archive was lost (3) or unreadable (4), linear interpolation for census years when a commune was not surveyed (5), and imputation from the previous (6) or next (7) census for communes that were created or suppressed during an inter-census year. Imputations remain limited to about 30 communes across all years after 1872, and to about a hundred in 1870–1871.

Finally, annual commune-level nomenclatures have a specific treatment of Paris: they provide both municipal *arrondissement* and commune-level information. Municipal *arrondissements* of Paris hold INSEE codes 75101–75120, and the commune of Paris, INSEE code 75056. Whenever Paris is the *chef-lieu* of a constituency, this prerogative falls onto the first municipal *arrondissement* as well as onto the commune. Whenever it is subdivided into constituencies, the commune of Paris holds the status of pseudo-constituency with corresponding codifications: it holds pseudo-canton code 7580, pseudo-military region and subdivision code 99, and pseudo-circonscription code 99.

4. Technical validation

4.1. TRF-GIS nomenclatures

Information supporting TRF-GIS nomenclatures documented in secondary sources were thoroughly verified against primary sources.⁴³ Appendix Table A.29 details the content of all 175 primary sources used for this verification procedure: the constituency they are relevant to, their codification (date of official issue and type), their title, their persistent identifier (PID) assigned by Gallica and the National Library of France (BNF) in archival resource key (ARK) form, and their origin. Primary sources' types take the following forms: law (L), decree (D), order (A, for *arrêté*), ordinance (O), and papal bull (B). Their origin is generally the *Journal Officiel*, though sometimes it is the *Bulletin des Lois*, the *Code Pénitentiaire* or *Code des Prisons*, the *Bulletin Administratif de l'Instruction Publique*, and the *Acta Apostolicae Sedis*. All these primary sources are available in PDF format in the TRF-GIS dataverse. This rigorous verification procedure combined with data integrity tests in the TRF-GIS source code enabled to identify 231 inaccuracies or errors in the HAC database and to rectify them.⁴⁴

Nomenclatures were also verified against periodical statistical publications of relevant administrations, among which censuses for general administrative constituencies, the *Compte Rendu sur le Recrutement de l'Armée* for military constituencies, the *Compte Général de l'Administration de la Justice Civile et Commerciale* and the *Compte Général de l'Administration de la Justice Criminelle* for judicial constituencies, the *Statistique Pénitentiaire* for penitentiary constituencies, and the *Tableau des Élections à la Chambre des Députés* for electoral constituencies.⁴⁵

4.2. TRF-GIS shapefiles

IGN's GEOFLA 2011 shapefile constitutes the underlying frame of TRF-GIS shapefiles. As discussed above, I use this dataset for quality and reproducibility purposes: it constitutes the official representation of France's communes in 2011 geography and is available under open licensing. Although IGN's BD CARTO[®] shapefiles are more precise than the GEOFLA (1:25,000 against 1:1,000,000), these files are not all freely accessible. Their use would therefore substantially alter the possibilities of reproduction and dissemination of TRF-GIS shapefiles. Moreover, their manipulation is more cumbersome: while GEOFLA is

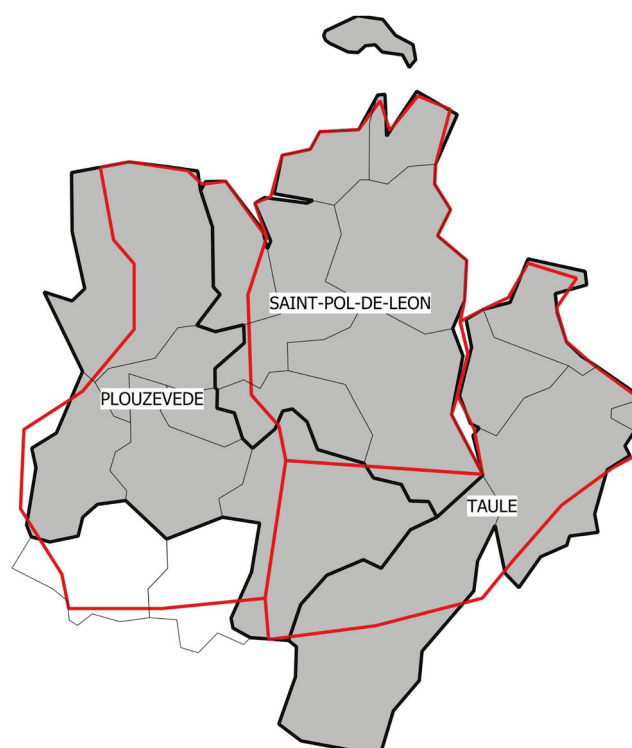


Figure 4. Comparison of TRF-GIS and LARHRA cantons shapefiles 1884. TRF-GIS cantons have thick black boundaries while LARHRA's (2011) cantons have thick red boundaries. Thin boundaries correspond to communes in the GEOFLA 2011. The three cantons are Plouzevé, Saint-Pol-De-Léon, and Taulé in the département of Finistère (29).

available as an ensemble, BD CARTO files are only available separately for each département, and their overall size approaches 2.5 GB against 150 MB for the GEOFLA. In the end, the lower precision of the GEOFLA need not be very problematic because TRF-GIS shapefiles aim at regional and national extents (see Sec. 2.2).

The general methodology used to construct TRF-GIS shapefiles offers several advantages over the manual vectorization of georeferenced historical maps (as in, e.g., LARHRA 2011; Perret, Gribaudo, and Barthélemy 2015; Ostafin et al. 2020). First, it yields more precise results than existing outputs, such as LARHRA's (2011) cantons shapefiles for 1884 and 1925: historical maps with national extents oftentimes lacked precision and had unspecified projection systems, making the resulting georeferencing potentially approximate, and urban centers, Corsica, and islands challenging to vectorize.⁴⁶ Typical inaccuracies that result from these methods are displayed in Figure 4, which plots three randomly selected cantons from TRF-GIS and LARHRA cantons shapefiles 1884. Five out of twenty communes (which territories did not change historically) are classified in the wrong canton

by LARHRA's (2011) shapefile.⁴⁷ Second, and more importantly, historical maps are generally not available annually for general administrative constituencies, and are even more rare for special administrative constituencies, which implies that relying uniquely on the vectorization of georeferenced historical maps would leave substantial gaps in a historical GIS of France. Finally, manual vectorization methods are quite time and resource consuming, making it challenging to build annual shapefiles with realistic resources and to verify their accuracy.

5. Usage notes

5.1. Data availability

The TRF-GIS database can be accessed through the TRF-GIS Dataverse under the CC-BY 4.0 license, which is hosted in the Harvard Dataverse and accessible at <https://dataverse.harvard.edu/dataverse/TRF-GIS>. The reproduction material is available in the "TRF-GIS Reproduction Material" dataset at <https://doi.org/10.7910/DVN/6FCK3W>. It contains source data, reproduction codes, and primary archival sources. Although each file can be accessed independently, I recommend downloading the full dataset for reproducibility purposes in order to preserve the integrity of the folders structure.

5.2. Code availability

The TRF-GIS database was produced using Stata SE version 16 as well as QGIS version 3.12 for its shapefiles. Stata `.do` files and python `.py` files needed to reproduce TRF-GIS data files are available in the "TRF-GIS Reproduction Material" dataset under the MIT open license. Using a standard processor of 4.00 GHz with 16 GB RAM, the total processing time of the master Stata `.do` file is 7 hours, and 8 hours for the python `.py` files constructing shapefiles.

5.3. Walk-through for map production

The production of maps is straightforward for constituencies that are not subdivided into pseudo-constituencies, as nomenclatures in this case exactly correspond to shapefiles.⁴⁸ Suppose that a user is willing to map rates of abstention during the first round of legislative elections of 1914 at the département level. They should proceed in the following four steps: (1) Download the TRF-GIS 1914 nomenclature and shapefile of départements at <https://doi.org/10.7910/DVN/ULQYM5>. (2) Input abstention data from the

Tableau des élections générales à la chambre des députés pour la onzième législature (26 avril et 10 mai 1914) onto the nomenclature, and save the data as a .xlsx file.⁴⁹ (3) In QGIS, join and relate département-level abstention rates from the .xlsx file to the attribute table of the DEPARTEMENTS_1914 shapefile based on the dep field. (4) Map département-level abstention rates (Appendix Figure A.13).

The production of maps for constituencies that are subdivided into pseudo-constituencies is more challenging.⁵⁰ Suppose that a user is willing to map rates of abstention during the first round of legislative elections of 1914 at the circonscription level. They should proceed in the following six steps: (1) Download the TRF-GIS 1914 nomenclature and shapefile of circonscriptions at <https://doi.org/10.7910/DVN/L2LGDW>. (2) Input abstention data from the *Tableau des élections générales à la chambre des députés pour la onzième législature (26 avril et 10 mai 1914)* onto the nomenclature, and save the data as a .xlsx file.⁵¹ (3) In QGIS, export the attribute table from the CIRCONSCRIPTIONS_1914 shapefile to an .xlsx file, in order to get the list of pseudo-circonscriptions and their types through the `pcirco_type` variable (unique or multiple). (4) Attribute the raw data for the circonscriptions not divided into pseudo-circonscriptions, and sum the raw data for the corresponding pseudo-circonscriptions.⁵² (5) In QGIS, join and relate circonscription-level abstention rates from the .xlsx file to the attribute table of the CIRCONSCRIPTIONS_1914 shapefile based on the `deppcirco` field. (6) Map circonscription-level abstention rates (Appendix Figure A.14).

6. Conclusion

This article describes a comprehensive geographic information system of Third Republic France: the TRF-GIS. This historical GIS provides annual nomenclatures along with shapefiles of the most significant administrative constituencies of metropolitan France from 1870 to 1940. By providing a comprehensively-curated common frame of reference that encompasses all aspects of the French society, the TRF-GIS will empower empirical research programs in the context of Third Republic France and enhance their interoperability. This frame of reference is also complementary to historical GISs of France focused on the development of the railroad network over this time period (Thévenin, Schwartz, and Sapet 2013; Mimeur et al. 2018). While its nomenclature component is exact,

the main limitation of its cartographic component lies in its reliance on delineations of current commune as its building blocks. Nevertheless, the availability of commune boundaries for this time period in the near future means that this limitation is only temporary.⁵³ A broader challenge a historical GIS such as the TRF-GIS faces, however, lies in its ability to provide researchers with tools to analyze and map consistent statistical time series over this seven-decades period, given the many changes in boundaries of administrative units documented in this article. The methodology developed by the ISTADS project for the U.S. for maximizing cross-temporal comparability across evolving administrative units offers interesting opportunities in this perspective (Noble et al. 2011).

Notes

1. Historical GISs on other aspects have been produced for France, for instance on the development of infrastructure networks (Thévenin, Schwartz, and Sapet 2013; Perret, Gribaudi, and Barthelemy 2015; Mimeur et al. 2018).
2. A notable exception is LARHRA's cantons shapefiles for 1884 and 1925, which were produced by the manual vectorization of georeferenced historical maps (LARHRA 2011).
3. Appendix Figure A.1 provides an overview of the construction logic of the TRF-GIS. INSEE is France's national statistics bureau and stands for *Institut National de la Statistique et des Études Économiques*. IGN is France's national bureau of geographical information and stands for *Institut Géographique National*.
4. The COG 2005 is available at <https://www.insee.fr/fr/information/2560656>, and the COG 2011, at <https://www.insee.fr/fr/information/2560625>. (accessed in July 2021).
5. The LaDéHiS is a research lab at EHESS and stands for *Laboratoire de Démographie et d'Histoire Sociale* (formerly *Laboratoire de Démographie Historique*). INED is the French Institute for Demographic Studies and stands for *Institut National d'Études Démographiques*.
6. Voisinlieu's record is available at http://cassini.ehess.fr/cassini/fr/html/fiche.php?select_resultat=40965 (accessed in July 2021).
7. The GEOFLA 2011 shapefile is available at <https://geoservices.ign.fr/geofla> (accessed in July 2021).
8. RGF93, which stands for *Réseau Géodésique Français 1993*, is the projection system used by IGN.
9. This is among the endeavors of the ongoing project COMMUNE HIS-DBD (<https://anrcommunes.hypotheses.org/>), which is scheduled to be completed by 2023.
10. For local extents, specific tools such as ALPAGE for Paris (<https://alpage.huma-num.fr/>) might be more relevant (Noizet, Bove, and Costa 2013).

11. More precisely, the National Constituent Assembly initially divided France into départements, districts, and cantons. Districts were suppressed in 1795 and substituted with arrondissements in 1800.
12. The composition of annexed territories is documented in Appendix Table A.1.
13. The reconfiguration of these territories' arrondissements and cantons between 1871 and 1873 is documented in Appendix Table A.2.
14. Changes that occurred under German rule between 1871 and 1918 are documented in Appendix Table A.3.
15. Territorial changes induced by this reform are documented in Appendix Table A.4.
16. The creation and modifications of cantons during the Third Republic is documented in Appendix Table A.5.
17. Among the main modifications, a 20th military region (Nancy) was created in 1898 through the splitting of the 6th military region (Châlons-sur-Marne) as well as a 21st military region (Épinal) by the end of 1913 through the splitting of the 7th military region (Besançon).
18. The territorial configuration of military divisions in 1870 is documented in Appendix Table A.6.
19. The 35 territorial modifications to military constituencies are documented in Appendix Table A.7.
20. For instance, from 1874 onward, all but one subdivision of the 2nd military region (Amiens) were composed of fractions of the cantons of Saint-Denis and Pantin (Seine), as well as fractions of the 10th, 19th, and 20th municipal arrondissements of Paris.
21. The courts of appeal of Metz and Colmar were suppressed in 1871, the later being reinstated in 1919.
22. Modifications that affected courts of appeal and courts of first instance, along with other details, are documented in Appendix Table A.8.
23. These modifications are documented in Appendix Table A.8.
24. Appendix Table A.9 documents all 13 redistricting laws enacted during the Third Republic.
25. In 1871, the Territoire-de-Belfort was transferred to the academy of Besançon; in 1888, the chef-lieu of the academy of Douai was transferred to Lille; in 1919, the recovered département of Moselle was transferred to the academy of Strasbourg; and in 1920, the academy of Chambéry was suppressed and its départements (Savoie and Haute-Savoie) transferred to the academy of Grenoble (Marichal 1962, 751–9).
26. The few territorial modifications to academies are documented in Appendix Table A.10.
27. The sieges of the 2nd and 4th constituencies were respectively transferred from Châteauroux to Tours and from Bar-le-Duc to Nancy in 1893; the département of Isère, initially split between the 10th and 11th constituency, was reunited into the 11th constituency in 1902; and the département of Somme was transferred from the 5th to the 6th constituency in 1911, then back to the 5th constituency in 1919. After World War I, a 12th constituency was created with Strasbourg as chef-lieu in order to structure labor inspections in the départements of Moselle, Bas-Rhin, and Haut-Rhin.
28. The diocese of Marseille, which corresponded to the arrondissement of Marseille; the diocese of Reims, which corresponded to the département of Ardennes and the arrondissement of Reims (Marne); the diocese of Fréjus, which corresponded to the département of Var and the arrondissement of Grasse (Alpes-Maritimes) until 1886; the dioceses of Lille and Cambrai, which split the département of Nord after 1913.
29. This configuration is documented in Appendix Table A.11.
30. Beyond those described above, the Territoire-de-Belfort was transferred from the diocese of Strasbourg to that of Besançon after the Franco-Prussian War, and the territories of the département of Meurthe-and-Moselle were attributed to the diocese of Nancy. Appendix Table A.12 documents these changes.
31. A description of each these data files' names, content, formats, sizes, and locations is available in Appendix Table A.20 for general administrative constituencies, in Appendix Table A.21 for military constituencies, in Appendix Table A.22 for judicial and penitentiary constituencies, and in Appendix Table A.23 for electoral, academic, labor inspection, and ecclesiastical constituencies. Appendix Table A.23 also describes data files of commune-level nomenclatures.
32. Appendix Table A.13 documents the ten toponymic modifications to arrondissements that occurred during the Third Republic.
33. Appendix Table A.14 documents the 206 toponymic modifications to cantons that occurred during the Third Republic.
34. This was the case for 36 communes of the annexed cantons of Gorze (Moselle), Vic-sur-Seille and Lorquin (Meurthe), and Saales and Schirmeck (Vosges).
35. Appendix Table A.15 documents the mapping between cantons and pseudo-cantons—this mapping is also available as a dataset alongside annual canton-level nomenclatures.
36. Appendix Table A.16 documents the mapping between pseudo-cantons and pseudo-military regions. This mapping is also available as a dataset alongside annual military region-level nomenclatures.
37. Appendix Table A.17 documents the mapping between pseudo-cantons and pseudo-military subdivisions. This mapping is also available as a dataset alongside annual military subdivision-level nomenclatures.
38. Appendix Table A.18 documents the composition of pseudo-military subdivisions in terms of military subdivisions.
39. Because shapefiles are stored in associated .dbf files, attributes names cannot be longer than 10 characters. Consequently, two field names are abbreviated: `pcirco_name` into `pcirco_nam` and `pcirco_type` into `pcirco_typ`.
40. The formal diocesan toponymy is documented in Appendix Table A.19.
41. Communes that disappeared between 1943 and 2005 hold the INSEE code assigned by the first official nomenclature, the COG of 1943. Those that disappeared before 1943 hold the INSEE code of their

- absorbing commune. The variable `com_act` discriminates between these cases.
42. These population counts correspond to the total population (*population totale*), which includes the municipal population (*population municipale*) and the population in collective dwellings (*population comptée à part*). Population counts for municipal arrondissements of Paris are from the *Dénombrements de la Population* volumes, which are available on <https://www.epsilon.insee.fr/jspui/> (accessed in July 2021).
 43. These secondary sources include the HAC database (Motte, Séguay, and Théré 2003; Motte and Vouloir 2007) as well as Marichal (1962), Masson (1984), Farcy (1992), Gaudillère (1995), Chetcuti and Le Noël (1998), Boulanger (2001), Chanet (2006), Ozouf-Marignier and Verdier (2009), Verdier (2009), Marty (2013), Ozouf-Marignier and Verdier (2013), Deloche (2017), and Duquesnoy (2020).
 44. These inaccuracies and errors are documented in Appendix Table A.30.
 45. Useful to interested users, an up-to-date listing of the availability of these periodical statistical publications is available at <https://progedo.hypotheses.org/514> (accessed in July 2021).
 46. For instance, historical cantons maps of 1884 and 1925 have a precision of 1:1,250,000 and 1:1,600,000, respectively.
 47. More precisely, misclassified communes are Plougar (Plouvézédé instead of Plouescat), Plougourvest (Plouvézédé instead of Landivisiau), Sibiril (Plouvézédé instead of Saint-Pol-de-Léon), Plouvorn (Taulé instead of Plouvézédé), and potentially Guiclan (Saint-Thégonnec instead of Taulé). Moreover, the commune of Ile-de-Batz (canton of Saint-Pol-de-Léon) is not represented by LARHRA's (2011) shapefile.
 48. Constituencies that are not subdivided into pseudo-constituencies are départements, arrondissements, military corps and military divisions (1870–1873), courts of appeal, courts of first instance, penitentiary constituencies, academic constituencies, labor inspections, and dioceses.
 49. More precisely, input the *Tableau récapitulatif des inscrits et des votants par départements* (pp. 93–5) and divide content of the *Abstentions* column by the content of the *Inscrits* column. This volume is available from the web page of the archives of the *Assemblée Nationale* at https://archives.assemblee-nationale.fr/dossiers/guerre1914-1918/docs/tableau_elections_1914.pdf (accessed in July 2021). Errors in the original source are documented in Panel A of Appendix Table A.31.
 50. Constituencies that are subdivided into pseudo-constituencies are cantons, electoral circonscriptions, and military regions and subdivisions (1874–1940).
 51. More precisely, input the *Inscrits* and *Abstentions* columns from the *Tableau récapitulatif des inscrits et des votants par départements* (pp. 17–89) and divide content of the *Abstentions* column by the content of the *Inscrits* column. Errors in the original source are documented in Panel B of Appendix Table A.31.
 52. For instance, the circonscription of Montluçon-1 had 21,015 registered voters and 4,866 abstentions (abstention rate of 23.2%), and the circonscription of Montluçon-2 had 21,799 registered voters and 5,455 abstentions (abstention rate of 25.0%). The pseudo-circonscription of Montluçon (1–2), which only concerns the commune of Montluçon (and hardly visible on a map with regional extent), can therefore be attributed an abstention rate of 24.1%.
 53. The ongoing project on this aspect, COMMUNE HIS-DBD (<https://anrcommunes.hypotheses.org/>), is scheduled to be completed by 2023.

Disclosure statement

The author declares that he has no relevant or material financial interests that relate to the research described in this article.

ORCID

Victor Gay  <http://orcid.org/0000-0001-9912-3841>

Data availability statement

The data and reproduction material that support the findings of this study are openly available in the Harvard Dataverse at <https://dataverse.harvard.edu/dataverse/TRF-GIS>.

Funding

This work was supported by the Agence Nationale de la Recherche (ANR) under Grant ANR-17-EURE-0010 (Investissements d'Avenir program).

References

- Boulanger, P. 2001. *La France Devant la Conscription. Géographie Historique d'une Institution Républicaine, 1914–1922*. Paris: Economica.
- Chanet, J.-F. 2006. *Vers l'Armée Nouvelle: République Conservatrice et Réforme Militaire, 1871–1879*. Rennes: Presses Universitaires de Rennes.
- Chetcuti, C., and M. Le Noël. 1998. Évolution et Organisation de l'Administration Centrale du Ministère du Travail de 1887 à 1940. In *Études et Documents pour Servir à l'Histoire de l'Administration du Travail*, edited by Claude Chetcuti and Marcel Le Noël, 1–138. Paris: Ministère de l'Emploi et de la Solidarité.
- Christensen, G., and E. Miguel. 2018. Transparency, Reproducibility, and the Credibility of Economics Research. *Journal of Economic Literature* 56 (3):920–80. doi: 10.1257/jel.20171350.
- Deloche, E. 2017. *Le Diocèse d'Annecy de la Séparation à Vatican II (1905–1962)*. Annecy: Académie Salésienne.
- Desrosières, A. 1998 [1993]. *The Politics of Large Numbers. A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.

- Duquesnoy, J.-P. 2020. *Atlas Historique des Diocèses en France*. Paris: Archives & Culture.
- Farcy, J. C. 1992. *Guide des Archives Judiciaires et Pénitentiaires: 1800–1958*. Paris: CNRS éditions.
- Fitch, C. A., and S. Ruggles. 2003. Building the National Historical Geographic Information System. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 36 (1):41–51. doi: [10.1080/01615440309601214](https://doi.org/10.1080/01615440309601214).
- Gaudillère, B. 1995. *Atlas Historique des Circonscriptions Électorales Françaises*. Genève: Librairie Droz.
- Gros, D. 2014. *Naissance de la Troisième République*. Paris: Presses Universitaires de France.
- Lang, G. 2016. “Le Code Officiel Géographique (COG), Avant, Pendant et Autour.” INSEE Note.
- LARHRA. 2011. *Les Cantons Français de 1884 à 1966*. Lyon: Laboratoire de Recherche Historique Rhône-Alpes.
- Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles. 2020. *IPUMS National Historical Geographic Information System: Version 15.0 [dataset]*. Minneapolis, MN: IPUMS.
- Marichal, P. 1962. F/17. Instruction Publique. In *État Sommaire des Versements Faits aux Archives Nationales par les Ministères et les Administrations qui en Dépendent (Série F–F/10 à F/17)*, edited by Direction des Archives, 742–1083. Paris: Ministère de l’Instruction Publique et des Beaux-Arts.
- Marty, T. 2013. De l’Espace Électoral à son Découpage: Circonscriptions Uninominales et Circonscriptions de Listes sous la Troisième République Française. *Noroi* 226 (226):91–106. doi: [10.4000/noroi.4606](https://doi.org/10.4000/noroi.4606).
- Masson, J.-L. 1984. *Provinces, Départements, Régions. L’Organisation Administrative de la France d’Hier à Demain*. Paris: Fernand Lanore.
- Mimeur, C., F. Queyroi, A. Banos, and T. Thévenin. 2018. Revisiting the Structuring Effect of Transportation Infrastructure: An Empirical Approach with the French Railway Network from 1860 to 1910. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 51 (2):65–81. doi: [10.1080/01615440.2017.1393358](https://doi.org/10.1080/01615440.2017.1393358).
- Motte, C., and M.-C. Vouloir. 2007. Le Site cassini.ehess.fr: Un Instrument d’Observation pour une Analyse du Peuplement. *Bulletin du Comité Français de Cartographie* 191:68–84.
- Motte, C., I. Ségué, and C. Théré. 2003. *Communes d’Hier, Communes d’Aujourd’hui. Les Communes de la France Métropolitaine, 1801–2001*. Dictionnaire d’Histoire Administrative. Paris: INED.
- Noble, P., D. Van Riper, S. Ruggles, J. Schroeder, and M. Hindman. 2011. Harmonizing Disparate Data across Time and Place: The Integrated Spatio-Temporal Aggregate Data Series. *Historical Methods* 44 (2):79–85. doi: [10.1080/01615440.2011.563228](https://doi.org/10.1080/01615440.2011.563228).
- Noizet, H., B. Bove, and L. Costa. 2013. *Paris de Parcelles en Pixels. Analyse Géomatique de l’Espace Parisien Médiéval et Moderne*. Paris: Presses Universitaires de Vincennes–Comité d’Histoire de la Ville de Paris.
- Ostafin, K., D. Kaim, T. Siwek, and A. Miklar. 2020. Historical Dataset of Administrative Units with Social-Economic Attributes for Austrian Silesia 1837–1910. *Scientific Data* 7 (1):1–14. doi: [10.1038/s41597-020-0546-z](https://doi.org/10.1038/s41597-020-0546-z).
- Ozouf-Marignier, M.-V. 1989. *La Formation des Départements. La Représentation du Territoire Français à la Fin du 18ème Siècle*. Paris: Éditions de l’École des Hautes Etudes en Sciences Sociales.
- Ozouf-Marignier, M.-V., and N. Verdier. 2009. Le Canton d’Hier à Aujourd’hui: Étude Cartographique d’un Maillage. In *Le Canton, un Territoire du Quotidien*, edited by Yann Lagadec, Jean Le Bihan, and Jean-François Tanguy, 281–95. Rennes: Presses Universitaires de Rennes.
- Ozouf-Marignier, M.-V., and N. Verdier. 2013. Les Mutations des Circonscriptions Territoriales Françaises. Crise ou Mutation? *Mélanges de L’École Française de Rome–Italie et Méditerranée Modernes et Contemporaines* 125 (2):271–84.
- Perret, J., M. Gribaudo, and M. Barthelemy. 2015. Roads and Cities of 18th Century France. *Scientific Data* 2: 150048. doi: [10.1038/sdata.2015.48](https://doi.org/10.1038/sdata.2015.48).
- Ruggles, S. 2018. The Importance of Data Curation. In *The Palgrave Handbook of Survey Research*, edited by David L. Vannette and Jon A. Krosnick, 303–8. London: Palgrave Macmillan.
- Ruggles, S. 2021. The Revival of Quantification: Reflections on Old New Histories. *Social Science History* 45 (1):1–25. doi: [10.1017/ssh.2020.44](https://doi.org/10.1017/ssh.2020.44).
- Sorrel, C. 2013. *Les “Matériaux Boulard” Trente Ans Après. Des Chiffres et des Cartes: Approches Sérielles et Spatiales en Histoire Religieuse*. Lyon: Laboratoire de Recherche Historique Rhône-Alpes.
- Southall, H. 2011. Rebuilding the Great Britain Historical GIS, Part 1: Building an Indefinitely Scalable Statistical Database. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 44 (3):149–59. doi: [10.1080/01615440.2011.589774](https://doi.org/10.1080/01615440.2011.589774).
- Southall, H. 2012. Rebuilding the Great Britain Historical GIS, Part 2: A Geo-Spatial Ontology of Administrative Units. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 45 (3):119–34. doi: [10.1080/01615440.2012.664101](https://doi.org/10.1080/01615440.2012.664101).
- Southall, H. 2014. Rebuilding the Great Britain Historical GIS, Part 3: Integrating Qualitative Content for a Sense of Place. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 47 (1):31–44. doi: [10.1080/01615440.2013.847774](https://doi.org/10.1080/01615440.2013.847774).
- Thévenin, T., R. Schwartz, and L. Sapet. 2013. Mapping the Distortions in Time and Space: The French Railway Network 1830–1930. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 46 (3):134–43. doi: [10.1080/01615440.2013.803409](https://doi.org/10.1080/01615440.2013.803409).
- Verdier, N. 2009. La Réforme des Arrondissements de 1926: Un Choix d’Intervention entre Espaces et Territoires. In *Les Territoires de l’Administration: Départir, Décentraliser, Déconcentrer*, edited by Pierre Allorant, 107–22. Orléans: Presses Universitaires d’Orléans.
- Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al. 2016. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data* 3 (1):1–9. doi: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).

Chapter 2

Ongoing Research

In this section, I provide extended abstracts for the following ongoing research projects:

- Building on the First World War:
 - World War I and Female Labor Force Participation. The Case of England and Wales (with Benjamin Milner).
 - Weapons of Mass Production. World War I and the Modernization of the French Economy (with Anne Alonzo and Ronan Tallec).
- Measuring Economic Activity in Early Twentieth-Century France (solo data collection project).
- Other ongoing research:
 - Who Became a Nazi? A Structured Database of the German Denazification Questionnaires, 1945–49 (with Jan Stuckatz, Selina Hofstetter, and Mikkel Dack).

2.1 Building on the First World War

2.1.1 World War I and Female Labor Force Participation. The Case of England and Wales (with Benjamin Milner)

How did the demographic shock of the First World War affect female labor force participation in England and Wales during the postwar period?¹ This question has attracted little attention in the economic history literature because, at the national level, female labor force participation in England and Wales appears to have changed little after

¹This work has been presented at the 2025 CNEH and 2026 EHS conferences.

the war relative to its prewar level (e.g., Hatton and Bailey, 2001, p. 88). National aggregates, however, may conceal substantial local heterogeneity. Boehnke and Gay (2022) show this in the French case: although aggregate female labor force participation rates remained broadly stable before and after the war, this pattern masked large regional shifts, with female labor force participation rates increasing markedly in regions that experienced higher military death rates. They argue that this effect was driven by a tightening of the marriage market for young single women and by the negative income shocks experienced by older widows. In this paper, we examine the case of England and Wales to assess the external validity of these findings.

England and Wales offer a useful comparative perspective on the French case in several respects. First, unlike France, the First World War was not fought on British soil, thereby removing concerns that infrastructure destruction or civilian fatalities might affect the results. Second, the composition of the female workforce differed markedly from that in France: whereas more than 50 percent of working women in France were employed in agriculture in 1911, the corresponding share in England and Wales was less than 8 percent. Finally, France's meagre pensions for war widows contrasted sharply with Britain's more generous policy, which by 1920 provided benefits roughly equivalent to the average earnings of a working woman.

By linking individual-level soldier death records to the 1911 census, we estimate the share of prewar male residents who died in the war for each of the nearly two thousand local government districts of England and Wales. We then combine these estimates with the newly released full-count 1921 census data (Schürer and Wakelam, 2025) to implement an individual-level difference-in-differences analysis of female labor force participation. As in France, we find that higher military death rates significantly increased postwar female labor force participation at the local level, with most of the effect driven by rising participation in previously male-dominated occupations. The underlying mechanisms, however, differ sharply across the two countries. In France, the share of widows in the labor force was strongly positively associated with military death rates and accounted for nearly half of the overall estimated effect on female labor force participation. In England and Wales, by contrast, widows' labor force participation was significantly negatively associated with military death rates, and the vast majority of the overall effect was instead driven by never-married women entering the labor force. These findings support the view that the generosity of war widows' pensions played an important role in shaping the contrasting patterns observed in the two countries.

2.1.2 Weapons of Mass Production. World War I and the Modernization of the French Economy (with Anne Alonzo and Ronan Tallec)

Did World War I accelerate the modernization of the French economy?² During the war, annual government spending in France rose to unprecedented levels, reaching nearly 50 percent of prewar GDP. A substantial portion of this spending was devoted to producing war matériel at a time when the country's industrial heartland in the northeast was unavailable for production. In this paper, we ask whether this massive but temporary industrial mobilization durably transformed the structure of the French economy.

Our analysis draws on two original datasets. First, we reconstruct the intensity of wartime industrial mobilization at the local level using archival information on 211,760 procurement contracts awarded to 10,368 private firms between 1914 and 1919 for the production of war matériel. Second, we build a département-level panel of economic activity at the three-digit sector level from the statistical returns of the population censuses spanning 1896 to 1936. We then implement a difference-in-differences strategy, comparing changes in manufacturing outcomes before and after the war across départements that were differentially exposed to wartime industrial mobilization.

We find clear evidence that wartime industrial mobilization fostered a sustained expansion of manufacturing during the interwar period. Départements that received more wartime investment experienced stronger postwar growth in both manufacturing employment and establishments. These effects were not limited to a simple scale expansion: the war also shifted economic activity toward manufacturing, increasing its share in total employment and establishments. In addition, wartime industrial mobilization altered the industrial organization of manufacturing, as establishments in more exposed départements became larger after the war. These départements also experienced an increase in the share of manufacturing workers employed as wage earners, pointing to a more salaried labor structure. By contrast, we find little evidence that the war stimulated capital deepening, suggesting that its medium-run effects operated primarily through labor reallocation and organizational change rather than through postwar capital accumulation.

Our analysis also sheds light on the mechanisms underlying these effects. We show that the modernization induced by industrial warfare was not driven by the expansion of armament production itself. Rather, wartime activity generated substantial spillovers across a wide range of manufacturing subsectors besides armament. More broadly, this transformation is largely explained by a reallocation of labor from agriculture to

²This work has been presented at seminars at Chapman University, Northwestern University, Université Paris-2, Université Paris Dauphine, UC Irvine, the University of Michigan, the University of Toronto, and Yonsei University.

manufacturing within—rather than across—départements. Finally, shifting the analysis to the municipality level, we find evidence for highly localized (within canton) population sorting from rural areas toward nearby manufacturing centers.

Overall, this paper makes three main contributions. First, it provides rare comparative evidence on war-induced industrialization, a topic for which systematic empirical evidence remains limited to the case of the United States during World War II despite many potential historical cases (Jaworski, 2017; Garin and Rothbaum, 2025).³ Second, it speaks to broader debates on place-based industrial policy by analyzing a historical setting in which large, temporary, and spatially exogenous public investments generated persistent effects on local economic structures (Juhász and Steinwender, 2024; Juhász, Lane, and Rodrik, 2024). Finally, it contributes to the literature on the economic consequences of World War I in France by interpreting it directly as an economic shock rather than as a demographic shock (Abramitzky, Delavande, and Vasconcelos, 2011; Boehnke and Gay, 2022; Bergeaud, Chaniot, and Malgouyres, 2025). Taken together, our findings suggest that industrial warfare was paradoxically a catalyst for the structural transformation of the French economy.

2.2 Measuring Economic Activity in Early Twentieth-Century France

Drawing on my extensive work on Third Republic France, I contend that progress in the field is constrained by the scarcity of spatially localized data on economic activity and structure.⁴ Although the production of administrative statistics expanded markedly at the turn of the twentieth century, the development of industrial statistics remained severely constrained by the institutional weakness of the Statistique Générale de la France (SGF) and by its repeated failure to conduct a proper industrial census (Huber, 1937; Touchelay, 1994; 2000; 2022; Desrosières, 1998 [1993]; Volle 1987).⁵

In response to this gap, several generations of scholars attempted to reconstruct retrospective measures of French economic activity. A first generation of statisticians at the SGF constructed national industrial indices from scattered administrative publications (Dugé de Bernonville, 1918; Dessirier, 1924; Sauvy, 1936), but these measures remained partial and approximate (Sauvy, 1941; Chossat, 1987; Toutain, 1996; Harnois, 2003). A second generation—the French school of Quantitative Economic History—broadened the basis of evidence and refined statistical reconstruction methods (e.g., Markovitch, 1965)

³See Bo, Liu, and Zhou (2023) and Mitrunen (2025) for rare exceptions.

⁴I have submitted this project to the 2025 ERC Starting Grant competition. While it passed Step 1, it received a “B” at Step 2. I will resubmit a revised application in the next call.

⁵The industrial survey conducted alongside the 1931 population census was itself widely regarded as flawed and incomplete (Statistique Générale de France, 1935; Huber, 1937; Markovitch, 1965; Volle, 1987).

yet their indices remained national in scope and limited to the few sectors for which administrative statistics were collected. More recent efforts have produced département-level benchmarks of value added by broad sector—most notably Combes et al. (2011)—but these estimates rest heavily on the problematic 1931 industrial survey and remain too aggregated to capture local structural change.⁶

This data-collection project addresses this major data bottleneck by constructing new measures of the local structure of the French economy for the period 1896–1936. As a first approach, I will exploit the nominative lists of the population censuses. While published census returns have long served as a valuable source for studying the French economy (Hau, 1987; Cahen, 1953; Sauvy, 1955; Lucchi, 1956; Vincent, 1965; Toutain, 1963; Marchand and Thélot, 1991), their use has remained limited by their aggregation at the département level. By contrast, the nominative lists contain individual-level information on names, birthplaces, household position, and occupations, and will soon become available at scale through the SOCFACE project. Building on this effort, I will normalize and curate the occupational data for the 1896–1936 censuses—as I am currently doing for the population censuses of Paris for 1926–36 (Brée and Gay, 2026). These data will make it possible to reconstruct the local occupational structure of France at very fine spatial levels.

The second—and more original—approach will develop an original data source based on firm-level archival tax records available at départemental archives throughout France. It will rely on the nominative lists of the *patente* for the prewar period and on establishment-level files from the tax on industrial and commercial profits for the postwar period (Koepke, 1980; Tristram, 2018; Delalande, 2011; Touchelay, 2011). These records contain information on firms' names, addresses, sectors of activity, and taxable values, and sometimes also on employment, wages, and legal status. They therefore provide a direct measure of firms' economic activity and final output, rather than workers' positions in employment. When completed, this new data will dramatically expand the possibilities for empirical research on the French economy during the Third Republic, making it possible to recover the local structure of economic activity at a level of spatial and sectoral detail that has so far been unavailable.

⁶Other proxies, based on wages, labor shares, fiscal data, or household expenditures, are similarly limited by coarse spatial granularity, sparse temporal coverage, or their inability to recover the sectoral structure of local economies (Caruana-Galizia, 2013; Rosés and Sanchis, 2019; Díez-Minguella and Sanchis Llopis, 2019; Lévêque, 1983; Bazot, 2014; Monnet, Riva, and Ungaro, 2021; Montalbo, 2022).

2.3 Other Ongoing Research

2.3.1 Who Became a Nazi? A Structured Database of the German Denazification Questionnaires, 1945–49 (with Jan Stuckatz, Selina Hofstetter, and Mikkel Dack)

The objective of the project “Who Became a Nazi?” is to assemble a new structured database of denazification questionnaires (*Fragebögen*) from the U.S. occupation zone in post-war Germany (1945–49) to study the social foundations, determinants, and consequences of support for the Nazi movement.⁷ Although the rise of Nazism is one of the most researched topics in modern history (e.g., Peukert, 1987; Browning, 1992; Mason, 1993; Aly, 2007), quantitative work still faces severe data limitations. Indeed, existing evidence is often aggregate, small-scale, or non-representative, while the rare individual-level sources typically contain only Nazi Party members and therefore lack both rich socioeconomic information and a credible control group of non-members (e.g., Brustein and Falter, 1995; Rass and Rohrkamp, 2009; Satyanath, Voigtländer, and Voth, 2017; Blum and de Bromhead, 2019). Our project seeks to address this gap by exploiting a source that records both Nazi and non-Nazi individuals and provides unusually detailed retrospective information on political behavior, organizational membership, education, occupations, religion, income, and other sociodemographics. This makes it possible to revisit a classic question—what led Germans to support and join the Nazi movement—with a much stronger empirical design than has previously been possible.

The source we use in this project is historically unique. After the Second World War, the Allies and the Soviets used denazification questionnaires to screen job-seeking Germans—approximately 20 million such forms were filled in, with the most systematic use occurring in the U.S. occupation zone. These questionnaires contain personal information, detailed records of membership in more than 50 Nazi organizations, but also voting, educational, occupational, and income histories—a total of 131 entries. Thanks to their richness, they make it possible to observe not only whether an individual joined the NSDAP or affiliated organizations, but also when they joined, what offices they held, how their employment and income changed over time from the late 1920s up to the early 1940s. The resulting data structure is therefore uniquely well suited to studying both selection into the Nazi movement and the material consequences of participation in it.⁸

⁷This is an ongoing data-collection project jointly funded by the ANR and the FRAL (DeNazDB, [ANR-21-FRAL-0005](#), 2022–26).

⁸The database itself is being constructed from archival denazification files preserved at the U.S. National Archives in College Park, Maryland. Our sampling design uses a weighting by county population using the 1950 West German census to ensure nation-wide representativeness. So far, we have sampled 13,810 individuals, of whom 10,267 have surviving *Fragebogen*.

The project has two main analytical ambitions. First, to provide new descriptive evidence on the socio-economic characteristics of Nazi support and members, as opposed to non-members. Second, to generate panel analyses of causal mechanisms through difference-in-differences designs. For instance, we aim at examining whether job loss increased subsequent Nazi membership, or whether Nazi membership in turn increased later income. These questions speak directly to long-standing debates over whether Nazi support was driven primarily by economic grievances, opportunism, careerism, or other factors, and whether participation in the regime generated tangible material rewards for ordinary Germans.

The first empirical results are promising. The time profile of NSDAP entry in our sample tracks well-established historical turning points, including the surge following Hitler's accession to power in 1933, the membership stop (*Aufnahmesperre*) beginning in April 1933, and renewed entry after the restriction was lifted in 1937. We also address a major concern with the retrospective denazification records: misreporting. To validate self-reported NSDAP membership, we compare a random sample of 1,000 individuals against the NSDAP membership files in the Bundesarchiv in Berlin, whose coverage represents about 80 percent of party members. The results suggest that self-reports are reasonably reliable: among self-reported non-members, only 4 percent appear to have lied, while the observed archival discovery rate for self-reported members aligns closely with the expected 80 percent coverage of archival membership files. This validation exercise supports the validity of our database as a source for studying Nazi participation.

Substantively, our project promises to make an important contribution to the literature on the origins and persistence of authoritarian mass support. By combining rich individual-level histories with a non-member comparison group, the project moves beyond the limits of both aggregate electoral studies and membership-only datasets and offers a powerful lens through which to study not only who became a Nazi, but also why—and with what consequences. More broadly, it opens the door to a new generation of research on political extremism, social mobility under dictatorship, and the material foundations of regime support.

Chapter 3

Research Supervision

Since joining the Toulouse School of Economics as an Assistant Professor in 2018, I have supervised 12 M1 theses, 6 M2 theses, and 3 PhD theses. I provide the details below.

3.1 Master Theses

3.1.1 M1 theses

1. Simon Postler. “Forced migration and human capital.” 2024.
2. Thomas Desirée. “State Capacity and the Mitigation of Natural Disaster Consequences.” 2023.
3. Marwan Mezdar. “Famine and Consequences of Households’ Coping Strategies.” 2021.
4. Aabir Moncef. “Assessing the Impact of the Black Death on Demographic Behaviors.” 2021.
5. Matteo Cassoret. “Skewed Sex-Ratio and Female Labor Participation.” 2021.
6. Zhihan Zhang. “Does Gender-Specific Parental Leave Length Improve the Situation of Women’s Labor?” 2021.
7. Andrew Boomer. “Application of Univariate Time Series Techniques to High Frequency Electricity Data.” 2020.
8. Faustine Charron. “The Influence of Scandals on MPs’ Elections in France.” 2020.
9. Moongyeom Kim. “The Impact of Varying Ratio between Solar and Wind on the Balancing Cost.” 2020.

10. Mariya Sakharova. "How Tsarist Agriculture can Elucidate the Relationship between Resources and Institutions." 2020.
11. Jiangyu Wei. "Natural Calamities or Man-made Misfortunes: The Political Economic Causal Analysis of China's Great Famine." 2019.
12. Yifan Zhang. "Religion, Religiosity, and Identity: The Income Gap between Religious Believers and Nonbelievers in China." 2019.

3.1.2 M2 theses

1. Zhihan Zhang. "The Value-Added of Students' Expectations for Future Job Categories on Their Performance." 2022.
2. Omar Feddag. "Democracy, State Capacity and Long-Run Growth." 2021.
3. Leo Portales. "The Role of Institutions for Long-Run Economic Growth." 2020.
4. Olivier Lison. "Demographic Movements in France in the Aftermath of World War 2." 2020.
5. Yifan Zhang. "Ancestor Worship and Fertility Decisions in China." 2020.
6. Guohui Jiang. "Financing Civil War through External Debts: Evidence from Chinese Warlords." 2019.

3.2 Doctoral Theses

1. Capucine Ferré (thesis in progress, external committee member, Paris Dauphine University).
2. Yifan Zhang. "Essays in Economic History and Political Economy." 2025.
3. Luisa Carrer. "Essays on the Economics of Gender." 2024.

Bibliography

- Abad, Leticia Arroyo, and Noel Maurer. 2021. "History Never Really Says Goodbye: A Critical Review of the Persistence Literature." *Journal of Historical Political Economy*, 1(1): 31–68.
- Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos. 2011. "Marrying Up: The Role of Sex Ratio in Assortative Matching." *American Economic Journal: Applied Economics*, 3(3): 124–57.
- Acemoglu, Daron, and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. Cambridge, UK: Cambridge University Press.
- Acemoglu, Daron, David H. Autor, and David Lyle. 2004. "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury." *Journal of Political Economy*, 112(3): 497–551.
- Albertus, Michael, and Victor Gay. 2017. "Unlikely Democrats. Economic Elite Uncertainty under Dictatorship and Support for Democratization." *American Journal of Political Science*, 61(3): 624–41.
- Albertus, Michael, and Victor Gay. 2019. "No Better Time Than Now. Future Uncertainty and Private Investment Under Dictatorship." *Economics & Politics*, 31(1): 71–96.
- Albertus, Michael, and Victor Gay. 2025. "State-Building and Rebellion in the Run-Up to the French Revolution." *American Political Science Review*, forthcoming.
- Alonzo, Anne, Victor Gay, and Ronan Tallec. 2026. "Weapons of Mass Production: World War I and the Modernization of the French Economy." Work in progress.
- Aly, Götz. 2007. *Hitler's Beneficiaries: Plunder, Racial War, and the Nazi Welfare State*. New York, NY: Metropolitan Books.
- Aneja, Abhay, Silvia Farina, and Guo Xu. 2024. "Beyond the War: Public Service and the Transmission of Gender Norms." NBER Working Paper 32639.
- Bazot, Guillaume. 2014. "Interregional Inequalities, Convergence, and Growth in France from 1840 to 1911." *Annals of Economics and Statistics*, 113/114: 309–45.
- Beach, Brian, and Walker W. Hanlon. 2023. "Culture and the Historical Fertility Transition." *The Review of Economic Studies*, 90(4): 1669–1700.

- Becker, Gary S., and Gregg H. Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2): S279–88.
- Bergeaud, Antonin, Jean-Baptiste Chaniot, and Clément Malgouyres. 2025. "Escaping Labor Scarcity: Innovation and Human Capital after WW1 in France." CEPR Discussion Paper 20492.
- Blockmans, Wim, and Jean-Philippe Genet, ed. 1995. *The Origins of the Modern State in Europe, 13th to 18th Centuries*. Strasbourg, France; Oxford, UK: European Science Foundation; Clarendon Press.
- Blockmans, Wim, André Holenstein, and Daniel Shläppi, ed. 2009. *Empowering Interactions: Political Cultures and the Emergence of the State in Europe 1300–1900*. Farnham, UK: Ashgate.
- Blum, Matthias, and Alan de Bromhead. 2019. "Rise and Fall in the Third Reich: Social Advancement and Nazi Membership." *European Economic Review*, 120: 103312.
- Boehnke, Jörn, and Victor Gay. 2022. "The Missing Men. World War I and Female Labor Force Participation." *The Journal of Human Resources*, 57(4): 1209–41.
- Bo, Shiyu, Cong Liu, and Yan Zhou. 2023. "Military Investment and the Rise of Industrial Clusters: Evidence from China's Self-Strengthening Movement." *Journal of Development Economics*, 161: 103015.
- Bourel, Lucas, Nathalie Hernandez, Nathalie Aussenac-Gilles, Victor Gay, and Sébastien Poublanc. 2022. "Knowledge Graphs for Representing and Analyzing the Evolution of Territories in History." In *Actes des journées humanités numériques et Web sémantique*, edited by Nicolas Lasolle, Oliver Bruneau and Jean Liever, 23–37. Nancy, France.
- Brodeur, Abel, Nikolai Cook, and Anthony Heyes. 2020. "Methods Matter: p-Hacking and Publication Bias in Causal Analysis in Economics." *American Economic Review*, 110(11): 3634–60.
- Browning, Christopher R. 1992. *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*. New York, NY: Harper Perennial.
- Brustein, William, and Jürgen W. Falter. 1995. "Who Joined the Nazi Party? Assessing Theories of the Social Origins of Nazism." *Zeitgeschichte*, 22(3/4): 83–108.
- Brée, Sandra, and the POPP Research Collective [11 researchers, including Victor Gay]. 2025. "Paris 100 Years Ago: More People than Today—and Mostly Born Elsewhere." *Population & Sociétés*, 636: 1–4.
- Brée, Sandra, and Victor Gay. 2026. "POPPOcc: Occupations in the Parisian Censuses, 1926–36." Work in progress.
- Brée, Sandra, Victor Gay, Marion Leturcq, Baptiste Coulmont, Yoann Doignon, Thomas Constum, Thierry Paquet, and Pierrick Tranouez. 2026. "POPP. An OCR-Generated

- Database of the Population Censuses of Paris (1926–1936).” *Historical Life Course Studies*, 16: 3–28.
- Cahen, Lucienne. 1953. “Évolution de la population active en France depuis cent ans d’après les dénombrements quinquennaux” [The evolution of the labor force in France over the past hundred years based on the quinquennial censuses]. *Economie et statistique*, 8–9(3): 230–88.
- Carranza, Eliana. 2014. “Soil Endowments, Female Labor Force Participation, and the Demographic Deficit of Women in India.” *American Economic Journal: Applied Economics*, 6(4): 197–225.
- Caruana-Galizia, Paul. 2013. “Estimating French Regional Income: Departmental per Capita Gross Value Added, 1872–1911.” *Research in Economic History*, 29: 71–95.
- Cavaillé, Charlotte, and Victor Gay. 2026. “Caught in the Crossfire: World War I and the Failure of Women’s Suffrage in France.” Work in progress.
- Chen, Keith M. 2013. “The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets.” *American Economic Review*, 103(2): 690–731.
- Chossat, Jean-Pierre. 1987. “L’indice de la production industrielle” [The industrial production index]. In *Pour une histoire de la statistique. Tome 2/Matériaux*, 469–78. Paris, France: Economica/INSEE.
- Collier, David, and Gerardo L. Munck, ed. 2022. *Critical Junctures and Historical Legacies: Insights and Methods for Comparative Social Science*. London, UK: The Rowman Littlefield Publishing Group, Inc.
- Combes, Pierre-Philippe, Miren Lafourcade, Jacques-François Thisse, and Jean-Claude Toutain. 2011. “The Rise and Fall of Spatial Inequalities in France: A Long-Run Perspective.” *Explorations in Economic History*, 48(2): 243–71.
- Davoine, Eva, Joseph Enguehard, Victor Gay, and Igor Kolesnikov. 2026a. “Extraction or Evasion? The Limits of Fiscal Capacity in Early Modern France.” Work in progress.
- Davoine, Eva, Joseph Enguehard, Victor Gay, and Igor Kolesnikov. 2026b. “Mapping the Grandes Gabelles in Early Modern France.” TSE Working Paper 26–1716.
- Davoine, Eva, Joseph Enguehard, Victor Gay, and Igor Kolesnikov. 2026c. *The Grandes Gabelles in 1665 France* (Version 1.0) [Dataset]. Harvard Dataverse. <https://doi.org/10.57745/D039SH>.
- Dazey, Margot, and Victor Gay. 2025. “The Mosque Nearby. Visible Minorities and Far-Right Support in France.” *Comparative Political Studies*, 58(9): 2000–48.

- Delalande, Nicolas. 2011. "Guerre et mobilisation fiscale (1914–1918)" [War and fiscal mobilization (1914–1918)]. In *Les batailles de l'impôt. Consentement et résistances de 1789 à nos jours*, 245–73. Paris, France: Seuil.
- Desrosières, Alain. 1998 [1993]. *The Politics of Large Numbers. A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
- Dessirier, Jean. 1924. "Essai de détermination d'indices mensuels de la production industrielle en France" [An attempt to construct monthly indices of industrial production in France]. *Bulletin de la Statistique générale de la France*, 14(1): 73–101.
- Díez-Minguella, Alfonso, and Teresa Sanchis Llopis. 2019. "Regional Income Inequality in France 1860–1954: Methods and Findings." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 52(1): 1–14.
- Dreber, Anna, Victor Gay, Maia Güell, Joan Llull, and Christophe Pérignon. 2024. "Replication Challenges in Economics and Econometrics." <https://doi.org/10.5281/zenodo.13626807>.
- Dugé de Bernonville, Léopold. 1918. "Essai d'évaluation de la production industrielle en France avant la guerre" [An attempt to estimate industrial production in France before the war]. *Bulletin de la Statistique générale de la France*, 8(1): 71–92.
- Elias, Norbert. 1994 [1939]. *The Civilizing Process*. Oxford, England: Basil Blackwell.
- Fernández, Raquel. 2011. "Does Culture Matter?" In *Handbook of Social Economics*, Volume 1, edited by Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 481–510. North-Holland.
- Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *The Quarterly Journal of Economics*, 119(4): 1249–99.
- Fernández, Raquel, and Alessandra Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics*, 1(1): 146–77.
- Franck, Raphaël, and Victor Gay. 2025. "Urbanization and Electoral Success: Lawyers and Workers in Interwar France." CEPR Discussion Paper 18737.
- Garin, Andrew, and Jonathan Rothbaum. 2025. "The Long-Run Impacts of Public Industrial Investment on Local Development and Economic Mobility: Evidence from World War II." *The Quarterly Journal of Economics*, 140(1): 459–520.
- Gay, Victor. 2020. *The TRF-GIS Dataverse. The Third-Republic France Geographic Information System (1870–1940) [Dataverse]*. Harvard Dataverse. <https://dataverse.harvard.edu/dataverse/TRF-GIS>.

- Gay, Victor. 2021. "Mapping the Third Republic. A Geographic Information System of France (1870–1940)." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 54(4): 189–207.
- Gay, Victor. 2023a. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility. A Verification and Reproduction of Fernández and Fogli (2009)." *Journal of Comments and Replications in Economics*, 2: 1–15.
- Gay, Victor. 2023b. "The Intergenerational Transmission of World War I on Female Labour." *The Economic Journal*, 133(654): 2303–33.
- Gay, Victor. 2025a. "La mesure du travail des femmes dans les recensements de la population au tournant du 20e siècle" [Measuring female labor in France at the turn of the twentieth century]. In *Economistes et historiens, dialogue de sourds?*, edited by Alain Trannoy and Arundhati Virmani, 65–78. Paris, France: Odile Jacob.
- Gay, Victor. 2025b. *The Jean Nicolas Database* (Version 2.1) [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/ANQXMQ>.
- Gay, Victor. 2025c. "The Jean Nicolas Database. The French Rebellion, 1661–1789." *Data & Corpus*, 1: 15892.
- Gay, Victor. 2025d. "Un data paper en SHS: pourquoi, pour qui, comment?" [A data paper in the social sciences: why, for whom, how?]. In *Publier, partager, réutiliser les données de la recherche: les data papers et leurs enjeux.*, edited by Christine Kosmopoulos and Joachim Schöpfel, 66–74. Lille, France: Presses universitaires du Septentrion.
- Gay, Victor, and Benjamin Milner. 2026. "World War I and Female Labor Force Participation: The Case of England and Wales." Work in progress.
- Gay, Victor, and Christophe Lévêque. 2026. "After the Front: Devastation and Structural Transformation in Post-WWI France." Work in progress.
- Gay, Victor, and Lionel Kesztenbaum. 2026. "Collateral Damage: How World War I Changed the Way Women Worked." Work in progress.
- Gay, Victor, and Pauline Grosjean. 2023a. "Morts Pour la France: A Database of French Fatalities of the Great War." *Explorations in Economic History*, 90: 101550.
- Gay, Victor, and Pauline Grosjean. 2023b. *The Morts Pour la France Database* (Version 4.0) [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/2APQZ6>.
- Gay, Victor, and Sébastien Poulanc. 2022. "ObARDI: un projet interdisciplinaire pour comprendre la construction de l'État moderne (1661–1789)" [ObARDI: an interdisciplinary project to understand the construction of the Modern State (1661–1789)]. *Les Cahiers de Framespa*, 41: 1–12.
- Gay, Victor, and Sébastien Poulanc. 2023. "Espaces construits par la puissance publique et espaces vécus par les justiciables dans le royaume de France (1661–1789)" [Constructed

- spaces by public authorities and experienced spaces by litigants in the Kingdom of France (1661–1789)]. In *Actes du congrès de la Fédération historique de la région Occitanie*.
- Gay, Victor, Daniel L. Hicks, and Estefania Santacreu-Vasut. 2016a. "Language and Gender Roles among Immigrants to the U.S. A Historical Perspective." In *I Mondi Delle Donne*, edited by Paola Paolini, 165–79. Rome, Italy: Edi Cusano.
- Gay, Victor, Daniel L. Hicks, and Estefania Santacreu-Vasut. 2016b. "Migration As A Window Into The Coevolution Between Language And Behavior." In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG 11)*, edited by Seán Roberts, Christine Cuskley, Luke McCrohon, Lluís Barceló-Coblijn, Olga Feher and Tessa Verhoef, 1–8.
- Gay, Victor, Daniel L. Hicks, Estefania Santacreu-Vasut, and Amir Shoham. 2017. *Replication Data for: Decomposing Culture: An Analysis of Gender, Language, and Labor Supply in the Household* (Version 2.0) [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/XIFPS5>.
- Gay, Victor, Daniel L. Hicks, Estefania Santacreu-Vasut, and Amir Shoham. 2018. "Decomposing Culture. An Analysis of Gender, Language, and Labor Supply in the Household." *Review of Economics of the Household*, 16(4): 879–909.
- Gay, Victor, Estefania Santacreu-Vasut, and Amir Shoham. 2013a. "Do Female/Male Distinctions in Language Matter? Evidence from Gender Political Quotas." *Applied Economics Letters*, 20(5): 495–8.
- Gay, Victor, Estefania Santacreu-Vasut, and Amir Shoham. 2013b. "The Grammatical Origins of Gender Roles." Berkeley Economic History Laboratory Working Paper 2013–03.
- Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2023a. *Bailliages in 1789 France* (Version 2.0) [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/T8UXHK>.
- Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2023b. *Customs in Ancien Régime France* (Version 2.0) [Dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/HYE209>.
- Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2024a. "The Atlas of Local Jurisdictions of Ancien Régime France." *Journal of Historical Geography*, 84: 49–60.
- Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2024b. "The Customary Atlas of Ancien Régime France." *Explorations in Economic History*, 93: 101588.
- Gay, Victor, Paula E. Gobbi, and Marc Goñi. 2025. "Revolutionary Transition: Inheritance Change and Fertility Decline." *Journal of Political Economy*, forthcoming.
- Goldin, Claudia. 2006. "The Quiet Revolution That Transformed Women's Employment, Education, and Family." *American Economic Review*, 96(2): 1–21.
- Grosjean, Pauline, and Rose Khattar. 2019. "It's Raining Men! Hallelujah? The Long-Run Consequences of Male-Biased Sex Ratios." *The Review of Economic Studies*, 86(2): 723–54.

- Guinnane, Timothy W. 2011. "The Historical Fertility Transition: A Guide for Economists." *Journal of Economic Literature*, 49(3): 589–614.
- Harnois, Jérôme. 2003. "Historique des indices français de la production industrielle" [History of French industrial production indices]. *INSEE Méthodes*, 104: 49–65.
- Hatton, Timothy, J., and Roy E. Bailey. 2001. "Women's Work in Census and Survey, 1911–1931." *The Economic History Review*, 54(1): 87–107.
- Hau, Michel. 1987. "L'étude de la croissance régionale de 1810 à nos jours: données et lacunes" [The study of regional growth from 1810 to the present: data and gaps]. In *Pour une histoire de la statistique. Tome 1. Contributions*, 325–67. Paris, France: Economica/INSEE.
- Hicks, Daniel L. 2013. "War and the Political Zeitgeist: Evidence from the History of Female Suffrage." *European Journal of Political Economy*, 31: 60–81.
- Huber, Michel. 1937. "Quarante années de la Statistique générale de la France, 1896–1936." *Journal de la Société statistique de Paris*, 78: 179–214.
- Huntington-Klein, Nick, Claus C. Pörtner, and the Many-Economists Research Collective [146 researchers, including Victor Gay]. 2025. "The Sources of Researcher Variation in Economics." NBER Working Paper 33729.
- Jaworski, Taylor. 2017. "World War II and the Industrialization of the American South." *The Journal of Economic History*, 77(4): 1048–82.
- Juhász, Réka, and Claudia Steinwender. 2024. "Industrial Policy and the Great Divergence." *Annual Review of Economics*, 16: 27–54.
- Juhász, Réka, Nathan Lane, and Dani Rodrik. 2024. "The New Economics of Industrial Policy." *Annual Review of Economics*, 16: 213–42.
- Koepke, Robert L. 1980. "The Loi des patentes of 1844." *French Historical Studies*, 11(3): 398–400.
- le Bris, David, and Victor Gay. 2023. "Distance to Innovations, Kinship Intensity, and Psychological Traits." *PLoS One*, 18(1): e0279864.
- Le Play, Frédéric. 1875. *L'organisation de la famille selon le vrai modèle signalé par l'histoire de toutes les races et de tous les temps* [Family organization according to the true model revealed by the history of all peoples across all ages]. Tours, France: A. Mame et Fils.
- Lévêque, Pierre. 1983. "La patente, indicateur de croissance économique différentielle au XIX^{ème} siècle ?" [The business tax as an indicator of differential economic growth in the nineteenth century?]. In *Entreprise et entrepreneurs (XIX–XX^è siècles)*, 47–72. Paris, France: Presses de l'Université de Paris Sorbonne.
- Lucchi, Michel. 1956. "Évolution de la population active entre 1906 et 1954" [The evolution of the labor force between 1906 and 1954]. *Etudes statistiques*, 3.

- Maniadis, Zacharias, Fabio Tufano, and John A. List. 2017. "To Replicate or Not to Replicate? Exploring Reproducibility in Economics through the Lens of a Model and a Pilot Study." *The Economic Journal*, 127(605): F209–35.
- Mann, Michael. 1984. "The Autonomous Power of the State." *European Journal of Sociology*, 25(2): 185–213.
- Marchand, Olivier, and Claude Thélot. 1991. *Deux siècles de travail en France* [Two centuries of work in France]. Paris, France: INSEE.
- Markovitch, Tihomir J. 1965. "L'industrie française de 1789 à 1964" [French industry from 1789 to 1964]. *Cahiers de l'I.S.E.A.*, AF 4.
- Mason, Tim. 1993. *Social Policy in the Third Reich: The Working Class and the 'National Community'*. Oxford, UK: Berg Publishers.
- Miske, Olivia, Brian A. Nosek, Timothy L. Errington, and the SCORE Research Collective [124 researchers, including Victor Gay]. 2026. "Investigating the Reproducibility of the Social and Behavioral Sciences." *Nature*, forthcoming.
- Mitrunen, Matti. 2025. "War Reparations, Structural Change, and Intergenerational Mobility." *The Quarterly Journal of Economics*, 140(1): 521–84.
- Monnet, Eric, Angelo Riva, and Stefano Ungaro. 2021. "The Real Effects of Bank Runs. Evidence from the French Great Depression (1930–1931)." CEPR Discussion Paper 16054.
- Montalbo, Adrien. 2022. "Primary Education and Economic Growth in Nineteenth-Century France." *Cliometrica*, 16: 277–332.
- Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics*, 8: 405–34.
- Pengl, Yannick I., Carl Müller-Crepon, Roberto Valli, Lars-Erik Cederman, and Luc Girardin. 2026. "The Train Wrecks of Modernization: Railway Construction and Separatist Mobilization in Europe." *American Political Science Review*, 120(1): 37–54.
- Peukert, Detlev J. K. 1987. *Inside Nazi Germany: Conformity, Opposition, and Racism in Everyday Life*. New Haven, CT: Yale University Press.
- Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance." *The Quarterly Journal of Economics*, 123(3): 1251–85.
- Rass, Christoph, and René Rohrkamp. 2009. *Deutsche Soldaten 1939–1945: Handbuch einer biographischen Datenbank zu Mannschaften und Unteroffizieren von Heer, Luftwaffe und Waffen-SS* [German Soldiers, 1939–1945: Handbook of a Biographical Database on Enlisted Men and Non-Commissioned Officers of the Army, Luftwaffe, and Waffen-SS]. Aachen, Germany: RWTH Aachen. Ergänzte Ausgabe.

- Roberts, Séan G., James Winters, and Keith M. Chen. 2015. "Future Tense and Economic Decisions: Controlling for Cultural Evolution." *PLoS One*, 10(7): e0132145.
- Rosés, Joan R., and Teresa Sanchis. 2019. "A Long-Run Perspective on French Regional Income Inequality, 1860–2010." In *The Economic Development of Europe's Regions. A Quantitative History since 1900*, edited by Joan R. Rosés and Nikolaus Wolf, 129–48. New York, NY: Routledge.
- Satyanath, Shanker, Nico Voigtländer, and Hans-Joachim Voth. 2017. "Bowling for Fascism: Social Capital and the Rise of the Nazi Party." *Journal of Political Economy*, 125(2): 478–526.
- Sauvy, Alfred. 1936. "Indices comparés de la production industrielle et de la production agricole en divers pays de 1870 à 1928" [Comparative indices of industrial production and agricultural production in various countries from 1870 to 1928]. *Bulletin de la Statistique générale de la France*, 25(3): 453–93.
- Sauvy, Alfred. 1941. "Les statistiques industrielles" [Industrial statistics]. *Journal de la Société statistique de Paris*, 82: 131–45.
- Sauvy, Alfred. 1955. "Les tendances de la population active en France" [Trends in the labor force in France]. *Population*, 10(3): 413–26.
- Schulz, Jonathan F, Duman Bahrami-Rad, Jonathan P. Beauchamp, and Joseph Henrich. 2019. "The Church, Intensive Kinship, and Global Psychological Variation." *Science*, 366(6466): eaau5141.
- Schürer, Kevin, and Alexander Wakelam. 2025. *Integrated Census Microdata (I-CeM), England and Wales, 1921* (Version 1.0). SN: 9280 [data collection]. UK Data Service. <http://doi.org/10.5255/UKDA-SN-9280-1>.
- Statistique Générale de France. 1935. *Enquêtes annexes du recensement de 1931. Enquête industrielle* [Supplementary surveys to the 1931 census: industrial survey]. Paris, France: Imprimerie nationale.
- Stuckatz, Jan, Selina Hofstetter, Victor Gay, and Mikkel Dack. 2026. "Who Became a Nazi? A Structured Database of the German Denazification Questionnaires, 1945–49." Work in progress.
- Teso, Edoardo. 2019. "The Long-Term Effect of Demographic Shocks on the Evolution of Gender Roles: Evidence from the transatlantic Slave Trade." *Journal of the European Economic Association*, 17(2): 497–534.
- Thébaud, Françoise. 1986. *Les femmes au temps de la guerre de 14* [Women during the First World War]. Paris, France: Stock.
- Touchelay, Béatrice. 1994. "Le pari de la création de l'INSEE en avril 1946. Concilier deux traditions distinctes" [The challenge of creating INSEE in April 1946: reconciling

- two distinct traditions]. *Études et documents du Comité pour l'histoire économique et financière de la France*, 6: 405–28.
- Touchelay, Béatrice. 2000. "L'INSEE, histoire d'une institution" [The age of numbers: statistical systems and national traditions]. In *L'ère du chiffre. Systèmes statistiques et traditions nationales*, edited by Jean-Pierre Beaud and Jean-Guy Prévost, 153–87. Montreal, Canada: Presses de l'Université du Québec à Montréal.
- Touchelay, Béatrice. 2022. *L'aventure statistique. Moderniser la France et le monde*. Rennes, France: Presses universitaires de Rennes.
- Touchelay, Béatrice. 2011. "La contribution exceptionnelle sur les bénéfices de guerre et ses effets immédiats (1916–1919)" [The exceptional tax on wartime profits and its immediate effects (1916–1919)]. In *L'Etat et l'entreprise. Une histoire de la normalisation comptable et fiscale à la française*, 35–64. Rennes, France: Presses universitaires de Rennes.
- Toutain, Jean-Claude. 1963. "La population de la France de 1700 à 1959" [The population of France from 1700 to 1959]. *Cahiers de l'I.S.E.A.*, AF 3.
- Toutain, Jean-Claude. 1996. "French Economic Growth, 1815–1938: Some Remarks on the Different Estimates of the Gross Domestic Product." *Revue économique*, 47(4): 893–919.
- Tripp, Aili Mari. 2023. "War, Revolution, and the Expansion of Women's Political Representation." *Politics & Gender*, 19(3): 922–27.
- Tristram, Frédéric. 2018. "La Première Guerre mondiale et la rénovation du système fiscal français" [The First World War and the reform of the French tax system]. In *Une fiscalité de guerre?*, edited by Florence Descamps and Laure Quennouëlle-Corre, 189–210. Paris, France: Institut de la gestion publique et du développement économique.
- Vincent, Louis-André. 1965. "Population active, production et productivité dans 21 branches de l'économie française (1896–1962)" [Labor force, production, and productivity in 21 branches of the French economy (1896–1962)]. *Etudes et conjoncture*, 20(2): 73–108.
- Volle, Michel. 1987. "Naissance de la statistique industrielle en France (1930–1950)" [The emergence of industrial statistics in France (1930–1950)]. In *Pour une histoire de la statistique. Tome 1/Contributions*, 327–65. Paris, France: Economica/INSEE.
- Waldinger, Maria. 2024. "'Let Them Eat Cake': Drought, Peasant Uprisings, and Demand for Institutional Change in the French Revolution." *Journal of Economic Growth*, 29(1): 41–77.
- Weber, Max. 1968. *Economy and Society*. New York, NY: Bedminster Press.

Appendix

A Doctoral Research on the First World War

This section provides the following analysis and data articles:

- Analysis article 1: Boehnke, Jörn and Victor Gay. 2022. “The Missing Men. World War I and Female Labor Force Participation.” *The Journal of Human Resources*, 57(4): 1209–41.¹
- Analysis article 2: Gay, Victor. 2023b. “The Intergenerational Transmission of World War I on Female Labour.” *The Economic Journal*, 133(654): 2303–33.²
- Data article 1: Gay, Victor. 2021. “Mapping the Third Republic. A Geographic Information System of France (1870–1940).” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 54(4): 189–207.³
- Data article 2: Gay, Victor and Pauline Grosjean. 2023a. “Morts Pour la France: A Database of French Fatalities of the Great War.” *Explorations in Economic History*, 90: 101550.⁴

¹The appendix to this article is available at <https://doi.org/10.3368/jhr.57.4.0419-10151R1>. The replication material for this article is available at <https://doi.org/10.7910/DVN/AP1HZ8>.

²The appendix to this article is available at <https://doi.org/10.1093/ej/uead029>. The replication material for this article is available at <https://doi.org/10.5281/zenodo.7762881>.

³The appendix to this article is available at <https://doi.org/10.1080/01615440.2021.1937421>. The dataverse associated with this article is available at <https://dataverse.harvard.edu/dataverse/TRF-GIS>.

⁴The appendix to this article is available at <https://doi.org/10.1016/j.eeh.2023.101550>. The dataset associated with this article is available at <https://doi.org/10.7910/DVN/2APQZ6>.

The Missing Men

World War I and Female Labor Force Participation

Jörn Boehnke
Victor Gay

ABSTRACT

Using spatial variation in World War I military fatalities in France, we show that the scarcity of men due to the war generated an upward shift in female labor force participation that persisted throughout the interwar period. Available data suggest that increased female labor supply accounts for this result. In particular, deteriorated marriage market conditions for single women and negative income shocks to war widows induced many of these women to enter the labor force after the war. In contrast, demand factors such as substitution toward female labor to compensate for the scarcity of male labor were of second-order importance.

Jörn Boehnke is an assistant professor in the Graduate School of Management at the University of California, Davis, CA and an associate research scholar in the Center of Mathematical Sciences and Applications at Harvard University, Cambridge, MA. Victor Gay is an assistant professor of economics at the Toulouse School of Economics and a research fellow at the Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, France. They thank Ran Abramitzky, Gani Aldashev, Pierre André, Aaron Bodoh-Creed, Nick Crawford, Richard Freeman, Cecilia Garcia Peñalosa, Claudia Goldin, Richard Hornbeck, Lionel Kesztenbaum, Scott Kominers, Steven Levitt, Gabriel Mesevage, Sreemati Mitter, Carolyn Moehling, Natalya Naumenko, Derek Neal, Maëlys de la Rupelle, Mohamed Saleh, Estefania Santacreu-Vasut, Mara Squicciarini, Françoise Thébaud, Alain Trannoy, Alessandra Voena, Arundhati Virmani, and Eugene White for fruitful discussions. They gratefully acknowledge funding from the Center of Mathematical Sciences and Applications at Harvard University, from the Social Science Division at the University of Chicago, and from the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). The authors declare that they have no relevant or material financial interests that relate to the research described in this article. The replication material for this article can be found at <https://doi.org/10.7910/DVN/AP1HZ8>.

[Submitted April 2019; accepted May 2020]; doi:10.3368/jhr.57.4.0419-10151R1

JEL Classification: J12, J16, J22, and N34.

ISSN 0022-166X E-ISSN 1548-8004 © 2022 by the Board of Regents of the University of Wisconsin System

 Color versions of some graphs in this article are available through online subscription at: <http://jhr.uwpress.org>

 Supplementary materials are freely available online at: <http://uwpress.wisc.edu/journals/journals/jhr-supplementary.html>

 This open access article is distributed under the terms of the CC-BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0>) and is freely available online at: <http://jhr.uwpress.org>.

Victor Gay <https://orcid.org/0000-0001-9912-3841>

The major fact will be a breakdown of the equilibrium between sexes. There will not be enough suitors for all young women searching for a husband. [...] The prospect of remaining single will induce most young women to worry about getting an occupation to make their living and to be self-sufficient.

—Girault (1915), p. 443–4

I. Introduction

The dramatic rise of female labor force participation has been one of the most significant changes to labor markets in the past century (Olivetti and Petrongolo 2016). This “quiet revolution” has been mostly interpreted as a consequence of long-run trends in technological change (Goldin 2006).¹ Less understood is the role of idiosyncratic historical events, such as shocks to the adult sex ratio, which can at times stall or propel the secular march toward gender equality by disrupting labor market conditions both in the short and the long run (Angrist 2002; Grosjean and Khattar 2019; Teso 2019). Indeed, implications of such imbalances and the mechanisms through which they materialize are often challenging to identify because they typically result from factors that gradually shape labor market structures as well (Qian 2008; Carranza 2014).

In this article, we provide evidence that jolts of history can generate rapid and long-lasting changes to women’s involvement in the economy. We overcome identification issues by using World War I (WWI) in France as a severe exogenous shock to the adult sex ratio and show that it generated a persistent upward shift in female labor force participation during the interwar period. While WWI ravaged continental Europe between 1914 and 1918, France suffered an especially high death toll relative to other belligerent countries. Because of a universal conscription system, most French male citizens were drafted throughout the war. Out of ten million men aged 15–50 before the war, eight million were drafted in the army; 1.3 million died in combat—a military death rate of 16 percent. As a result, the sex ratio among adults aged 15–50 dropped from 98 men per 100 women at the onset the war to 88 by the end of the war. It was not until after World War II (WWII) that the sex ratio reverted back to balance (Figure 1).

Our empirical strategy exploits differential changes in female labor force participation rates before and after the war across départements that experienced different military death rates.² Key to our contribution, we collected the individual military records of these 1.3 million missing men to build a precise measure of military death rates at the département level. While the relationship between military death rates and changes in female labor force participation rates was flat between 1901 and 1911, it exhibits a positive slope of 0.4 between 1911 and 1921 (Figure 2). Difference-in-differences estimates confirm this relationship. In départements that experienced military death rates of 20 percent rather than 10 percent—equivalent to switching from the 25th to the

1. For instance, the increasing availability of household appliances liberated women’s time for market work (Greenwood, Seshadri, and Yorukoglu 2005). Alternatively, the structural transformation increased the supply of service sector jobs, in which women have a comparative advantage (Ngai and Petrongolo 2017).

2. Départements constitute the second level of France’s administrative subdivisions, between regions and arrondissements. They are broadly comparable to English and U.S. counties, German Landkreises, and Spanish provincias. There were 87 départements before the war.



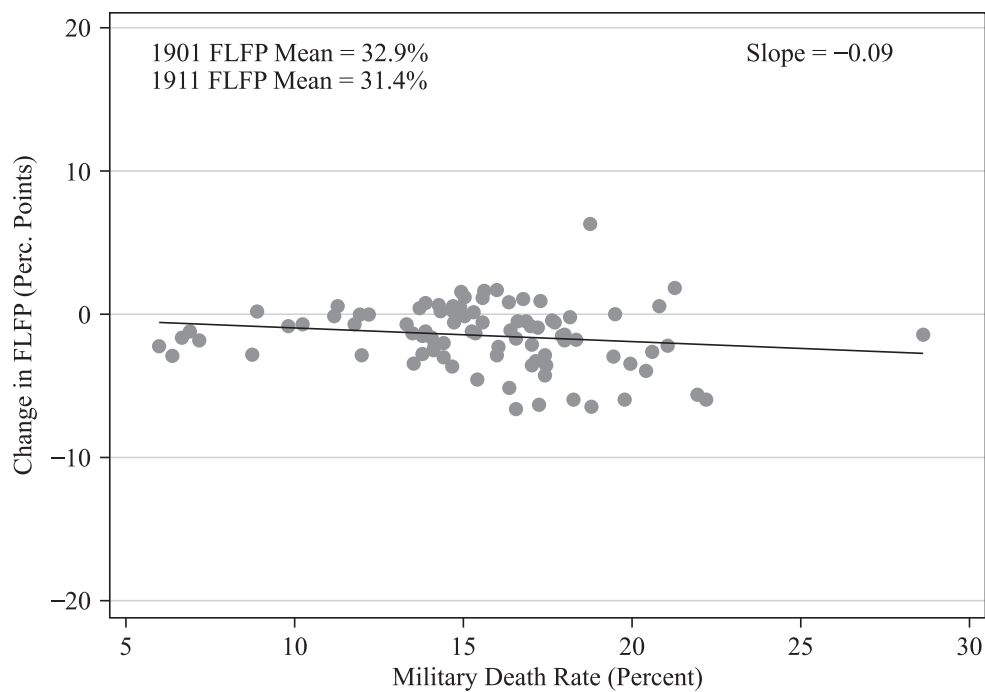
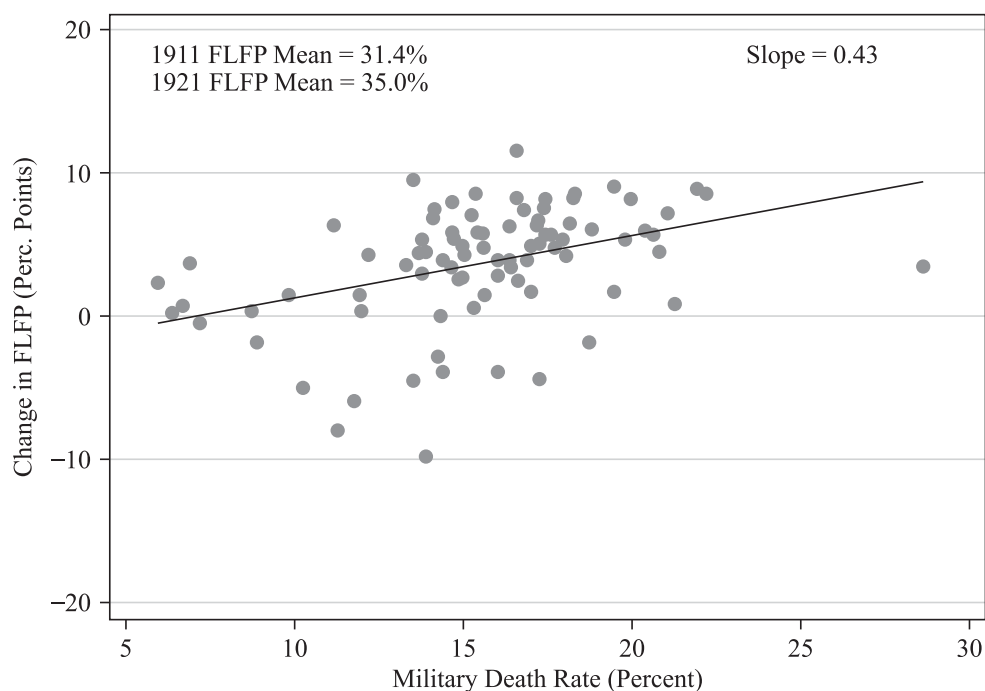
Figure 1
Adult Sex Ratio (1900–2012)

Notes: This figure displays the sex ratio among French adults aged 15–50. Data are from the censuses 1900–2012. Vertical lines indicate WWI (1914–1918) and WWII (1939–1945).

75th percentile of the distribution—female labor force participation rates were 3.5 percentage points higher throughout the interwar period, an increase of 11 percent relative to prewar levels. Eighty percent of the effect we identify stems from women entering the industrial sector, with a shift of the labor force toward blue-collar occupations and self-employment.

Next, we explore the validity of identifying assumptions. Military death rates were not randomly distributed; they were greater in rural départements due to the policies implemented by the Ministry of War to sustain the industrial war effort. Importantly, this does not invalidate our identification strategy, as military death rates were not correlated with pre-war trends in female labor force participation. Allowing for département-specific time trends, region-by-year fixed effects, or time-varying heterogeneity across départements (Bonhomme and Manresa 2015) generates results that are in line with the baseline estimate. Moreover, an instrumental variables strategy that exploits discontinuities in the timing of military service across cohorts yields similar results.

Data available for this time period suggest that supply factors related to changes in post-war marriage market conditions constitute a potentially important explanation for the patterns we identify. In particular, information on female labor force participation among widowed women indicates that these women were responsible for nearly half of the overall impact of WWI military fatalities on female labor force participation. We interpret this result as the consequence of increased labor supply due to negative income shocks experienced by war widows, whose pensions remained very low until the early

Panel A**Panel B****Figure 2*****WWI Military Death Rates and Changes in FLFP***

Notes: FLFP denotes female labor force participation rates in percent. Each dot represents one of 87 départements. The vertical axis represents changes in female labor force participation rates in percentage points between 1901 and 1911 in Panel A and between 1911 and 1921 in Panel B.

1930s. We also find that single women in départements more affected by the war delayed marriage, which likely induced them to enter the labor force while searching longer for a husband. In contrast, our analysis suggests that labor demand factors were of second-order importance. While post-war female employment rose, female wages declined in the manufacturing sector, as well as in the domestic services sector, across occupations in which men and women were closer substitutes. Therefore, substitution of firms from male labor to female labor was likely limited. To compensate for the scarcity of the male labor input, firms instead slightly increased their stock of physical capital. Finally, we find no evidence that female wartime employment was correlated with military death rates, nor that it generated a rise in female labor force participation after the war.

The remainder of the article is organized as follows. Section II discusses our contributions and the related literature, Section III describes the data and historical context, Section IV presents the main results, Section V explores the mechanisms, and Section VI concludes.

II. Contributions and Related Literature

We contribute to the literature that explores implications of war mobilization and fatalities on female labor force participation. Starting with Goldin (1991), this literature has extensively focused on the case of WWII in the United States, broadly suggesting that wartime mobilization generated an inflow of women into the labor force and that some of them kept working after the war (Acemoglu, Autor, and Lyle 2004; Goldin and Olivetti 2013; Doepke, Hazan, and Maoz 2015). Rose (2018) nuances these findings. Using direct measurement of female work during the war and of the drafting process, he shows that female wartime employment had little overall effect on female labor force participation after the war beyond a reallocation of female jobs from the nondurable manufacturing sector to the durable manufacturing sector. He further highlights that the correlation between female wartime employment and war mobilization is much weaker than previously thought. Finally, he finds no relationship between WWII military fatalities and post-war female labor force participation. Our study provides renewed evidence for the consequences of wars on female labor force participation, albeit in a different context. While, as Rose (2018), we find no evidence for a relationship between female wartime employment and post-war female labor, we show that the permanent loss of men due to the war generated a substantial inflow of women into the labor force after the war.

In that respect, our analysis helps clarify historical debates. Perhaps prompted by aggregate trends in female labor force participation that rapidly reverted to pre-war levels, contemporaneous historiography has emphasized that WWI in France was far from an engine of liberation for women (Thébaud 2013, p. 419–29).³ The surge in female employment during the war was indeed short-lived. Nevertheless, by comparing relative changes in female labor force participation across départements that were

3. For instance, Françoise Thébaud concludes her seminal study by “[t]he war, which brought hundreds of thousands of women into factories and male sectors, appears as a parenthesis” (Thébaud 2013 [1986], p. 406). Her historical analyses, however, highlight that the consequences of the war varied across women depending on their social origin, residence location, and age (Thébaud 1992).

differentially affected by military fatalities, our study paints a nuanced picture: the war increased women's presence in the labor force, but only as a result of disruptions that mostly materialized after the war.

Implications of WWI in France for marriage and fertility outcomes have also been the subject of recent research. Most related to our article, Abramitzky, Delavande, and Vasconcelos (2011) show that women in regions that experienced greater military death rates faced deteriorated post-war marriage prospects.⁴ To explain post-war patterns in fertility and female marriage choices, Vandenbroucke (2014) and Knowles and Vandenbroucke (2019) build and calibrate models of fertility choices and marital matching. All three studies rely on military fatalities data from Huber (1931, p. 426), which are only available across 22 regions and the accuracy of which has been challenged by historians (Prost 2008). Besides studying alternative consequences of the war over a longer time horizon, our analysis employs a measure of military death rates that builds upon the collection of 1.3 million individual military records (Gay and Boehnke 2020). Our measure is therefore more accurate than that in the literature as it varies across 87 units rather than just across 22. In fact, we show that variation in the Huber (1931) data is insufficient to precisely identify the effect of WWI military fatalities on female labor force participation.

Our study also contributes to our understanding of the consequences of permanent sex ratio imbalances on female labor force participation. Economic theories of marriage imply that the scarcity of one gender impacts women's working behaviors through its effects on marriage market conditions. For instance, in Grossbard's (2014) demand and supply model of marriage, a scarcity of men decreases the implicit market price of women's work in the household, which in turn increases women's supply of labor through an income effect. Collective models of household labor supply yield similar conclusions (Chiappori 1992). These theoretical predictions have been tested using various sources of variation in sex ratios, such as natural fluctuations in cohort sizes or migration shocks. For instance, exploiting sex ratio differences across cohorts in the United States between 1965 and 2005, Amuedo-Dorantes and Grossbard (2007) find a negative correlation between sex ratios and women's participation in the labor force. Alternatively, Angrist (2002) shows that changes in immigrants' sex ratios in the United States between 1910 and 1940 induced second-generation immigrant women to marry more often, contributing to a decline in their participation in the labor force.⁵ These permanent disruptions to sex ratios usually materialize progressively, generating equilibrium responses over time. Furthermore, they are typically the product of factors that also shape labor market structures, making it challenging to identify the mechanisms through which they translate (Qian 2008; Carranza 2014). We overcome these identification issues by using a permanent source of variation in adult sex ratios that is sharp—military fatalities were concentrated within a period of four years—large in magnitude, and exogenous to the outcome under scrutiny.

4. The consequences of male scarcity due to WWII for family formation and fertility have also been studied in the contexts of Bavaria, the United States, and Russia (Bethmann and Kvasnicka 2013; Jaworski 2014; Brainerd 2017).

5. Another strand of literature relies on temporary variations in sex ratios. For instance, Charles and Luoh (2010) find that rising male incarceration rates in the United States affected women's working behaviors through an impact on marriage market conditions.

Imbalances in sex ratios can further have far-reaching consequences for women's involvement in the economy. For instance, the scarcity of men in Africa due to the transatlantic slave trade between the 15th and 19th centuries resulted in higher participation of women in the labor force today (Teso 2019). Conversely, the scarcity of women in Australia due to the arrival of predominantly male British convicts throughout the 19th century resulted in lower participation of women in the labor force today (Grosjean and Khattar 2019). Little is known about the mechanisms that induced these long-run relationships to emerge in the first place: the historical nature of these phenomena generally implies a substantial lack of data at the time of the initial imbalance, preventing a proper analysis of the short-run mechanisms at play around the historical shock. This lack of quantitative evidence might jeopardize the validity of such persistence studies and, in particular, the credibility of their exclusion restrictions (Cantoni and Yuchtman 2020). By uncovering the initial channels through which sex ratio imbalances due to the war affected women's working behaviors in interwar France, our analysis provides sound foundations for exploring the mechanisms of long-run persistence of this historical episode. Such is done in Gay (2021), which shows that this historical shock to female labor transmitted to subsequent generations until today, mainly through parental intergenerational transmission channels. Overall, our article constitutes an important foundation to studies focused on the medium- and long-run consequences of gender imbalances in that it provides quantitative evidence for the mechanisms through which the relationship between male scarcity due to the war and female labor force participation emerged in the first place.

III. Data and Historical Context

A. *Female Labor Force Participation (1901–1936)*

Female labor force participation data at the département level are from the seven censuses between 1901 and 1936.⁶ While farmers' wives were to be classified as labor force participants, not all census enumerators did so in 1901 (Maruani and Meron 2012, p. 33–5). For consistency in measurement, the analysis focuses on female labor force participation net of farmers' wives. Because these women were systematically classified as farm owners whenever recorded, we avoid potential measurement error by removing them from the labor force, as nearly all female farm owners were farmers wives. Moreover, this transformation enables us to focus on paid work.

Female labor force participation is the share of women aged 15 and older in the labor force.⁷ Table 1 reports average female labor force participation rates for 1901–1936. While many women entered the labor force after the war, at least as many had dropped out by the late 1920s. Table 1 further motivates our focus on female labor net of farm owners: while the corrected measure remains stable at 33 percent between 1901 and 1906, the uncorrected measure increases by seven percentage points between these two censuses. Since there was no major shock to labor market conditions during this period,

6. Census years are 1901, 1906, 1911, 1921, 1926, 1931, and 1936, the last census before WWII. [Online Appendix J](#) provides comprehensive details about the sources of data used in this article.

7. This includes employed and unemployed women. Few unemployment benefits existed so individuals seldom had incentives to register—female unemployment rates were below 1 percent.

Table 1*Average Female Labor Force Participation Rates (1901–1936)*

	1901	1906	1911	1921	1926	1931	1936
FLFP (net of farm owners)	32.9	32.7	31.4	35.0	29.9	30.1	28.1
FLFP (uncorrected)	45.0	51.9	51.5	55.7	49.6	49.4	47.0
Difference	12.1	19.2	20.1	20.7	19.7	19.3	18.9

Notes: This table reports average female labor force participation (FLFP) rates in percent across 87 départements.

the discrepancy appears solely due to inconsistent measurement in 1901. After 1901, the difference between the two measures remains stable, around 20 percentage points, suggesting that the transformation we operate does not introduce systematic biases.

B. Military Death Rates

To build a precise measure of military death rates at the département level, we collected the individual military records for the 1.3 million French soldiers who died because of the war from the *Mémoire des Hommes* archive maintained by the Ministry of Defense.⁸ We then extracted their dates and départements of birth. The military death rate in a département is calculated as the ratio of deceased soldiers born in that département to the size of its drafted population, which we approximate by the male population aged 15–44 in 1911. This approximation is reasonable because all French male citizens aged 20–48 were subject to conscription at the onset of the war.

In Figure 3, we map the distribution of military death rates.⁹ Military death rates range from 6 percent in Belfort to 29 percent in Lozère, with an average of 15 percent. Throughout the article, we interpret regression coefficients by comparing differences in outcomes across départements that experienced high military death rates (20 percent) rather than low military death rates (10 percent). This roughly corresponds to switching from a median département in the low military death rates group (25th percentile) to a median département in the high military death rates group (75th percentile).

Two types of inaccuracies could potentially affect the measure of military death rates. First, we assign military fatalities to a département through soldiers' départements of birth. However, these might have differed from their départements of residence at the onset of the war, as 19 percent of men aged 15–44 resided outside their département of birth in 1911. This could be problematic if pre-war migration flows were correlated with

8. This archive is accessible at <http://www.memoiredeshommes.sga.defense.gouv.fr> (accessed November 30, 2021). *Online Appendix I* provides more details about this database and discusses its advantage over Huber (1931, p. 426). The number of soldiers who ultimately died as a result of the war remains uncertain as some passed several years after the war due to injuries or illnesses contracted during the conflict, but the figure of 1.3 million is the consensus (Prost 2008). It is similarly difficult to assess the number of civilian fatalities; they are usually evaluated at 40,000 (Huber 1931, p. 310–4).

9. Data are missing for the three départements that belonged to Germany before the war—Bas-Rhin, Haut-Rhin, and Moselle. They are excluded throughout the analysis.

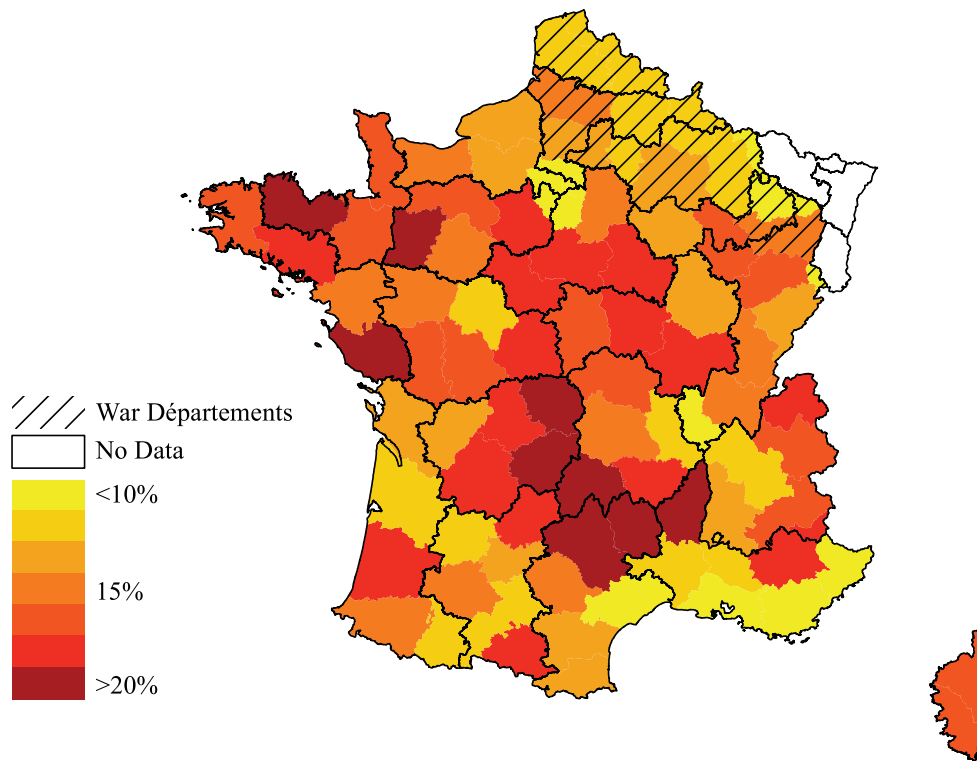


Figure 3
Distribution of Military Death Rates across 87 Départements

Notes: Data are missing for Bas-Rhin, Haut-Rhin, and Moselle. Shaded areas in the Northeast experienced war combats on their soil. Darker lines delineate military regions.

trends in female labor force participation. To alleviate this concern, we build a measure of military death rates that corrects for these patterns. Estimates with this corrected measure are similar to the baseline.

A second potential issue concerns the approximation of the pool of drafted men. We assume that men subject to conscription were drafted at similar rates across départements. However, 21 percent of them were initially exempted, mainly due to poor health (Huber 1931, p. 93).¹⁰ Using military recruitment data by cohort together with health information, we show that differential recruitment rates across départements do not affect the results.

C. Sources of Variation in Military Death Rates

The distribution of military death rates was determined by the territorial organization of military recruitment and by demographic and economic factors. Rural départements

10. Recruitment rates nevertheless increased throughout the conflict as many conscripts previously deemed “unfit” were eventually recalled to compensate for heavy military casualties. For instance, 92 percent of the cohort aged 20 in 1914 was eventually drafted (Boulanger 2001, p. 118–28). Another potential concern might be that men younger than 20 and older than 48 voluntarily enlisted. These were relatively rare cases. While 26,000 men out of 188,000 conscripts voluntarily enlisted in 1914, only 11,000 out of 211,000 did so in 1915 (Boulanger 2001, p. 128–36).

experienced greater military death rates, a correlation generated by the policies implemented by the Ministry of War to sustain the industrial war effort. Nevertheless, the distribution of military death rates was not correlated with pre-war trends in female labor force participation.

1. The territorial organization of military recruitment

The territorial organization of the military structured both the recruitment and composition of military units. These were initially composed with soldiers residing in the same military region, so that soldiers from the same region were initially sent to the same battlefields. As a result, départements in different military regions experienced disparate military death rates. However, as casualties accumulated, the military command changed this policy: after only five months into the war, soldiers were allocated based on each unit's needs, so troops from different regions were increasingly mixed together (Boulanger 2001, p. 253). The resulting intraregional correlation in military death rates is therefore low at 0.12.

2. Economic and demographic factors

The systematic part of the variation in military death rates can be explained by demographic and economic factors. To see that, we regress military death rates on pre-war characteristics and report estimates in Table 2.¹¹ Départements that experienced greater military death rates had lower female labor force participation rates before the war (Column 1) and were more rural (Column 2). Rurality is captured through two measures: the share of rural population and the share of population born in the département.¹² Together, these two measures explain 74 percent of the variation in military death rates across départements. Pre-war differences in female labor force participation that are correlated with military death rates are fully captured by differences in rurality; the coefficient on female labor force participation is nonsignificant and close to zero once rurality is controlled for (Column 3). When 17 additional pre-war characteristics are further included, only rurality exhibits statistical significance, and corresponding coefficients barely change (Column 4). Finally, including 21 military-region fixed effects to compare neighboring départements generates similar estimates (Column 5).

The policies implemented by the Ministry of War to sustain the industrial war effort explain the correlation between military death rates and rurality. As the war lingered, the military command realized that its plan for supplying troops with weapons and machinery was dramatically insufficient (Porte 2005, p. 73–82). For instance, the plan of military mobilization did not mention the production of new military equipment, providing only for 50,000 workers allocated across 30 factories (Porte 2006, p. 26). To cope with the shortage of civilian labor and the German occupation of the industrial Northeast,

11. The full results for Table 2, Column 4, are available in [Online Appendix Table A1](#).

12. Censuses define as “rural” the population that resides in municipalities with less than 2,000 inhabitants. The share of population born in the département is tied to fundamental aspects of French rurality as a measure of the intensity of the rural exodus during the late 19th century (White 2002). Average personal wealth, the share of active population in agriculture, and the share of cultivated land also capture some aspects of rurality, but all the variation in these variables is subsumed in variations in the two measures we propose.

Table 2
Military Death Rates and Pre-War Characteristics

	Dependent Variable: Military Death Rate				
	(1)	(2)	(3)	(4)	(5)
FLFP	−0.15*** [0.05]		0.03 [0.03]	−0.06 [0.07]	−0.01 [0.07]
Rural		0.12*** [0.01]	0.12*** [0.01]	0.11*** [0.03]	0.16*** [0.04]
Born in dép.		0.12*** [0.03]	0.13*** [0.03]	0.15*** [0.04]	0.11* [0.06]
Other characteristics	No	No	No	Yes	Yes
Military region FE	No	No	No	No	Yes
Départements	87	87	87	87	87
Adj. R^2	0.120	0.739	0.739	0.790	0.838

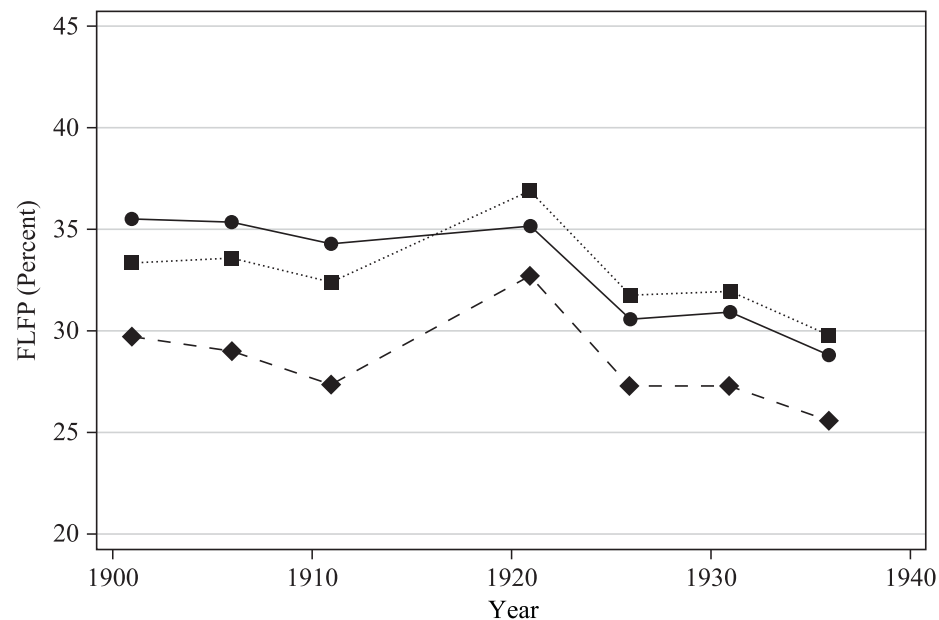
Notes: This table reports OLS estimates from regressing military death rates on pre-war département characteristics measured in 1911, otherwise noted. FLFP denotes female labor force participation in percent; Rural, the share of rural population in percent; Born in dép., the share of population born in the département in percent. Other characteristics consist of population in thousands, population density per km², average age, average height of 1911 conscripts in cm, the share of active population in industry in percent, km of roads and km of rails per km² in 1913, the share of cultivated land in 1912 in percent, personal wealth in 1908, banking deposits, and direct taxes in francs per inhabitant, the share of 1911 conscripts that could read and write and with primary education in percent, the minimum distance to the war in km, the share of students in religious schools in 1906 in percent, and turnout in 1914 in percent. Robust standard errors are in brackets. Significance: * $p < 0.10$, *** $p < 0.01$.

as early as August 1915, the Ministry of War began to withdraw soldiers with manufacturing skills from the front and sent up to 560,000 of them into war factories and mines.¹³ Moreover, administrative jobs were mostly occupied by soldiers with higher formal education, who are from urban areas (Ridel 2007). Because of these policies, soldiers from industrial and urban départements were less likely to be on the battlefield and thus had lower chances of dying in combat.

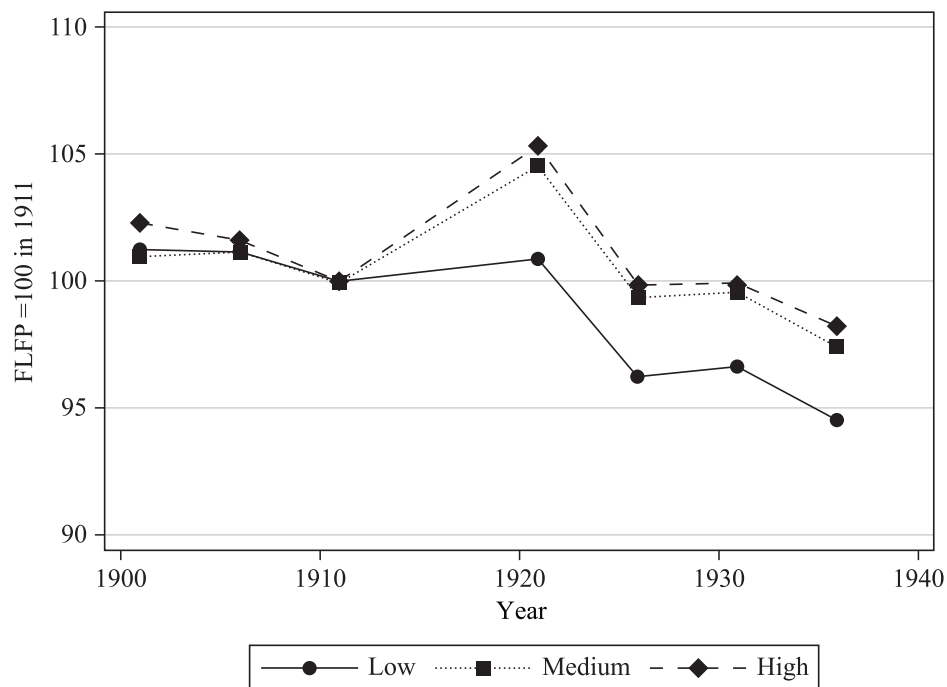
Such correlation in levels need not threaten the identification as long as the distribution of military death rates is not correlated with trends in female labor force participation. Figure 4 displays absolute and relative trends in female labor force

13. The Dalbiez law of August 17, 1915, stipulates: “The Ministry of War is authorized to allocate to corporations, factories, and mines working for the national defense men belonging to a mobilized or mobilizable age class, industrial managers, engineers, production managers, foremen, workers, and who will justify to have practiced their job for at least a year in those corporations, firms and mines, or in comparable corporations, firms, and mines” (art. 6, *Journal Officiel de la République Française, Lois et Décrets*, 47(223), p. 5785–7, dated August 19, 1915). [Online Appendix Table A2](#) provides a detailed account of the number of mobilized soldiers outside of armed services throughout the war.

Panel A



Panel B

**Figure 4***Trends in Female Labor Force Participation*

Notes: This figure displays absolute and relative trends in female labor force participation rates between 1901 and 1936 across groups of 29 départements with high, medium, and low military death rates. In Panel B female labor force participation rates are normalized to 100 in 1911.

participation rates across three groups, each composed of 29 départements that experienced high, medium, and low military death rates. Départements with different military death rates had different levels of female labor force participation rates (Panel A), but they had little differential trends in female labor force participation in the pre-war period (Panel B).

IV. The Missing Men and Female Labor Force Participation

To analyze the effect of military fatalities on female labor force participation, we use a difference-in-differences strategy and estimate the following baseline specification:

$$(1) \quad \text{FLFP}_{d,t} = \beta \text{death_rate}_d \times \text{post}_t + \theta' X_{d,t} + \gamma_d + \delta_t + \varepsilon_{d,t},$$

where $\text{FLFP}_{d,t}$ denotes the female labor force participation rate in département d and year t in percent, and post_t , an indicator variable for $t > 1918$. We cluster standard errors at the département level throughout the analysis. Département fixed effects γ_d control for département-specific unobservable characteristics that are fixed over time and might generate systematic differences in levels of female labor force participation. For instance, some départements might hold more traditional views about gender roles than others and exhibit systematically lower female labor force participation rates. Time fixed effects δ_t control for year-specific shocks that are common to all départements. $X_{d,t}$ is a vector containing the two time-varying measures of rurality.

Identification stems from relative changes in female labor force participation rates across départements with different military death rates. The baseline estimate is reported in Table 3, Column 1. In départements that experienced military death rates of 20 percent rather than 10 percent, female labor force participation was 3.5 percentage points higher after the war, an increase of 11 percent relative to pre-war levels. Removing time-varying controls does not affect this estimate (Column 2).

Next, we relax the assumption that the effect was constant over time and estimate year-specific coefficients:

$$(2) \quad \text{FLFP}_{d,t} = \sum_{\substack{\tau=1901 \\ \tau \neq 1911}}^{1936} \beta_{\tau} \text{death_rate}_d \times \text{year}_{\tau} + \theta' X_{d,t} + \gamma_d + \delta_t + \varepsilon_{d,t},$$

where we exclude the year 1911, and where year_{τ} is set of indicator variables for each year between 1901 and 1936. Estimates are displayed in Figure 5. They are stable throughout the interwar period. Coefficients on pre-war years are close to zero and non-significant, suggesting that differential pre-war trends in female labor force participation do not drive the results. They are nonetheless slightly positive. For instance, the coefficient on the lead of 1906 is 0.03 (standard error of 0.05). This implies that départements that experienced greater military death rates had a slight relative downward trend in female labor force participation before the war, which could bias post-war estimates

Table 3
Impact of WWI Military Fatalities on FLFP

	Dependent Variable: FLFP										MLFP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Death rate $\times$ post	0.35*** [0.07]	0.37*** [0.08]	0.40** [0.16]	0.21*** [0.08]	0.23** [0.09]	0.38*** [0.07]	0.45*** [0.10]	0.28*** [0.07]	0.55*** [0.13]	0.48*** [0.07]	-0.01 [0.06]
Controls	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Specification	Baseline	No controls	Dép.-specific trends	Region $\times$ year FE	Residual measure	Empl.	Female farm owner	No war dép.	Corr. measure	Pop. weights	Male placebo
Difference from baseline	0.00	0.02 [0.10]	0.05 [0.17]	-0.15 [0.10]	-0.13 [0.11]	0.03 [0.10]	0.10 [0.12]	-0.07 [0.10]	0.20 [0.15]	0.13 [0.10]	-0.36*** [0.09]
Observations	609	609	609	609	609	609	522	532	609	609	609
Départements	87	87	87	87	87	87	87	76	87	87	87
Within R^2	0.581	0.578	0.824	0.798	0.550	0.584	0.606	0.633	0.569	0.636	0.675
1911 mean	31.4	31.4	31.4	31.4	31.4	31.1	51.5	30.5	31.4	35.7	93.2

Notes: This table reports OLS coefficients from estimating Specification 1. The dependent variable is female labor force participation rate (FLFP) in percent except in Column 6, where it is female employment rate in percent, and in Column 11, where it is male labor force participation rate (MLFP) in percent. All regressions include département and year fixed effects. Controls include the share of rural population in percent and the share of population born in the département in percent. Census years are 1901, 1906, 1911, 1921, 1926, 1931, and 1936. Column 3 includes département-specific linear time trends, and Column 4 includes region-by-year fixed effects. In Column 5, military death rates are purged from pre-war trends between 1901 and 1911 in FLFP and rurality. In Column 7, female labor force participation includes female farm owners and excludes census year 1901. In Column 8, all 11 départements that experienced war combats on their territory are excluded. In Column 9, military death rates are corrected for pre-war migration patterns (see [Online Appendix E](#)). In Column 10, départements' relative population sizes in 1911 are used as weights. Standard errors are in brackets and are clustered at the département level. Significance: ** $p < 0.05$, *** $p < 0.01$.

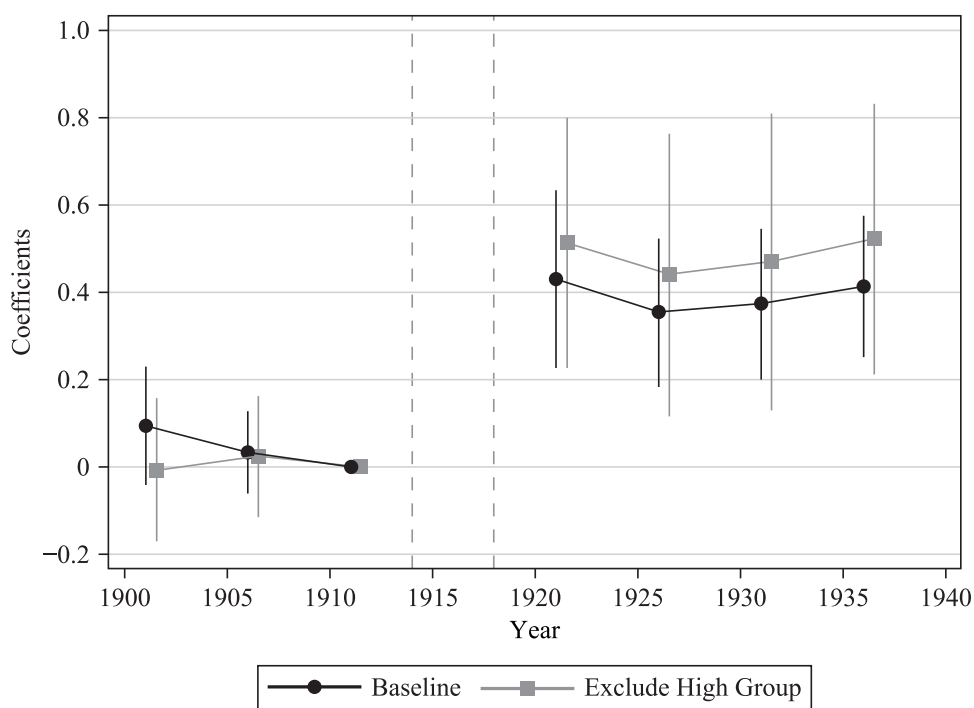


Figure 5

Impact of WWI Military Fatalities on FLFP

Notes: This figure reports year-specific OLS coefficients from estimating Specification 2. The dependent variable is female labor force participation (FLFP) in percent. Controls include the share of rural population in percent and the share of population born in the département in percent. Vertical lines represent 95 percent confidence intervals. “Exclude high group” corresponds to estimates when excluding the 29 départements with highest military death rates (among 87 départements).

downward. In Section IV.A, we show that the parallel-trends assumption is nonetheless justified in this context.

1. Decomposition by age

Female labor force participation rates for decennial age groups 20–29, 30–39, and 40–49 are available in the census of 1901 and in all post-war censuses. We reestimate the baseline specification on the sample of women aged 20–49 and obtain a coefficient of 0.40, close to the baseline (Online Appendix Table A3, Column 1). Although older women responded relatively more intensively to military fatalities than younger women did (Columns 2–4), base rates in labor force participation across groups imply that all ages contributed equally to the overall effect (Columns 5–7).

2. Decomposition by sector and occupation

In Online Appendix B, we decompose the overall effect by sector and occupation. Eighty percent stems from women entering the industrial sector, while the remainder essentially stems from women entering the domestic services sector. Moreover, we

observe a displacement of the female labor force toward blue-collar occupations and self-employment in départements that experienced greater military death rates, especially in the industrial sector. These results suggest that women mostly entered low-skilled jobs after the war.¹⁴

A. Robustness

We perform a wide range robustness checks that are summarized in Table 3, Columns 3–11. These checks, as well as additional ones provided in the [Online Appendix](#), are described below.

1. Parallel-trends assumption

We provide evidence that the parallel-trends assumption is reasonable in this context. The baseline point estimate does not change significantly when we control for département-specific linear time trends or region-by-year fixed effects, nor when we purge military death rates from pre-war trends in female labor force participation and rurality (Columns 3–5).¹⁵ Excluding départements in the higher group of military death rates results in pre-war trends that are flat at zero and in estimates that are only slightly larger than the baseline (Figure 5). In [Online Appendix C](#), we further show that our results are robust to relaxing explicitly the assumption that time fixed effects are common to all départements. Herein, we allow for time-varying heterogeneity across départements using the Bonhomme and Manresa (2015) grouped fixed effects strategy, which imposes no a priori structure on group membership.

2. Measurement of female labor force participation

Using alternative measurements of female labor force participation does not alter the results. Estimates are similar to the baseline when we exclude unemployed women and when we include female farm owners (Columns 6–7).

3. War départements

Consequences of the war might have been different in départements that experienced war combats directly if military death rates were correlated with the intensity of war destruction or the reconstruction. Excluding these 11 départements from the analysis decreases the point estimate to 0.28 (Column 8). This is because these départements were predominantly industrial, and female labor was mostly responsive to military fatalities in the industrial sector. It nevertheless suggests that war départements are not driving the results. Using data from Michel (1932), we further show that the distribution of war destruction and of the reconstruction effort across these départements was not correlated with military death rates ([Online Appendix D](#)).

14. In [Online Appendix H](#), we use individual-level data on the cohorts 1899–1908 from the census of 1968 to show that this did not affect these women's educational attainment.

15. Controlling for département-specific quadratic, cubic, or quartic time trends yields identical point estimates at 0.47, with standard errors of 0.16.

4. *Pre-war migration patterns*

We assign military fatalities to a département through soldiers' départements of birth, which might differ from their départements of residence on the eve of the war. This could introduce bias in the baseline estimate if military death rates and pre-war migration patterns were correlated. To alleviate this concern, we construct a measure that takes into account bilateral migration flows between départements in 1911 ([Online Appendix E](#)). Using this corrected measure generates a point estimate close to the baseline (Column 9).¹⁶ We also show that a measure based on départements of military recruitment rather than départements of birth is contaminated with substantial measurement error because the geography of military recruitment did not overlap département boundaries.

5. *Additional robustness checks*

In [Online Appendix F](#), we show that differences in enlistment rates due to pre-war differences in health conditions were uncorrelated with military death rates and changes in female labor force participation. Herein, we further provide evidence that the distribution of death rates due to the Spanish Flu of 1918–1919 was uncorrelated with military death rates and did not affect changes in female labor force participation. We also show in [Online Appendix Table A4](#) that our results are robust to alternative specifications of standard errors: two-way clustering on départements and years (Cameron and Miller 2015), region-level clustering, and spatial autocorrelation (Conley 1999). Furthermore, Column 10 of Table 3 reports a population-weighted regression estimate. It is slightly larger than baseline at 0.48 as more populated départements were relatively more industrial. Finally, a placebo test shows that military death rates did not affect male labor force participation rates (Column 11).

B. *Instrumental Variables Strategy*

To further support the validity of the baseline ordinary least squares (OLS) point estimate, we use an instrumental variables (IV) strategy that exploits discontinuities in the timing of military service across cohorts. At the onset of the war, the active army was composed of four age cohorts: men aged 20–23, born between 1891 and 1894. We label an age cohort by the year in which they first drafted—the year they reached age 20. For instance, the class 1914 denotes the cohort of 1894. Hence, in 1914, the active army was composed of classes 1911–1914. While the class 1914 had just been recruited, the class 1911 had gone through three years of military training and was about to be transferred to the reserve of the active army. As a result, men of classes 1911–1914 had different levels of military training at the onset of the war. They nevertheless belonged to the same military units and were initially sent to the same battlefields. Intuitively, men with more

16. Trends in female migration patterns were not altered by the war ([Online Appendix Figure A1](#)). Reestimating the baseline specification with the share of women born in their département of residence as the dependent variable results in a weak and nonsignificant estimate of 0.08 (standard error of 0.11). This alleviates the concern that post-war labor mobility might confound the results.

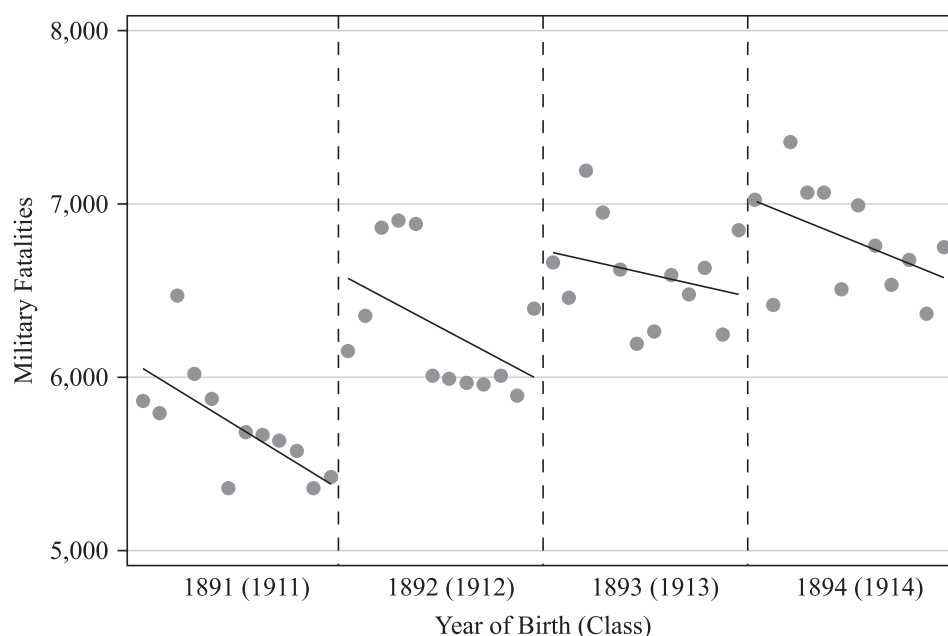


Figure 6

Military Fatalities by Month of Birth, Classes 1911–1914

Notes: Each dot represents the number of military fatalities relative to soldiers born during the same month of the same year. Black lines are regression lines for each class.

military training should be more “efficient” on the battlefield and die at lower rates.¹⁷ Consistent with this idea, discontinuities in the number of military fatalities across cohorts are apparent in Figure 6.¹⁸ We argue that these discontinuities are due to differences in military training.

Other reasons could potentially explain these differences. First, members of each class could have different initial physical or intellectual abilities. To examine this possibility, we collected height and education data from yearly recruitment reports of the army. Summary statistics, reported in [Online Appendix Table A5](#), clearly reject this possibility. Second, older soldiers might have died at lower rates than younger ones because of better physical abilities or some form of seniority.¹⁹ Averaging military death rates over an entire class in this case might yield the patterns we observe in the data. However, differences in military fatalities across classes are not driven by an averaging effect, as cohort-specific regression lines in Figure 6 do not display positive slopes.

We build on these discontinuities and create three instruments, each representing the size of a cohort relative to the next in 1911. In [Online Appendix G](#) we show that these

17. The contribution of each class to military fatalities is indeed monotonically increasing from the class 1911 to the class 1914: the class 1911 contributed 5.7 percent to overall military fatalities; the class 1912, 6.2 percent; the class 1913, 6.5 percent; the class 1914, 6.7 percent.

18. This is not driven by cyclical birth patterns, as the same pattern holds when we weight military fatalities by the number of births in each month.

19. Guillot and Parent (2018, p. 428) show that soldiers in higher ranks had a longer life expectancy during the war, although this advantage remains quantitatively limited.

Table 4
Instrumental Variables Estimates

	Dependent Variable: Military Death Rate $\times$ Post				
	(1)	(2)	(3)	(4)	
Panel A: First Stage					
Ratio class 1911–1912 $\times$ post	−0.28*** [0.05]			−0.17*** [0.03]	
Ratio class 1912–1913 $\times$ post		−0.39*** [0.05]		−0.31*** [0.06]	
Ratio class 1913–1914 $\times$ post			−0.31*** [0.10]	−0.20*** [0.05]	
Controls	Yes	Yes	Yes	Yes	
Observations	609	609	609	609	
Départements	87	87	87	87	
Dependent Variable: FLFP					
	OLS	IV			
	(1)	(2)	(3)	(4)	(5)
Panel B: Second Stage					
Death rate $\times$ post	0.35*** [0.07]	0.80*** [0.22]	0.48*** [0.14]	0.37** [0.18]	0.54*** [0.13]
Instruments	No	1	2	3	1–3
Controls	Yes	Yes	Yes	Yes	Yes
Difference with OLS	0.00	0.44* [0.23]	0.13 [0.16]	0.02 [0.19]	0.19 [0.15]
Observations	609	609	609	609	609
Départements	87	87	87	87	87
Within R^2	0.581	0.507	0.575	0.581	0.567
KPW F -statistic		29.159	55.847	10.638	27.350

Notes: This table reports first-stage coefficients in Panel A, and IV coefficients from estimating Specification 1 with class ratios as instruments in Panel B. Instrument 1 is the ratio of the class 1911 to the class 1912; Instrument 2, the ratio of the class 1912 to the class 1913; Instrument 3, the ratio of the class 1913 to the class 1914. All regressions include département and year fixed effects and controls for the share of rural population in percent and the share of population born in the département in percent. FLFP denotes female labor force participation in percent. All regressions contain 87 départements. Census years are 1901, 1906, 1911, 1921, 1926, 1931, and 1936. The KPW F -statistic is the Kleibergen–Paap Wald rk F -statistic. Standard errors are in brackets and are clustered at the département level. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

instruments are nearly randomly distributed and are not correlated with pre-war levels in female labor force participation and rurality. They are, however, strongly correlated with military death rates: first-stage estimates reported in Panel A of Table 4 imply that the larger the size of a class relative to the next, the smaller the military death rate. Moreover, they exhibit Kleibergen–Paap F -statistics between 11 and 56, suggesting little bias at the second stage. Using all three instruments together implies that in départements that experienced military death rates of 20 percent rather than 10 percent, female labor force participation was 5.4 percentage points higher after the war (Table 4, Panel B, Column 4). This is larger than the baseline OLS estimate—though not statistically different—for potentially three reasons: downward bias in the OLS estimate due to the slight relative downward trend in female labor force participation in départements that experienced greater military death rates, measurement error in military death rates, or the larger responsiveness of female labor force participation to the death of younger soldiers if the main mechanism operates through the marriage market.

In [Online Appendix G](#), we further show through a series of placebo tests that only the class ratios we consider generate meaningful results. We also show herein that IV estimates are robust to the same robustness checks as those reported in Table 3 and to alternative specifications of the instruments.

V. Mechanisms

We now investigate the mechanisms that could explain the impact of WWI military fatalities on female labor force participation. Both changes in the supply and demand for female labor could account for the pattern we identify. On the one hand, the scarcity of men might have induced firms to increase their demand for female labor, especially in sectors in which women and men were close substitutes. On the other hand, shocks to the marriage market might have induced women to increase their overall supply of labor through two mechanisms. The first consists of differential base rates in labor force participation across marital statuses. As apparent on [Online Appendix Figure A2](#), at the national level, single and widowed women had much larger propensities to work than married women. Therefore, a decline in the proportion of married women due to the war should have mechanically increased overall female labor supply. The second consists of changes in the relative intensity of female labor supply within each marital group due to negative income shocks. For instance, some war widows might have entered the labor force to compensate for the loss of their husbands' incomes as their pensions remained low until the early 1930s (Bette 2017). This negative income shock might also have induced their daughters to enter the labor force as secondary earners in their families.

We explore whether supply (Section V.A) or demand channels (Section V.B) help explain the pattern we identify. Empirical evidence points towards a supply-side explanation, although the absence of systematic information on female labor force participation by marital status at the département level prevents us from providing definitive evidence. In Section V.C, we also show that female wartime employment cannot explain post-war changes in female labor force participation.

A. Supply Factors: A Marriage Market Channel?

1. Military fatalities and the post-war marriage market

We first document the consequences of military fatalities for the post-war marriage market.²⁰ The primary channel through which the war affected the marriage market is changes in sex ratios. To illustrate the impact of military death rates on adult sex ratios, we estimate the following specification:

$$(3) \quad \text{sex_ratio}_{a,d,t} = \sum_{\substack{\tau=1901 \\ \tau \neq 1911}}^{1946} \beta_{\tau} \text{death_rate}_{a,d} \times \text{year}_{\tau} + \gamma_d + \delta_t + \mu_a + \varepsilon_{a,d,t},$$

where $\text{sex_ratio}_{a,d,t}$ denotes the sex ratio among age group a in département d and year t in percentage points, and year_{τ} , a set of indicator variables for each year between 1901 and 1946. To improve the precision of estimates, we compute decennial-cohort-specific sex ratios and military death rates for age groups 20–29, 30–39, and 40–49, and we include age-group fixed effects μ_a . Moreover, we include the year 1946 to assess the long-run consequences of the war. Estimates are displayed in Figure 7. Coefficients on pre-war years are flat at zero but strongly negative after the war. In particular, in départements that experienced military death rates of 20 percent rather than 10 percent, the sex ratio among adults aged 20–49 decreased by 6.2 percentage points between 1911 and 1921. Coefficients then revert to balance over time, reaching zero in 1946. This suggests that the marriage market remained disrupted throughout the interwar period, but that these disruptions had dissipated by the end of WWII.

Consistent with these sex-ratio effects, the share of married women aged 20–49 declined from 71 percent in 1911 to 66 percent in 1921. At the same time, the share of single women among this age group increased from 22 percent to 24 percent, while the share of widowed women increased from 6 percent to 9 percent ([Online Appendix Figure A3](#)).

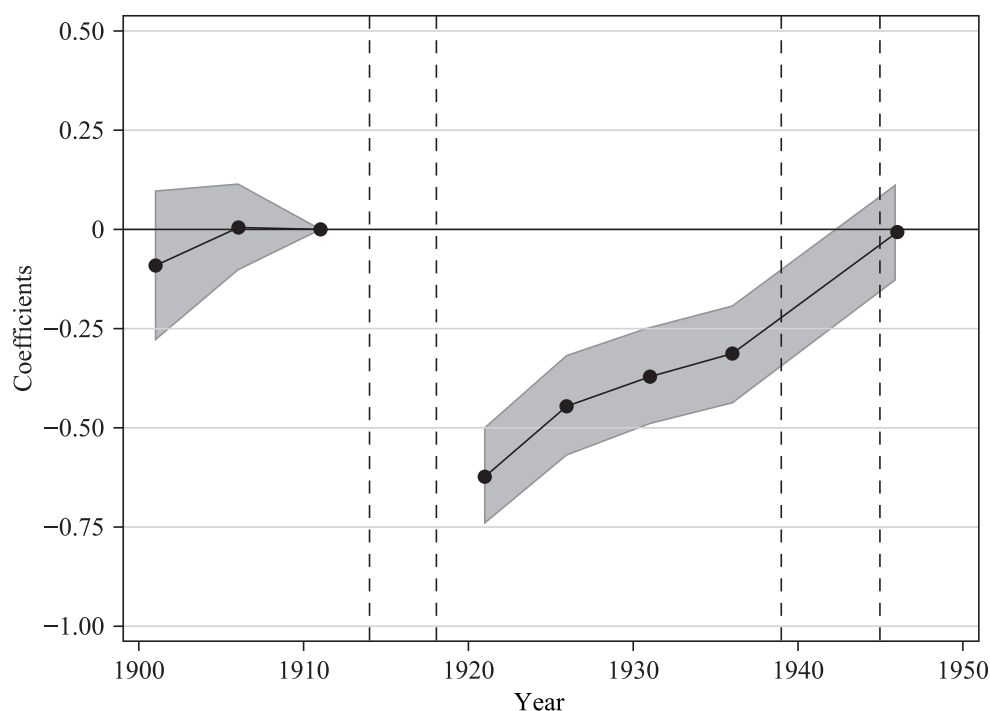
To analyze the impact of WWI military fatalities on the interwar marriage market, we estimate the following specification:

$$(4) \quad Y_{m,d,t} = \beta \text{death_rate}_d \times \text{post}_t + \theta' X_{d,t} + \gamma_d + \delta_t + \varepsilon_{m,d,t},$$

where $Y_{m,d,t}$ denotes the share of women aged 20–49 of marital status m , in département d , and year t in percent.²¹ We report results in Panel A of Table 5. Estimates imply that in départements that experienced military death rates of 20 percent rather than 10 percent, the share of single women was 2.3 percentage points higher after the war, an increase of 10 percent relative to pre-war levels (Column 1). In these départements, the share of

20. We are not the first to document this phenomenon. Using a more aggregated source of data for military death rates (Huber 1931, p. 426), Abramitzky, Delavande, and Vasconcelos (2011) show that women were less likely to marry after the war in regions that experienced greater military death rates. We replicate their main result (Table 2, p. 136) using our measure of military death rates in [Online Appendix Table A6](#).

21. We focus on this sample because age groups outside these bounds are not consistently defined across censuses. Marital statuses “widowed” and “divorced” are generally not available separately in the censuses, so we group widowed and divorced women into the same category. Moreover, because age groups within the 20–49 bounds were defined differently in the census of 1906, this year is excluded from the sample.

**Figure 7***Impact of WWI Military Fatalities on Adult Sex Ratios*

Notes: This figure reports year-specific OLS coefficients from estimating Specification 3. The dependent variable is the sex ratio among adults aged 20–49 in percentage points. Shaded areas represent 95 percent confidence intervals.

widowed women was 0.7 percentage point higher after the war, an increase of 13 percent (Column 2). Mirroring these trends, the share of married women in these départements was 3.1 percentage point lower after the war, a decline of 4 percent (Column 3). Our analysis suggests that three-quarters of these “missing married women” were single women who did not marry, while the rest were married women who became widows.

Estimates generally do not change significantly when we control for département-specific linear time trends or region-by-year fixed effects, or when we purge military death rates from pre-war trends in female labor force participation and rurality ([Online Appendix Table A7](#)). Moreover, year-specific coefficients reveal no pre-war differential trends in marriage market outcomes, suggesting that the parallel-trends assumption is again reasonable ([Online Appendix Figure A4](#)).

To explore which age groups quantitatively contribute the most to these estimates, we repeat the analysis by decennial age group, keeping the denominator equal to the female population aged 20–49. We find that the increase in the share of single women was mostly driven by younger women, aged 20–29, while the increase in the share of widowed women was driven by older women, aged 40–49 ([Online Appendix Table A8](#)).

Marriage market conditions remained disrupted throughout the interwar period, with rates of singlehood still increasing in the 1930s in départements relatively more affected by the war. In [Online Appendix H](#), we explore the consequences of the war for family

Table 5
The Marriage Market Channel

Dependent Variable:	Panel A: Marital Status			Panel B: Labor Force Participation		
	Single Women/ All Women	Widowed Women/ All Women	Married Women/ All Women	Active Women/ All Women	Active Widows/ All Women	Active Widows/ Widows
	(1)	(2)	(3)	(4)	(5)	(6)
Death rate × post	0.23*** [0.04]	0.07*** [0.02]	−0.31*** [0.05]	0.26*** [0.07]	0.12*** [0.03]	0.54*** [0.11]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	522	522	522	261	261	261
Départements	87	87	87	87	87	87
Within R^2	0.693	0.864	0.763	0.604	0.325	0.609
Pre-war mean	22.3	5.6	71.4	32.9	5.7	34.5

Notes: This table reports OLS coefficients from estimating Specification 4 in Panel A and Specification 1 in Panel B. All regressions include département and year fixed effects, and controls for the share of rural population in percent and the share of population born in the département in percent. Shares are in percent and are defined with respect to the female population aged 20–49 in Panel A and the female population aged 15 and older in Panel B. Census years are 1901, 1911, 1921, 1926, 1931, and 1936 in Panel A, and 1901, 1926, and 1936 in Panel B. Standard errors are in brackets and are clustered at the département level. Significance: *** $p < 0.01$.

formation in more details by using individual-level data from the family survey of 1954 and the census of 1968. We focus on the cohorts most directly affected by the war, the cohorts 1899–1908. Confirming findings in Abramitzky, Delavande, and Vasconcelos (2011, p. 147–8), we show that women in départements relatively more affected by the war delayed marriage and child bearing. These marriage market effects were not permanent, however, as rates of permanent singlehood among these women were not affected. This suggests that marriage market disruptions due to the war remained confined to the interwar period.

2. *The marriage market as a transmission channel*

Data to directly identify labor supply channels through changes in marriage market conditions are limited as censuses do not provide information on female labor force participation by marital status at the département level. The family surveys of 1901, 1926, and 1936 nevertheless provide information on employment status of widows at this level of aggregation.²² This enables us to assess the role of war widows in the overall effect of WWI military fatalities on female labor force participation.

We first reestimate the baseline effect when using only censuses of 1901, 1926, and 1936, and obtain an estimate of 0.26 (Table 5, Column 4). Repeating the analysis with the share of active widows among all women as the dependent variable generates an estimate of 0.12, which implies that widowed women accounted for nearly half of the overall effect of WWI military fatalities on female labor force participation (Column 5). This can be explained not only by the increase in the share of widowed women in the population, which are structurally more likely to work than married women, but also by the increase in labor force participation rates among these women. Indeed, we find that in départements that experienced military death rates of 20 percent rather than 10 percent, labor force participation rates among widowed women increased by 5.4 percentage points, an increase of 16 percent relative to pre-war levels (Column 6).

We interpret these results as the consequence of increased labor supply due to negative income shocks experienced by war widows, whose pensions remained low until the early 1930s (Bette 2017). Tracking social security laws promulgated throughout the interwar period, we estimate that pensions to a war widow amounted to a quarter of the average labor income of a working woman during the 1920s (Online Appendix Figure A5). This suggests that a substantial number of war widows had to enter the labor force to compensate for the loss of their husbands' incomes.

Although département-level information on labor force participation rates of single women do not exist, our analysis in Online Appendix H highlights that they delayed marriage, potentially spending more time searching for a husband. This might have induced some of them to enter the labor force as well, at least temporarily.²³

22. Family surveys provide information on widows but not on married or single women because they focus on family heads, and widows constitute the only category of women who were considered as such by official statistics.

23. Historical accounts support the idea that the marketplace was a platform to meet a husband. For instance, a female factory superintendent recounts the following in the 1930s: "[...] the young [female workers] prefer working at the factory than in their homes. Young women consider [the factory] as an occasion to get married" (Delagrangé 1934, p. 39).

B. Demand Factors: A Substitution Channel?

The increase in female labor force participation during the interwar period might also be explained by firms substituting male labor with female labor to cope with the scarcity of men. In a partial equilibrium framework, an increase in female wages could uncover this phenomenon. However, we documented that women increased their overall supply of labor after the war because of changes in marriage market conditions. As a result, changes in wages can only provide a partial view: on the one hand, rising female wages would imply that the increase in the demand for female labor was strong enough to overcompensate the depressing effect of increased female labor supply on wages, and, on the other hand, declining female wages would imply that the potential increase in the demand for female labor was not large enough to compensate for the depressing effect of increased female labor supply on wages.

To overcome general equilibrium effects, we analyze changes in female wages across occupations with different degrees of substitutability between male and female labor. We first consider occupations in the textile manufacturing sector that were almost exclusively occupied by women: ironers, seamstresses, and milliners.²⁴ Hourly wage rates for these occupations are available at the city level between 1901 and 1926.²⁵ Focusing on these occupations enables us to fix the demand curve for female labor: because male and female labor were not substitutes in these occupations, the scarcity of men is unlikely to have affected the demand for female labor differentially across departments. As a result, only shifts in the female labor supply curve should have influenced female wages in these occupations.

We aggregate city-level hourly wage rates at the département level and use a difference-in-differences strategy analogous to Specification 1. We report results in Panel A of Table 6. Consistent with our argument, female wages declined across all occupations in départements that experienced greater military death rates. Year-specific estimates reveal no differential pre-war trends in wage rates across départements, suggesting that the parallel-trends assumption is reasonable ([Online Appendix Figure A6](#)). While these occupations are not representative of all female occupations, they are representative of a large share of jobs women held throughout this time period, especially in the manufacturing sector. Given that the impact of the war was especially salient in that sector of activity, these results imply that labor supply factors alone might constitute a first-order explanation.

Next, we consider occupations in the domestic services sector in which male and female labor were closer substitutes: cooks and housekeepers. Annual wage rates for these occupations are available at the city level in 1913 and 1921. We transform these into hourly wage rates, assuming 2,808 annual working hours (Bayet 1997, p. 26). Focusing on these occupations provides an upper bound for the potential role of changes in labor demand through substitution. Similar to the analysis above, we aggregate city-level hourly wage rates at the département level and use a difference-in-differences strategy.

24. In 1911, there were about 260 women per man in these occupations (*Résultats Statistiques du Recensement Général de la Population 1911*, Tome I, Partie 3, p. 28).

25. In fact, these are the only female occupations for which wage rates are available throughout this time. Available years are 1901, 1906, 1911, 1921, and 1926. Wage information for other female occupations in the manufacturing sector (laundresses, lacemakers, embroiderers, and vest makers) is only available for the 1920s. Such wage information is not available for the 1930s.

Table 6
Impact of WWI Military Fatalities on Wages, Foreign Labor, and Capital

	Panel A: Manufacturing Log Wages			Panel B: Domestic Log Wages		Panel C: Foreign Labor		Panel D: Engine Power	
	Ironer (1)	Seamstress (2)	Milliner (3)	Cook (4)	Housekeeper (5)	LFP (6)	Share Pop. (7)	Log Total (8)	Per Worker (9)
Death rate × post	−0.011*** [0.004]	−0.011*** [0.003]	−0.007** [0.004]	−0.006*** [0.002]	−0.002* [0.001]	−0.11 [0.22]	−0.09 [0.06]	0.03** [0.01]	6.58*** [2.37]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	355	355	355	171	171	261	261	261	261
Départements	87	87	87	87	87	87	87	87	87
Within R ²	0.951	0.953	0.951	0.900	0.948	0.468	0.646	0.858	0.249
Pre-war levels	0.21	0.23	0.25	0.20	0.13	35.0	2.26	29,782	35.76

Notes: This table reports OLS coefficients from estimating Specification 1. All regressions include département and year fixed effects, and controls for the share of rural population in percent and the share of population born in the département in percent. The dependent variable is log hourly wage rate in Columns 1–5, labor force participation rates of foreigners in percent in Column 6, the share of foreigners in the population in percent in Column 7, the log total power of engines in kW in Column 8, and the power of engines per 100 workers in the industrial sector in kW in Column 9. Survey years are 1901, 1906, 1911, 1921, and 1926 in Columns 1–3; 1913 and 1921 in Columns 4 and 5; 1911, 1921, and 1926 in Columns 6 and 7; and 1901, 1906, and 1926 in Columns 8 and 9. Standard errors are in brackets and are clustered at the département level. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We report results in Panel B of Table 6. Again, female wages declined across all occupations in départements that experienced greater military death rates.

These results suggest that increased female labor supply might have been the driving force behind the post-war increase in female labor force participation. Nevertheless, increased labor demand through substitution appears to have played a (limited) role in the domestic services sector. Indeed, the net magnitude of the negative impact of military fatalities on female wages was smaller in this sector than in the textile manufacturing sector: while in départements that experienced military death rates of 20 percent rather than 10 percent, female wages declined by 7–11 percent in the textile manufacturing sector, they declined by 2–6 percent in the domestic services sector.

A concern might be that wage movements were driven by declining demand for goods and services due to a decline in income in départements more affected by military fatalities. If this were the case, then at least male labor force participation rates would have decreased. But they remained unchanged (Table 3, Column 9), while female labor force participation rates increased. Moreover, wages in sectors producing nontradable goods would have declined relatively more than in sectors producing tradable goods, but female wages in the domestic services sector declined relatively less than those in the manufacturing sector.

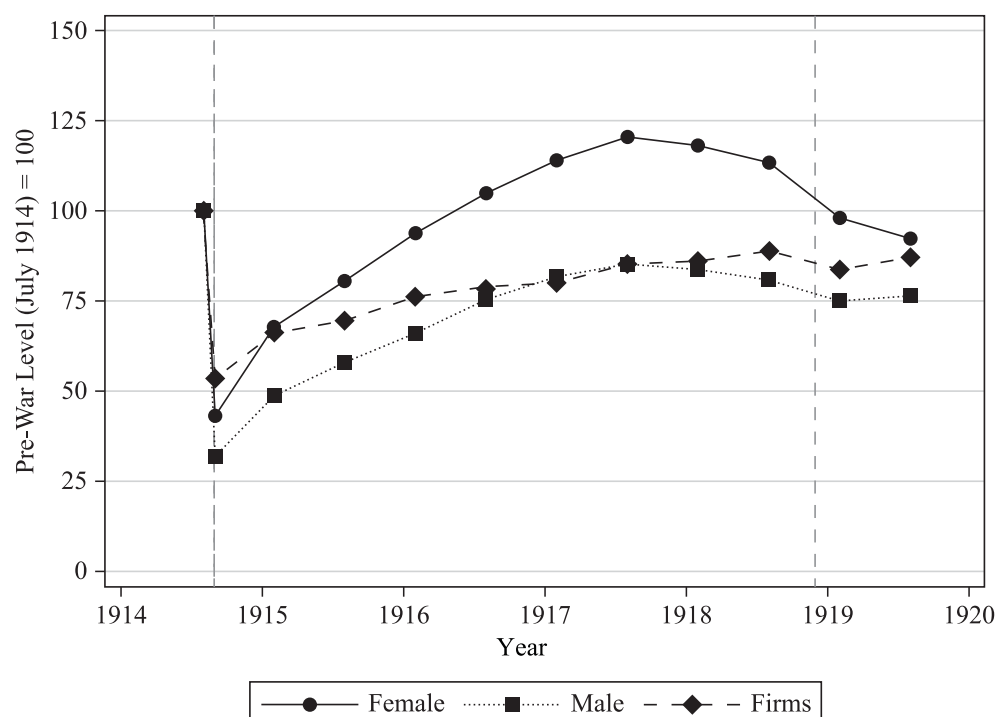
Finally, firms did not substitute toward foreign labor, as neither labor force participation rates of foreigners nor their share in the population changed across départements that experienced different military death rates (Table 6, Panel C). Available data suggest that firms instead compensated for the scarcity of male labor by increasing their stock of physical capital, as total engine power and engine power per worker in the industrial sector increased more after the war in départements that experienced greater military death rates (Table 6, Panel D). However, the magnitudes of these effects remained limited.

C. Female Wartime Employment

We now examine whether WWI military fatalities affected female labor force participation through female wartime employment. Belligerent nations anticipated a short war: since the middle of the 19th century, military strategies used strong initial attacks to defeat opponents rapidly (Van Evera 1984). Consistent with this doctrine, the French plan of military mobilization did not specify an industrial organization that would support a potentially long war, so that the industrial system nearly came to a halt in August 1914. Figure 8 displays trends in operating firms along with male and female employment in the industrial sector during the war.²⁶ Half of industrial firms ceased operating in August 1914, and male and female employment declined to 32 percent and 43 percent, respectively, of their levels of July 1914.

By the end of August 1914, 200,000 French soldiers had died in combat. The military command soon realized that the war would last longer than anticipated and that its industrial plan to support the ongoing war effort was dramatically insufficient. For instance, while 13,000 shells were produced daily at the beginning of the war, troops were using 150,000 shells per day by January 1915 (Porte 2005, p. 66–7). To manage the

26. Data are from five industrial surveys conducted in July 1917, January 1918, July 1918, January 1919, and July 1919.

**Figure 8***Labor during World War I (August 1914–July 1919)*

Notes. Female denotes employed women; Male, employed men; Firms, operating firms. Data are relative to the industrial sector. Levels are normalized to 100 in July 1914.

extended needs of the army, the military command centralized the industrial war effort under the State Secretariat of Artillery and Ammunition in November 1915 and started to coordinate a vast network of public and private industrial firms (Bostrom 2016). Moreover, the government incentivized firms to employ alternative forms of labor, such as women, immigrants, and war prisoners. As a result, the number of women employed in the industrial sector exceeded its pre-war level by July 1916. This was especially salient in sectors that directly supplied weapons and machinery to the army. For instance, in the metallurgic sector, the number of employed women exceeded its pre-war level as early as January 1915 and was nearly 700 percent higher by July 1917.²⁷

The need for new military equipment vanished at the end of the war. Moreover, the government issued laws to help soldiers return to their pre-war jobs, even offering monetary lump sums of a month's pay to women who would quit their jobs in war industries.²⁸ As a result, female employment in the industrial sector dropped below its pre-war level by January 1919.

27. Online Appendix Table A9 provides an overview of the evolution of the number of women employed across various industries during the war.

28. The law of November 22, 1918 ensured that soldiers could claim their pre-war job: "The administrations, offices, public, or private firms must guarantee to their mobilized personnel [...] the occupation that all had at the moment of its mobilization" (*Journal Officiel de la République Française, Lois et Décrets*, 50(320), p. 10120–1, dated November 24, 1918). In November 1918, the Ministry of Armament was telling female workers: "[B]y

Table 7
Military Fatalities and Female Wartime Employment

	Panel A: Death Rate		Panel B: FLFP			
	(1)	(2)	(3)	(4)	(5)	(6)
Wartime employment	−0.00 [0.01]	−0.00 [0.01]				
Wartime employment × post			−0.00 [0.01]	−0.01 [0.03]	−0.00 [0.01]	−0.02 [0.04]
Death rate × post			0.35*** [0.07]	0.35*** [0.07]	0.35*** [0.07]	0.36*** [0.07]
Wartime employment × death rate × post				0.00 [0.00]		0.00 [0.00]
Reference date	07–1917	07–1918	07–1917	07–1917	07–1918	07–1918
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	87	87	609	609	609	609
Départements	87	87	87	87	87	87
Adj./Within R^2	0.745	0.746	0.582	0.582	0.581	0.582

Notes: Panel A reports OLS coefficients from regressing military death rates on changes in female employment in percent between July 1914 and July 1917 in Column 1 and between July 1914 and July 1918 in Column 2. Panel B reports OLS coefficients from estimating Specification 1. All regressions include controls for the share of the rural population in percent and the share of the population born in the département in percent. In Panel B, regressions include département and year fixed effects. FLFP denotes female labor force participation in percent. In Panel A, we report the adjusted R^2 , and robust standard errors are in brackets. In Panel B, we report the within R^2 , and standard errors are in brackets and are clustered at the département level. Significance: *** $p < 0.01$.

Women who entered the labor force during the war might have kept working after the war because they acquired valuable skills and experience, updated their beliefs about the benefits from working, or improved their information about labor market conditions. We capture the relative intensity of female wartime employment with the percentage change of women working in the industrial sector between July 1914 and July 1917, or June 1918. Départements that experienced greater increases in female employment during the war did not experience different military death rates (Table 7, Panel A). Hence, the potential impact of female wartime employment on subsequent female labor force participation is orthogonal to the mechanisms we highlight.

coming back to your previous occupations, you will be useful to your country as you have been by working in war industries in the past four years. [...] Each [female] worker who expresses the will to quit one's firm before December 5, 1918, will receive the amount of thirty days of salary as a severance pay" (*Bulletin du Ministère du Travail*, 1919, p. 45*–6*).

Using a difference-in-differences strategy, we further find that départements with greater increases in female wartime employment did not experience a post-war rise in female labor force participation (Table 7, Panel B). Including an interaction term between military death rates and increases in female wartime employment does not affect the results. Similarly, we find no effect of female wartime employment on the distribution of occupations or sectors of activity (Online Appendix Table A10). These findings are consistent with contemporaneous reports of labor inspectors, which recount how managers systematically divided tasks of female workers.²⁹ Because of this division of labor, women could hardly acquire human capital transferable to other sectors after the war (Downs 1995).³⁰ These results parallel those of Rose (2018), who finds that female wartime employment during WWII in the United States was orthogonal to soldiers mobilization and did not affect post-war female labor.

VI. Conclusion

We show that the scarcity of men due to World War I in France induced many women to enter the labor force after the war. In départements that experienced military death rates of 20 percent rather than 10 percent, female labor force participation increased by 11 percent relative to pre-war levels. This effect is stable throughout the interwar period and robust to alternative empirical strategies. Available data for this time suggest that labor supply factors (changes in marriage market conditions) rather than labor demand factors (substitution from male to female labor) help explain the pattern we identify.

This study provides evidence that jolts of history can generate rapid and long-lasting changes to women's involvement in the economy. Yet, the response of female labor to sex ratio imbalances we identify was arguably amplified by the historical context in which it occurred—most women were not in the labor force at the time, and low-skilled jobs in the manufacturing sector were increasingly available because of the transition toward Taylorism during the interwar period (Downs 1995). As a result, changes in marriage market conditions profoundly affected the extensive margin of female labor. Analyzing the impact of WWI military fatalities on female labor across countries with different characteristics might be crucial to better understand dependencies between the mechanisms we highlight and the historical context and to gauge the external validity of our findings.

29. For instance, a labor inspector in a report of January 1918 describes: “[T]o make female labor possible and enable [women] to replace men, industrialists have, in many regions, modified and improved their managing methods. They divide labor to the extreme, organize production in series and assign female workers to very delimited tasks” (*Bulletin du Ministère du Travail et de la Prévoyance sociale*, 25(1), 1918, p. 11).

30. Historians have further pointed out that instead of an inflow of women into the labor force, women employed in war factories were already working before the war. For instance, Downs (1995, p. 48) writes: “In the popular imagination, working women had stepped from domestic obscurity to the center of production, and into the most traditionally male of industries. In truth, the war brought thousands of women from the obscurity of ill-paid and ill-regulated works as domestic servant, weavers and dressmakers into the brief limelight of weapons production” (cited in Vandenbroucke 2014, p. 118).

References

- Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos. 2011. "Marrying Up: The Role of Sex Ratio in Assortative Matching." *American Economic Journal: Applied Economics* 3(3):124–57.
- Acemoglu, Daron, David H. Autor, and David Lyle. 2004. "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury." *Journal of Political Economy* 112(3):497–551.
- Amuedo-Dorantes, Catalina, and Shoshana Grossbard. 2007. "Cohort-Level Sex Ratio Effects on Women's Labor Force Participation." *Review of Economics of the Household* 5(3):249–78.
- Angrist, Josh. 2002. "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation." *Quarterly Journal of Economics* 117(3):997–1038.
- Bayet, Alain. 1997. "Deux Siècles d'Évolution des Salaires en France." Insee Document de Travail F9702.
- Bethmann, Dirk, and Michael Kvasnicka. 2013. "World War II, Missing Men and Out of Wedlock Childbearing." *Economic Journal* 123(567):162–94.
- Bette, Peggy. 2017. *Veuves Françaises de la Grande Guerre*. Paris: Peter Lang.
- Bonhomme, Stéphane, and Elena Manresa. 2015. "Grouped Patterns of Heterogeneity in Panel Data." *Econometrica* 83(3):1147–84.
- Bostrom, Alex. 2016. "Supplying the Front: French Artillery Production during the First World War." *French Historical Studies* 39(2):261–86.
- Boulanger, Philippe. 2001. *La France Devant la Conscription: Géographie Historique d'une Institution Républicaine, 1914–1922*. Paris: Economica.
- Brainerd, Elizabeth. 2017. "The Lasting Effect of Sex Ratio Imbalance on Marriage and Family: Evidence from World War II in Russia." *Review of Economics and Statistics* 99(2):229–42.
- Cameron, Colin A., and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2):317–72.
- Cantoni, Davide, and Noam Yuchtman. 2020. "Historical Natural Experiments: Bridging Economics and Economic History." NBER Working Paper 26754. Cambridge, MA: NBER.
- Carranza, Eliana. 2014. "Soil Endowments, Female Labor Force Participation, and the Demographic Deficit of Women in India." *American Economic Journal: Applied Economics* 6(4):197–225.
- Charles, Kerwin Kofi, and Ming Ching Luoh. 2010. "Male Incarceration, the Marriage Market, and Female Outcomes." *Review of Economics and Statistics* 92(3):614–27.
- Chiappori, Pierre-André. 1992. "Collective Labor Supply and Welfare." *Journal of Political Economy* 100(3):437–67.
- Conley, Timothy. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92(1):1–45.
- Delagrangé, Juliette. 1934. "Le Travail de la Femme Mariée." In *Bulletin du Comité Permanent des Associations des Surintendantes d'Usines et de Services Sociaux*, ed. Association des Surintendantes d'Usines et de Services Sociaux, 31–48. Paris: Elbé.
- Doepke, Matthias, Moshe Hazan, and Yishay D. Maoz. 2015. "The Baby Boom and World War II: A Macroeconomic Analysis." *Review of Economic Studies* 82(3):1031–73.
- Downs, Laura L. 1995. *Manufacturing Inequality: Gender Division in the French and British Metalworking Industries, 1914–1939*. Ithaca: Cornell University Press.
- Gay, Victor. 2021. "The Legacy of the Missing Men: The Long-Run Impact of World War I on Female Labor Force Participation." <https://hal.archives-ouvertes.fr/hal-02523129> (accessed April 4, 2022).

- Gay, Victor, and Jörn Boehnke. 2020. "Replication Data for: The Missing Men. World War I and Female Labor Force Participation." Version 3. Harvard Dataverse. <https://doi.org/10.7910/DVN/AP1HZ8>
- Girault, Arthur. 1915. "La Diminution de la Population Adulte Male en Europe et ses Conséquences Économiques et Sociales." *Revue d'Économie Politique* 29(6):431–46.
- Goldin, Claudia. 1991. "The Role of World War II in the Rise of Women's Employment." *American Economic Review* 81(4):741–56.
- Goldin, Claudia. 2006. "The Quiet Revolution That Transformed Women's Employment, Education, and Family." *American Economic Review* 96(2):1–21.
- Goldin, Claudia, and Claudia Olivetti. 2013. "Shocking Labor Supply: A Reassessment of the Role of World War II on Women's Labor Supply." *American Economic Review* 103(3): 257–62.
- Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. 2005. "Engines of Liberation." *Review of Economic Studies* 72(1):109–33.
- Grosjean, Pauline, and Rose Khattar. 2019. "It's Raining Men! Hallelujah? The Long-Run Consequences of Male-Biased Sex Ratios." *Review of Economic Studies* 86(2):723–54.
- Grossbard, Shoshana. 2014. *The Marriage Motive: A Price Theory of Marriage: How Marriage Markets Affect Employment, Consumption, and Savings*. New York: Springer.
- Guillot, Olivier, and Antoine Parent. 2018. "'Farwell Life, Farwell Love': Analysis of Survival Inequalities among Soldiers Who 'Died for France' During World War I." *Population* 73(3):411–44.
- Huber, Michel. 1931. *La Population de la France Pendant la Guerre*. New Haven, CT: Yale University Press.
- Jaworski, Taylor. 2014. "'You're in the Army Now': The Impact of World War II on Women's Education, Work, and Family." *Journal of Economic History* 74(1):169–95.
- Knowles, John, and Guillaume Vandenbroucke. 2019. "Fertility Shocks and Equilibrium Marriage-Rate Dynamics." *International Economic Review* 60(4):1505–37.
- Maruani, Margaret, and Monique Meron. 2012. *Un Siècle de Travail des Femmes en France*. Paris: La Découverte.
- Michel, Edmond. 1932. *Les Dommages de Guerre de la France et leur Réparation*. Paris: Berger-Levrault.
- Ngai, Rachel L., and Barbara Petrongolo. 2017. "Gender Gaps and the Rise of the Service Economy." *American Economic Journal: Macroeconomics* 9(4):1–44.
- Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual review of Economics* 8:405–34.
- Porte, Rémy. 2005. *La Mobilisation Industrielle: Premier Front de la Grande Guerre?* Paris: 14–18 Éditions.
- Porte, Rémy. 2006. "Mobilisation Industrielle et Guerre Totale: 1916, Année Charnière." *Revue Historique des Armées* 242:26–35.
- Prost, Antoine. 2008. "Compter les Vivants et les Morts: l'Évaluation des Pertes Françaises de 1914–1918." *Le Mouvement Social* 222(1):41–60.
- Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance." *Quarterly Journal of Economics* 123(3):1251–85.
- Ridel, Charles. 2007. *Les Embusqués*. Paris: Armand Colin.
- Rose, Evan K. 2018. "The Rise and Fall of Female Labor Force Participation during World War II in the United States." *Journal of Economic History* 78(3):673–711.
- Teso, Edoardo. 2019. "The Long-Term Effect of Demographic Shocks on the Evolution of Gender Roles: Evidence from the Transatlantic Slave Trade." *Journal of the European Economic Association* 17(2):497–534.
- Thébaud, Françoise. 1992. "La Grande Guerre: Le Triomphe de la Division Sexuelle." In *Histoire des Femmes en Occident*, Volume 5, ed. Françoise Thébaud, 31–74. Paris: Plon.

- Thébaud, Françoise. (1986) 2013. *Les Femmes au Temps de la Guerre de 14*. Paris: Payot.
- Vandenbroucke, Guillaume. 2014. "Fertility and Wars: the Case of World War I in France." *American Economic Journal: Macroeconomics* 6(2):108–36.
- Van Evera, Stephen. 1984. "The Cult of the Offensive and the Origins of the First World War." *International Security* 9(1):58–107.
- White, Paul. 2002. "Internal Migration in the Nineteenth and Twentieth Centuries." In *Migrants in Modern France*, ed. Philip Ogden and Paul White, 13–33. Crows Nest, Australia: Unwin Hyman.

THE INTERGENERATIONAL TRANSMISSION OF WORLD WAR I ON FEMALE LABOUR*

Victor Gay

Demographic shocks tied to World War I's high death toll induced many women to enter the labour force in the immediate post-war period. I document a positive impact of these newly employed women on the labour force participation of subsequent generations of women until today. I also find that the war permanently altered attitudes toward the role of women in the labour force. I decompose this impact into three channels of intergenerational transmission: transmission from mothers to daughters, transmission from mothers-in-law to daughters-in-law via their sons and transmission through local social interactions.

Over a million French men never came back from World War I (WWI) trenches. This unprecedented demographic shock induced many women to enter the labour force in the immediate post-war period to make ends meet (Boehnke and Gay, 2022). What implications, if any, did these new working women have on the French society? I show that women induced to work by the war affected the working behaviours of subsequent generations of women until today. I hypothesise three channels of intergenerational transmission underpinning this result: these women's impact on their daughters, on their daughters-in-law via their sons and on their surrounding communities. These mechanisms of intergenerational transmission have been pointed out as instrumental in explaining the dramatic growth in female labour force participation rates that occurred across industrialised economies throughout the twentieth century (Hazan and Maoz, 2002; Fernández *et al.*, 2004; Fogli and Veldkamp, 2011; Fernández, 2013; Hiller and Baudin, 2016).¹ Yet, empirical evidence for these is still scarce. This article shows that all three mechanisms were simultaneously at work in France throughout the second half of the twentieth century, long after sex ratios had reverted to balance.

To identify intergenerational transmission independent from confounding changes to local labour market structures induced by the war, I use an empirical strategy that relies on the working

* Corresponding author: Victor Gay, Department of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole, 1, Esplanade de l'Université, 31000, Toulouse, France. Email: victor.gay@tse-fr.eu.

This article was received on 7 September 2021 and accepted on 31 March 2023. The Editor was Ekaterina Zhuravskaya.

The data and codes for this article are available on the Journal repository. They were checked for their ability to reproduce the results presented in the article. The author was granted an exemption to publish parts of his data because access to these data is restricted. However, the author provided the Journal with temporary access to the data, which enabled the Journal to run his codes. The codes for the parts subject to exemption are also available on the Journal repository. The restricted access data and these codes were also checked for their ability to reproduce the results presented in the article. The reproduction package for this article is available at the following address: <https://doi.org/10.5281/zenodo.7762881>.

I am grateful to Richard Hornbeck for his guidance and support, as well as to Derek A. Neal, Steven D. Levitt and James A. Robinson. This article has also benefited from fruitful discussions with Ran Abramitzky, Jessica S. Bean, Jörn Boehnke, Oliver Browne, Charlotte Cavaillé, Eszter Czibor, Raquel Fernández, Mohamed Saleh and Estefania Santacreu-Vasut. I gratefully acknowledge the CISSR and the Department of Economics at the University of Chicago, as well as the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program) for financial support.

¹ Olivetti and Petrongolo (2016) describe the unprecedented increase in female labour force participation across industrialised economies during the twentieth century. See Goldin (1990; 2006; 2014) for a focus on the United States and Maruani and Meron (2012) for a focus on France.

behaviours of women who make decisions under similar local labour market conditions, but whose geographic origins differ in their exposure to the war. This approach implies comparing women born in counties (*départements*) that experienced different military death rates during the war—the ratio of deceased soldiers to the size of the drafted population—and therefore different historical shocks to female labour supply, but who reside in the same county and therefore face similar local institutional constraints when making decisions.² To assess the long-run effect of WWI on female labour force participation through intergenerational transmission, I apply this approach to individual-level census data from 1962 to 2012, complemented with labour force survey data from 1982 to 2012. I find that women born in counties that experienced military death rates of 20% rather than 10%—equivalent to switching from the 25th to the 75th percentile of the distribution—were 5 percentage points more likely to be working. This explains half of the overall long-run impact of WWI on female labour force participation.

To account for this finding, I explore three channels: transmission through parents, through marriage and through local social interactions. To identify parental transmission, I use the extended labour force surveys 2005–12 and compare women born and residing in the same county, but whose parents were born in counties that experienced different military death rates. I find evidence for a strong mother-to-daughter transmission channel. Next, to identify transmission through marriage, I compare women born and residing in the same county, but whose husbands were born in different counties. I find that women married to men born in counties that experienced greater military death rates are more likely to work. Consistent with this result, I further find evidence for a strong mother-in-law to daughter-in-law transmission channel. Finally, using municipality-level variation in the exposure to WWI, I find that this historical shock to female labour transmitted through local social interactions as well. Comparing estimates' magnitudes suggests that mothers and mothers-in-law were primary agents in the transmission of the legacy of the war.

To provide a rationale for the long-run impact of WWI on female labour force participation, I argue that women induced to enter the labour force during the interwar period altered attitudes toward female labour of their daughters, sons and neighbours, and that these changes translated into the working behaviours of women in subsequent generations. Since individuals form preferences and beliefs early in life from learning and socialising with their parents and peers (Bisin and Verdier, 2011; 2023), men and women who grew up with a working mother, or in an environment in which many women worked, should form more progressive attitudes toward female labour.³ Consistent with this argument, using the Generation and Gender Survey 2005, I find that men and women born in counties that experienced greater military death rates hold more progressive attitudes toward the role of women in the labour force today.

² *Départements* represent the second level of subnational government. France was divided into 87 *départements* before the war, 90 during the interwar period and 95 today. Throughout this article, I use the term *county* to designate *départements*.

³ Many studies find a correlation between both beliefs toward female labour and the working behaviours of mothers and their daughters', of mothers-in-law and their daughters-in-law' and of mothers and their daughters' peers. Farre and Vella (2013) show that US mothers' attitudes toward gender roles and their working behaviours are correlated with their daughters' attitudes when young and working behaviours when adults, a phenomenon corroborated for the United States (Morrill and Morrill, 2013), Great Britain (Johnston *et al.*, 2014) and Mexico (Campos-Vazquez and Velez-Grajales, 2014), though not for Egypt (Miyata and Yamada, 2016). Transmission from mothers-in-laws' labour choices to their daughters-in-law' has been found for Japan (Kawaguchi and Miyazaki, 2009), Switzerland (Bütikofer, 2013), the United States (Morrill and Morrill, 2013), China (Chen and Ge, 2018; Li and Liu, 2019) and West Germany (Schmitz and Spiess, 2022). Olivetti *et al.* (2020) show that the working behaviours of the mothers of a woman's friends when growing up affect her own working behaviour later in life.

To account for the revolution of female labour that occurred during the twentieth century (Goldin, 2006), models of intergenerational transmission typically highlight one specific mechanism—e.g., transmission from mothers to daughters (Fernández, 2011), from mothers-in-law to daughters-in-law (Fernández *et al.*, 2004) or through local social interactions (Fogli and Veldkamp, 2011).⁴ This article is the first to provide comprehensive empirical evidence that these mechanisms were simultaneously at work throughout the revolutionary period, over three generations. For instance, Fernández *et al.* (2004) argue that one explanation for the revolution of female labour is the increasing number of men who grew up with a working mother. They show that women's wartime employment during World War II (WWII) in the United States, which they proxy by wartime mobilisation, propagated to their daughters-in-law after the war—though Rose (2018) nuances these findings.⁵ Consistent with Fernández *et al.* (2004), I find evidence for a mother-in-law to daughter-in-law transmission channel. But this mechanism constitutes only one among those I am able to identify. Importantly, in contrast to Fernández *et al.* (2004) and the broader literature, the empirical strategy I use enables me to identify intergenerational transmission independent from confounding changes to local labour market conditions. Furthermore, using WWI rather than WWII as a source of variation provides the opportunity to investigate the diffusion dynamics of women's working behaviours across three generations, from the beginning to the end of the revolution of female labour. The case of WWI in France is also particularly advantageous to explore these mechanisms compared to the case of WWII in the United States because American black male labour alleviated the consequences of the war on female labour force participation (Ferrara, 2022).

This article also uncovers mechanisms of historical persistence for which the evidence is still scarce. Because of the methodological focus of the literature on a location-based approach to history—how historical events in a location shape long-run outcomes in that location—domestic institutions have been pointed out as primary channels of historical persistence (Acemoglu *et al.*, 2005; Nunn, 2014; Michalopoulos and Papaioannou, 2017). Partly because of a lack of appropriate data, a crucial vehicle of historical persistence has received less attention: individuals. Understanding how individuals transmit legacies of history across generations requires moving from a location-based to a lineage-based approach. Applying this methodology, Michalopoulos *et al.* (2019) show that pre-modern economic lifeways in Africa still affect individuals' economic outcomes today.⁶ Alesina *et al.* (2013, p. 524) find that half of the overall impact of historical plough use on contemporaneous gender-role attitudes is due to the transmission of cultural norms from ancestors to descendants rather than to long-run changes to institutional structures. Using a more systematic approach, I show that lineage aspects also account for half of the long-run impact of WWI on female labour force participation. This implies an important role for individuals relative to local institutional structures in generating historical persistence. My results further suggest that local social interactions are not to be neglected to understand how history transmits through individuals. Another novelty of this article is the use of a lineage approach

⁴ Traditional explanations for the growth in female labour force participation rates generally focus on the role of technological change, such as the greater availability of labour-saving consumer durable goods (Greenwood *et al.*, 2005), oral contraceptives (Bailey, 2006), the fall in child care cost induced by the marketisation of home production (Attanasio *et al.*, 2008), improvements to maternal health (Albanesi and Olivetti, 2016) and the structural transformation (Ngai and Petrongolo, 2017).

⁵ A large empirical literature suggests that wartime mobilisation during WWII in the United States generated an increase in female labour supply after the war (Acemoglu *et al.*, 2004; Goldin and Olivetti, 2013; Doepke *et al.*, 2015; Brodeur and Kattan, 2022). Using more accurate measurement, Rose (2018) nuances these findings and shows that manpower mobilisation and WWII military fatalities had limited effects on post-war female labour force participation.

⁶ Nunn and Wantchekon (2011) and Teso (2019, pp. 26–8) use a similar approach.

within a country's regions to study the transmission of a historical shock. For instance, some specifications compare second-generation internal migrants that originate from neighbouring counties. This reduces potential biases that might arise when comparing movers to stayers, which is often an issue for studies using a lineage approach.

Finally, this article complements the literature that investigates how cultural norms emerge and change over time (Giuliano and Nunn, 2021). Various studies show that historical shocks alter the economic roles of women in subsequent generations, particularly shocks to sex ratios (Giuliano, 2017, pp. 20–3). Teso (2019) finds that sex ratio imbalances generated by the transatlantic slave trade induced current generations of women to enter the labour force in Sub-Saharan Africa. Grosjean and Khattar (2019) show that the scarcity of women in nineteenth-century Australia, due to the arrival of predominantly male British convicts, had adverse long-run consequences for women's position in the labour force and for attitudes toward gender roles today. A strand of recent studies further explores the long-run consequences of historical shocks to sex ratios on gender roles today, such as the repercussions of the Triple Alliance War of 1864–70 in Paraguay (Boggiano, 2020; Alix-Garcia *et al.*, 2022) or of the Portuguese Colonial War of 1961–74 (Cardoso and Morin, 2018). I similarly find that historical sex ratio imbalances had permanent implications for the economic roles of women. However, in contrast to these studies, the historical shock I consider is sharper—it was generated over a period of four years—which enables me to rule out alternative channels through which it might have translated over the long run. More importantly, two elements differ substantially in the analysis offered in this article compared to the broader literature. First, the empirical strategy based on contrasting the roles of geographic origins from that of incentives arising from local labour market structures helps disentangle the role of culture from that of institutions. Second, by examining outcomes continuously from the early 1960s to the present, the analysis makes it possible to assess the nature of the persistence of intergenerational effects over time. In particular, results imply that the persistence of the implications of the war has been constant and is not dependent upon the time point selected for the analysis.⁷

In the remainder of this article, I first provide an overview of the impact of WWI on female labour during the interwar period (Section 1). I then describe the datasets used in the analyses (Section 2) and show that this shock persisted (Section 3) through mechanisms of intergenerational transmission (Section 4). Finally, I show that this shock also altered attitudes toward female labour (Section 5).

1. World War I and Female Labour in the Interwar Period

In this section, I first describe the county-level measure of military death rates used throughout this article and analyse its sources of geographic variation (Section 1.1). I then discuss the short-run effects of the war on female labour force participation (Section 1.2).

1.1. *World War I Military Fatalities*

1.1.1. *Measuring military death rates*

My empirical strategy exploits the war as a severe exogenous shock to the adult sex ratio. Of the 8 million French men drafted throughout the conflict, 1.3 million died because of the war—a

⁷ Several recent studies have highlighted how the past might have time-dependent implications on the present (see, e.g., Ochsner and Roesel, 2017; Cantoni *et al.*, 2020 and Fouka and Voth, 2022).

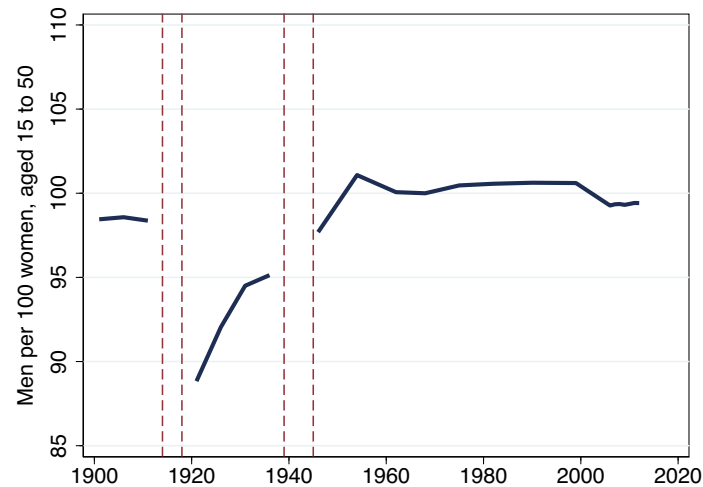


Fig. 1. *Adult Sex Ratio (1900–2012)*.

Notes: This figure displays the sex ratio among French adults aged 15 to 50. Data are from the censuses 1901–2012. Vertical lines indicate WWI (1914–18) and WWII (1939–45).

military death rate of 16%. As a result, it was not until the 1950s that the adult sex ratio recovered its balance (Figure 1).⁸ To measure the local intensity of this shock, I construct a county-level measure of WWI military death rates using the individual military records of all 1.3 million French soldiers who died because of the war.⁹ The military death rate in a county is then defined as the ratio of deceased soldiers born in the county to the size of its drafted population, which I capture by its male population aged 15 to 44 in the census of 1911, the last census before the war, as conscription was universal for all men aged 20 to 48.¹⁰ The distribution of the resulting county-level military death rates is displayed in Figure 2. Military death rates are 15% on average and range from 6% in Belfort to 29% in Lozère. Throughout the analysis, I interpret regression coefficients by comparing differences in outcomes across counties that experienced military death rates of 20% rather than 10% during the war, which corresponds to switching from the 25th to the 75th percentile of the distribution. Echoing national trends displayed in Figure 1, disruptions to sex ratios caused by WWI military fatalities remained circumscribed to the interwar period (Online Appendix Figure A1).

1.1.2. *Sources of variation in military death rates*

Two elements explain the systematic part of the variation in military death rates across counties: the territorial organisation of military recruitment and the industrial war effort. The territorial organisation of the military structured both recruitment and army corps, which initially comprised soldiers from the same military region.¹¹ Soldiers from the same region were therefore sent to

⁸ World War II affected the adult sex ratio in France only marginally as 115,000 French soldiers died because of the war, a military death rate of 1.5% (Lagrou, 2002).

⁹ Military records are from the *Mémoire des Hommes* (MDH) archive held by the Ministère des Armées. Online Appendix H.1 provides more details on the construction of this dataset. See also Boehnke and Gay (2022) and the associated data (Boehnke and Gay, 2020).

¹⁰ While conscription was universal, 21% of men aged 15 to 44 were initially exempted due to poor health conditions (Huber, 1931, p. 93). Nevertheless, local health conditions are entirely captured by the measures of rurality controlled for throughout the analysis in this article (see Boehnke and Gay, 2022, Appendix F).

¹¹ This territorial structure was inherited from the reorganisation of the army after the Franco-Prussian War of 1870. The structure in place in 1914 was defined by the law of December 22, 1913 (Boulanger, 2001, pp. 16–24).

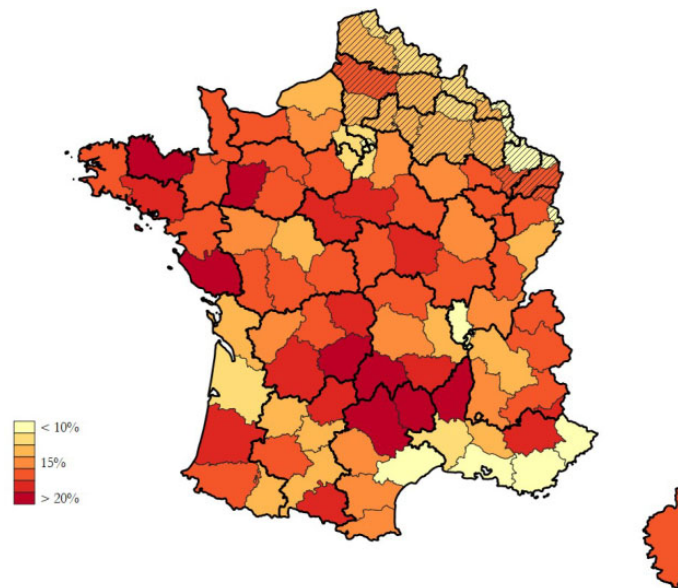


Fig. 2. *Distribution of Military Death Rates across Counties.*

Notes: This figure displays military death rates across counties in 1914 geography. Shaded areas indicate counties where war combats occurred. Darker lines indicate military region boundaries. Shapefiles for counties and military regions are from Gay (2020a) and Gay (2020b), respectively (Gay, 2021).

the same battlefields at the onset of the war.¹² However, as military fatalities soon accumulated, the military command changed its allocation policy: after five months into the war, soldiers were allocated to depleted army corps by priority, effectively pooling soldiers from different regions (Boulanger, 2001, p. 253). As a result, military death rates varied greatly across counties, even within regions—military region fixed effects explain only 16% of the variation in military death rates across counties.

As the war lingered, the military command's plan to supply troops with war equipment proved dramatically insufficient (Porte, 2005; Bostrom, 2016).¹³ To cope with the ongoing war effort, the Ministry of War recalled soldiers with manufacturing skills, so that up to 12% of soldiers were employed in war factories during the conflict, and an additional 8% were employed in the military administration.¹⁴ As a result, soldiers from industrial and urban areas were less likely to die in combat than those from rural areas. This pattern clearly emerges when regressing military death rates on pre-war county characteristics (Table 1): rural counties experienced greater military death rates, where rurality is captured by the share of rural population and the share of population born in the county.¹⁵ Rurality explains 75% of the variation in military death rates (column (1)). Determinants of subsequent female labour force participation available in historical statistics—

¹² The mobilisation plan—the *Plan XVII*—designed in 1912, assigned each army corps to a specific area upon invasion by German troops. The structure of this plan was constrained by the railway network (Gonzalez-Feliu and Parent, 2016).

¹³ For example, the *Plan XVII* procured 13,600 75 mm shells per day. Nearly 40,000 were used daily during the Race to the Sea in October 1914. By then, half of the stock of 75 mm shells had been depleted (Bostrom, 2016, p. 264). See also Porte (2005, pp. 73–82).

¹⁴ This allocation policy was enacted by the Dalbiez law of August 17, 1915. Soldiers were also allocated to steel production and harvesting from 1917 onward.

¹⁵ The census of 1911 defines the share of rural population as the share of population that resides in municipalities with fewer than 2,000 inhabitants. The share of population born in the county is tied to fundamental aspects of French rurality as a measure of the intensity of the rural exodus during the late nineteenth century.

Table 1. *Military Death Rates and Pre-War County Characteristics.*

Dependent variable	Military death rate		
	(1)	(2)	(3)
Share rural population	0.12*** [0.01]	0.12*** [0.03]	0.14*** [0.03]
Share born in county	0.12*** [0.03]	0.13*** [0.03]	0.15*** [0.04]
Female labour force participation		−0.03 [0.03]	−0.02 [0.04]
Share girls aged 5–19 in school		0.09 [0.07]	0.04 [0.07]
Total fertility rate		0.01* [0.01]	0.02 [0.01]
Personal wealth (10,000 francs)		0.04** [0.02]	0.06* [0.04]
Other characteristics	No	Yes	Yes
Military region FEs	No	No	Yes
Counties	87	87	87
R^2	0.745	0.827	0.875

Notes: This table reports OLS coefficients from regressing military death rates on various pre-war county characteristics measured in 1911. *Other characteristics* include average height, population, the shares of men working in industry and in agriculture, the share of the literate population and the average direct taxes per inhabitant. Robust standard errors are in brackets. *** Significant at the 1% level. ** Significant at the 5% level. * Significant at the 10% level.

pre-war female labour force participation, female education, total fertility, personal wealth per inhabitant—as well as additional characteristics do not appear correlated with military death rates, and only explain an additional 8% of their variation (column (2)).¹⁶ Including 20 military region fixed effects explains an additional 5% (column (3)). I further show in Online Appendix B that pre-war attitudes toward gender roles are not correlated with military death rates, although measuring these accurately remains challenging. Throughout the analysis, I rely on the residual variation in military death rates following the specification in column (3). I interpret this residual variation as non-systematic and related to the randomness at which soldiers encountered violence on the battlefield. Panel A of Table 2 provides summary statistics for these pre-war county-level characteristics.

1.2. *Female Labour During the Interwar Period*

To assess the short-run effect of WWI military death rates on women's working behaviours during the interwar period, Boehnke and Gay (2022) exploit differential changes in female labour force participation rates before and after the war across counties that experienced different military death rates. Their results imply that in counties that experienced military death rates of 20% rather than 10%, female labour force participation rates were 4 percentage points higher during the interwar period, an increase of 12% relative to pre-war levels.

This inflow of women into the labour force occurred after rather than during the war, as counties with greater female wartime employment did not experience a post-war rise in female labour force participation. It was instead driven by increased labour supply of war widows, whose pensions remained limited until the early 1930s (Bette, 2017), as well as of young single

¹⁶ These additional characteristics include the share of active population in the industrial sector, average height, total population, the share of literate population and average direct taxes per inhabitant.

Table 2. *Summary Statistics across Datasets.*

Panel A. Pre-war county characteristics (87 counties)								
	Mean	SD	Min.	Max.				
Military death rate	0.16	0.04	0.06	0.29				
Share rural population	0.67	0.18	0.00	0.89				
Share born in county	0.80	0.12	0.27	0.94				
Female labour force participation	0.52	0.09	0.34	0.69				
Share girls 5–19 in school	0.58	0.05	0.47	0.73				
Total fertility rate	2.42	0.44	1.56	3.73				
Personal wealth (10,000 francs)	0.16	0.04	0.06	0.29				
Panel B. Census respondents' characteristics (1962–2012)								
	Non-migrants (3,818,911 respondents)				First gen. migrants (3,213,921 respondents)			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Active	0.65	0.48	0	1	0.68	0.46	0	1
Employed	0.61	0.49	0	1	0.63	0.48	0	1
High school +	0.22	0.42	0	1	0.37	0.48	0	1
Number of children	1.93	1.36	0	14	1.80	1.25	0	14
Panel C. Extended labour survey respondents' characteristics (2005–12)								
	Non-migrants (117,331 respondents)				Second gen. migrants (15,091 respondents)			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Active	0.81	0.39	0	1	0.85	0.36	0	1
Employed	0.78	0.42	0	1	0.81	0.39	0	1
High school +	0.36	0.48	0	1	0.50	0.50	0	1
Number of children	1.27	1.11	0	13	1.34	1.09	0	9

Notes: This table reports summary statistics across datasets: pre-war characteristics across counties in panel A, characteristics of married women aged 30–49 in censuses 1968–2012 in panel B and characteristics of married women aged 30–59 in extended labour force surveys 2005–12 in panel C. *SD* denotes standard deviation, *Min.* denotes minimum and *Max.* denotes maximum. In panel B, non-migrants are individuals who reside in their county of birth, while first generation migrants are those who reside outside their county of birth. In panel C, non-migrants are individuals who reside in the same county as their county of birth as well as the county of birth of both their parents, while second generation migrants are individuals who reside in their county of birth, but whose mothers and fathers were both born in another county. Summary statistics are computed using sample weights provided in the censuses and in the labour force surveys. *Employed* and *Active* are indicator variables for whether the respondent is working or in the labour force, respectively. *Number of children* corresponds to the number of children of the respondent's family in the household. *High school +* corresponds to an indicator variable for whether the respondent achieved at least high school.

women, who had to enter the labour force and become secondary earners in their families while searching longer for a husband (Abramitzky *et al.*, 2011).¹⁷ In contrast, labour demand factors were of second-order importance: comparing female wages across sectors with different degrees of substitutability between male and female labour, Boehnke and Gay (2022) find that substitution of firms from male labour to female labour was likely limited.

¹⁷ Perturbed conditions on the marriage market were however temporary and circumscribed to the interwar period: using the family survey of 1954 and the census of 1968, Boehnke and Gay (2022, Appendix H) show that, while women of the cohorts 1899–1908 in counties with higher military death rates married older and had slightly less children, they had similar rates of definitive singlehood by the end of their lives, suggesting that they only delayed marriage due to the war.

2. Data

To assess the long-run implications of WWI beyond the interwar period, I combine several datasets that capture female labour force participation and beliefs toward female labour throughout the second half of the twentieth century: the population censuses 1962–2012, the labour force surveys 1983–2012, the extended labour force surveys 2005–12 and the Generation and Gender Survey 2005. I describe these datasets below and provide summary statistics in Table 2. Online Appendix H provides extensive details on the construction of these datasets.¹⁸

2.1. *The Population Censuses 1962–2012*

The main analyses in this article rely on the combination of eight individual-level population censuses, from 1962 to 2012.¹⁹ These censuses represent 20%–25% samples of the population, except those of 1962 and 1999, which are 5% samples. They provide respondents' city of residence, county and year of birth, position in the labour force, occupation, educational attainment and marital status. The availability of household identifiers further makes it possible to proxy for fertility through the number of children present in the household.

The population of interest consists in married women aged 30 to 49, born and residing in France—close to 7 million respondents in the censuses 1962–2012. I focus on this age group to ensure that human capital investments are completed and to abstract from retirement decisions. I restrict the sample to married women because their entrance in the labour force represented the core of the revolution of female labour: the share of working women in this sample increased from 30% in the early 1960s to 80% in the late 2000s. Panel B of Table 2 provides summary statistics and Online Appendix Figure A2 displays sample means for labour, fertility and education outcomes over time.

2.2. *The Labour Force Surveys 1982–2012*

To corroborate results obtained with the censuses, I further combine 31 annual labour force surveys from 1982 to 2012.²⁰ Labour force surveys present both drawbacks and advantages over the censuses. On the one hand, information about respondents' counties of birth is not available in the surveys prior to 1982, while it is available from 1962 onward in the censuses. Moreover, the sample size of the surveys is smaller than that of the censuses: while most censuses are 20%–25% samples of the population, the surveys represent 1.5%–3% samples. On the other hand, the labour force surveys provide a broader range of labour market outcomes: weekly hours worked, ever worked status, time with current firm and monthly wage rate. Therefore, the use of the labour force surveys makes it possible to extend the main analysis thanks the richer labour information they provide.

¹⁸ Beyond the data described in this section, the analysis in this article relies on the following additional sources and datasets, all described in Online Appendix H: Flaus (1947), the Commission Consultative des Dommages et des Réparations (1951), INSEE (1953; 2008; 2010; 2012c; 2018; 2019), Boehnke and Gay (2020) and Gay (2020a,b).

¹⁹ The census of 1962 is the earliest for which individual-level data are available in computer format. Census files are produced by INSEE and are available for 1962; 1968; 1975; 1982; 1990; 1999; 2006; 2012a. From 2004 onward, censuses are conducted annually on a rotating basis for a fifth of the French population. Rigorous comparisons across censuses after 2004 can therefore only be made five years apart (Brilhault and Caron, 2016).

²⁰ Labour force surveys are produced by INSEE and are available in three waves: 1982–9; 1990–2002; 2003–2012.

2.3. *The Extended Labour Force Surveys 2005–12*

To explore intergenerational transmission from parents to children, I use INSEE's extended annual labour force surveys 2005–12. Compared to the regular labour force surveys, starting in 2005, their extended versions further contain the counties of birth of respondents' parents as well as other parental information, including the labour status of respondents' parents during their childhood. This provides a unique opportunity to analyse the role of parental origins in the intergenerational diffusion process of the war. Panel C of Table 2 provides summary statistics.

2.4. *The Generation and Gender Survey 2005*

Finally, to assess the long-run implications of WWI on attitudes, I use the Generation and Gender Survey 2005 (INSEE and INED, 2005; Régnier-Loilier, 2016). It provides information about respondents' attitudes toward gender roles, religion, the family and marriage, as well as information about the distribution of household tasks and child care within the family. Importantly, it is the only survey available that provides both respondents' counties of birth and of residence, which is necessary to implement the empirical strategy.²¹ Details of survey questions and summary statistics are discussed in Section 5.

3. The Overall Intergenerational Effects of WWI on Female Labour

I now explore the overall long-run effect of WWI on female labour force participation. I first describe the empirical strategy, which relies on contrasting the working behaviour of women residing in the same county, but born in counties that experienced different military death rates (Section 3.1). I then discuss the results based on census data (Section 3.2) as well as their robustness (Section 3.3). Finally, I provide further results based on the richer outcomes from the labour force surveys (Section 3.4). Later on, in Section 4, I decompose these overall effects into three mechanisms: transmission through parents, marriage and local social interactions.

3.1. *Empirical Strategy*

To assess the long-run effect of WWI military death rates on female labour force participation, a first approach is to compare labour force participation rates of married women aged 30 to 49 across counties, and estimate the following specification separately for each census between 1962 and 2012:

$$\text{employed}_{irt} = \beta \text{death_rate}_r + \alpha \mathbf{X}'_{1911,r} + \theta_{i:1914,r} + \varepsilon_{irt}, \quad (1)$$

where employed_{irt} is an indicator variable that denotes whether woman i residing in county r in year t is employed and $\mathbf{X}_{1911,r}$ is a set of county-level pre-war controls that capture the systematic determinants of military death rates and subsequent female labour force participation.²² To make tighter comparisons within groups of women with similar characteristics and reinforce the credibility of the conditional independence assumption, I further include a set of cohort by military region fixed effects, $\theta_{i:1914,r}$, as soldiers from the same military region were initially

²¹ Two other datasets with information on attitudes toward gender roles are the opinion barometer of the DRESS and the International Social Survey (ISSP), but they do not provide respondents' counties of birth.

²² Vector $\mathbf{X}_{1911,r}$ contains the share of rural population, the share of population born in the county, female labour force participation, a measure of female education, total fertility and personal wealth per inhabitant, all measured in 1911.

sent to the same battlefields, making the distribution of military death rates more likely to be idiosyncratic within these regions.²³ I cluster standard errors at the level of counties. In this specification, β is identified from variations in the working behaviours of married women of the same cohort and residing in neighbouring counties that experienced different military death rates.

Differences in female labour force participation rates across counties identified through the above strategy could be resulting from two channels: intergenerational cultural transmission, which is the focus of this article, but also changes to local labour market structures. Indeed, interwar changes in female labour force participation were driven by increased female labour supply, which slightly depressed female wages in counties that were relatively more affected by the war (Boehnke and Gay, 2022, pp. 21–3). Lower female wages might have incentivised firms to specialise in female labour-intensive activities. Had these changes persisted, the incentive structure faced by women would differ systematically across counties that experienced different military death rates, explaining why labour force participation rates of women are higher in these counties.

To isolate the role of intergenerational cultural transmission from such institutional factors, I use an empirical strategy that relies on the working behaviours of women who make decisions under similar local labour market conditions, but whose geographic origins differ in their exposure to the war.²⁴ I implement this strategy by comparing internal migrant married women aged 30 to 49 residing in the same county, but born in counties that experienced different military death rates:

$$\text{employed}_{ibrt} = \beta \text{death_rate}_b + \alpha \mathbf{X}'_{1911,b} + \theta_{i;r;1914,b} + \varepsilon_{ibrt}, \quad (2)$$

where $\theta_{i;r;1914,b}$ are cohort by military region of birth by county of residence fixed effects.²⁵ I use two-way clustering and cluster standard errors at the levels of counties of birth and of residence. Here, β is identified from variations in the working behaviours of women of the same cohort and residing in the same county, but born in different neighbouring counties. By fixing variations arising from differences in local labour market structures, β solely captures intergenerational cultural transmission channels.

3.2. Results

The initial cross-county approach following Equation (1) implies that women residing in counties that experienced military death rates of 20% rather than 10% were 6–10 percentage points more likely to be working between 1962 and 2012 (Figure 3). Now, focusing on internal migrant women and holding constant their counties of residence following Equation (2), reveals that intergenerational transmission channels account for half of this overall effect: estimates reported in Figure 4 imply that women residing in the same county, but born in neighbouring counties that experienced military death rates of 20% rather than 10%, were 4–5 percentage points more likely to be working between 1962 and 2012. These estimates are significant at the 1% level,

²³ There were 20 military regions in 1914, containing five counties on average—they would be fairly close-by, if not strictly neighbouring in most cases (Figure 2). This strategy generates 400 comparison groups with 450–2,800 observations per group on average, depending on the census.

²⁴ This approach is sometimes called the ‘epidemiological approach’ to culture (Fernández, 2011). It has been used, for instance, to identify the role of culture for labour choices of immigrant women to the United States (Fernández and Fogli, 2009; Blau *et al.*, 2011; Gay *et al.*, 2018).

²⁵ This strategy generates 36,800 comparison groups with 4–17 observations per group on average, depending on the census.

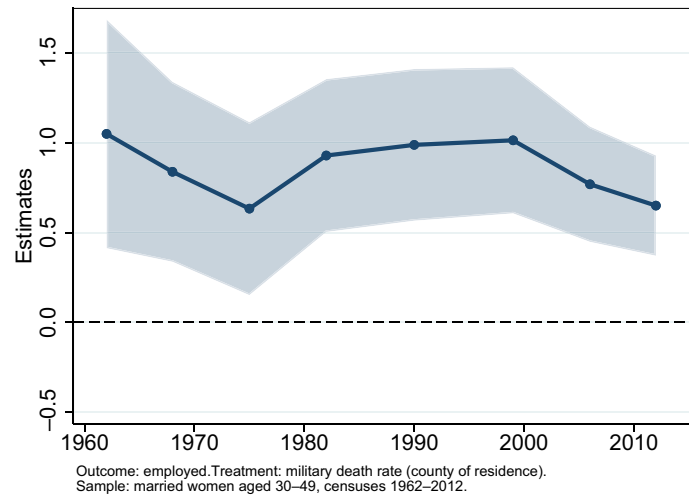


Fig. 3. *The Effect of WWI Military Death Rates on Female Labour Force Participation across Counties.*
Notes: This figure displays OLS coefficients from estimating Equation (1) along with 95% confidence intervals on the sample of married women aged 30 to 49 in the censuses 1962–2012. They result from regressing an employed indicator on the military death rate of respondents' counties of residence, along with county-level pre-war controls and cohort by military region fixed effects. Estimates are computed using sample weights provided in the censuses. Standard errors are clustered at the level of counties of residence.

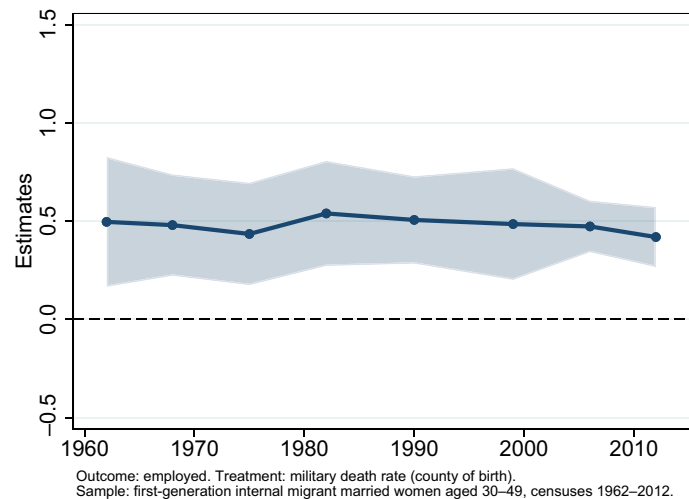


Fig. 4. *The Effect of WWI Military Death Rates on Female Labour Force Participation across Counties of Birth, Holding Counties of Residence Constant.*

Notes: This figure displays OLS coefficients from estimating Equation (2) along with 95% confidence intervals on the sample of internal migrant married women aged 30 to 49 in the censuses 1962–2012. They result from regressing an employed indicator on the military death rate of respondents' counties of birth, along with county-level pre-war controls and cohort by military region by county of residence fixed effects. Estimates are computed using sample weights provided in the censuses. Standard errors are clustered at the level of counties of residence and birth.

stable across time and represent 47%–69% of overall cross-county estimates of Figure 3. Since base rates increased from 34% in 1962 to 84% in 2012, their relative size declined from 15% of the mean to 5% over the period (Online Appendix Figure A5).

3.3. *Robustness*

I now perform a series of robustness checks that support the credibility of baseline estimates reported in Figure 4. The full set of results summarised below is discussed in detail in Online Appendix C. I also show in Online Appendix D that military fatalities and destruction from WWII do not affect the results.

3.3.1. *Alternative specifications*

Estimates are robust to alternative regression models: marginal coefficients from probit and logit models are similar to coefficients from a linear probability model (OLS). They are also similar when the outcome is a labour force participant indicator, when widening age bounds to 25–59 and when including women of all marital statuses in the regression sample.

3.3.2. *Selective migration patterns*

Because coefficients are identified from behaviours of internal migrants, selective migration might account for part of the results if out-migration patterns were correlated with both military death rates and labour market outcomes. This is unlikely. First, half of married women aged 30 to 49 were internal migrants as of 2012.²⁶ Moreover, trends in the proportion of married women among internal migrant women are identical to trends in the proportion of married women among all women (Online Appendix Figure A4). Importantly, internal migrant and non-migrant women are broadly alike along observable characteristics: regressing observable characteristics on a migration status indicator reveals that internal migrant women are equally likely to be working, were born in counties with similar military death rates and have the same number of children. They are, however, slightly positively selected in terms of education. This source of selection might bias the results if it was operating differently across individuals originating from counties that experienced low and high military death rates. Including an interaction term between the internal migrant indicator and the military death rate of respondents' counties of birth reveals that selection along all observable characteristics—including education—does not operate systematically differently across counties of birth.

Beyond selection through out-migration patterns, results might be biased through selective in-migration patterns, that is, through the destination location of internal migrants. In particular, individuals born in counties that experienced greater military death rates might settle in more dynamic labour markets—though removing the 10 most urban counties of destination from the sample leaves the results unchanged. Moreover, in-migration selection might operate within destination counties, with those born in counties that experienced greater military death rates sorting systematically into the more dynamic local labour markets within these counties. Comparing internal migrants within their local labour market of residence through employment zones and municipality fixed effects also leaves baseline estimates unchanged.

²⁶ Counties are relatively small, with a median radius of 43 km. The median migration distance is 150 km. The *Geographic Mobility and Urban Concentration* study of 1961 reports that, among married women aged 30 to 49, the primary migration motive was family related for two-fifths of them and work related for another two-fifths (Girard *et al.*, 1964).

To further alleviate potential issues of selective migration patterns, I control for the relative attractiveness of origin and destination counties through dyadic measures that capture pull and push forces between each pair of counties, together with the distance between them. I also restrict the sample to individuals who migrated at least a decade earlier—the one-and-a-half generation—since they might be subject to less biases due to recent migration. Both strategies generate estimates that are slightly lower than the baseline, though not statistically different. Overall, these robustness checks suggest that selective migration cannot account for the results.

3.3.3. *Inaccurate assignment of military death rates*

The empirical strategy used throughout this article requires regressors to be assigned at the level of geographic origin, i.e., the county in which internal migrants' ascendants were residing during the interwar period. However, censuses do not provide ascendants' counties of birth. I am therefore constrained to assume that counties of birth and of origin are identical. Here, I relax this assumption by repeating the analysis on the extended labour force surveys 2005–12, which provide parental counties of birth. I restrict the regression sample to internal migrants born before the mid-1960s in the same county as both their parents. These restrictions ensure that these individuals' parents were likely born in the 1920s and mid-1930s, and that their counties of birth trace to the interwar period.²⁷ Estimates are slightly larger than the baseline—though not significantly different—suggesting that inaccurate assignment of military death rates creates some attenuation bias through measurement error. A related concern is that counties of birth and childhood might differ. Comparing women with similar migration histories generates results identical to the baseline.

3.3.4. *Education and fertility*

Labour market outcomes are endogenous to human capital investment and fertility choices. As a result, these decisions might mediate the relationship between women's working behaviours and WWI military death rates. Controlling for educational attainment and the number of children leaves results unchanged, suggesting that the long-run impact of WWI military death rates was direct rather than mediated by education and fertility choices.

3.4. *Further Results*

Further results are reported in Online Appendix E. Herein, I show that women born in counties that experienced greater military death rates do not make different fertility or education choices, corroborating that these decisions are not mediators of the effect identified above. Moreover, they do not hold more male-biased occupations, suggesting that the war did not have long-lasting implications on gender occupational segregation. Also, these women do not marry at different rates, and when they marry, their husbands do not differ along observable characteristics. This suggests that marriage market disequilibria due to the war highlighted by Abramitzky *et al.* (2011) remained circumscribed to the interwar period. I also find little heterogeneity across women with different numbers of children, education, age or marital status. In particular, life-cycle patterns in labour force participation are not affected by WWI military death rates. I further

²⁷ The share of French men born in France that remained in their county of birth was 75.0% in 1921, 74.7% in 1926, 74.2% in 1931 and 73.2% in 1936 according to the censuses, suggesting limited migration movements between the end of the war and the mid-1930s. For women, these figures are 75.2, 74.9, 73.5 and 73.2, respectively.

provide cohort-specific estimates from cohort 1910 to cohort 1970, as well as a placebo test using the male sample.

Next, I combine the 31 annual labour force surveys from 1982 to 2012 and run a reproduction exercise that corroborates results obtained with the censuses.²⁸ The analysis of labour force surveys further reveals that women born in counties that experienced greater military death rates are more attached to the labour force, but that this influence on the extensive margin does not translate to the intensive margin, as they work shorter hours. A final concern is that, although there is no evidence of a mediating role of observable educational attainment, results might still be driven by unobserved heterogeneity in human capital. Should this be the case, it would be reflected in wage rates. However, I find that women born in counties that experienced greater military death rates do not earn higher wages.

4. Intergenerational Transmission Mechanisms

To account for the overall intergenerational effect of the war on female labour identified above, I explore three mechanisms of intergenerational transmission: transmission through parents (Section 4.1), transmission through marriage (Section 4.2) and transmission through local social interactions (Section 4.3).

4.1. *Transmission through Parents*

4.1.1. *Empirical strategy*

To isolate intergenerational transmission from parents to daughters, I use the same empirical strategy, but one generation earlier: I compare the working behaviours of women who make decisions under similar local labour market conditions, but whose parents' geographic origins differ in their exposure to the war. I now focus on the sample of second-generation internal migrant married women aged 30 to 59 in the extended labour force surveys 2005–12, which provides the counties of birth of respondents' parents.²⁹ In particular, I restrict second-generation internal migrants to women who reside in their county of birth, but whose mothers and fathers were both born in another county.³⁰ Focusing on second- rather than first-generation internal migrants improves the credibility of identification as a typical respondent in the sample has parents born in the 1930s. Their counties of birth therefore plausibly trace to the war (see footnote 27). Moreover, since the location of second-generation internal migrants was determined prior to their birth, results are further less likely to be driven by selective migration.³¹

²⁸ I run the same specification as the baseline on a different sample of the same underlying population—in this sense, it is a reproduction test (Clemens, 2017, p. 327).

²⁹ I widen the upper bound of the age restriction to 59 compared to 49 for analyses using large-scale census data because extended labour force surveys have limited sample sizes, which might create issues of low statistical power in specifications that include parental region of birth fixed effects.

³⁰ In the extended labour force surveys, 19% of married women aged 30 to 59 who reside in their county of birth have mothers born outside their county of birth, 20% have fathers born outside their county of birth and 9% have both parents born outside their county of birth.

³¹ Second-generation migrants and natives are broadly alike: they are equally likely to be employed, have the same number of children and their mothers were born in counties with similar military death rates. Second-generation migrants are slightly more educated in general, but not differentially so when from counties that experienced greater military death rates (Online Appendix Table A1).

I first explore the role of mothers' origins. To that end, I compare second-generation immigrant women whose mothers were born in different counties:

$$\text{employed}_{imfr} = \beta \text{death_rate}_m + \alpha \mathbf{X}'_{1911,m} + \gamma_i + \delta_r + \mu_{1914,m} + \omega_f + \varepsilon_{imfr}, \quad (3)$$

where military death rates death_rate_m , historical controls $\mathbf{X}_{1911,m}$ and military region fixed effects $\mu_{1914,m}$ are assigned at the level of mothers' counties of birth. γ_i is a set of year of birth and survey year indicators. Consistent with the empirical strategy, I include county of residence fixed effects, δ_r . To neutralise the influence of fathers' origins, I further include fathers' county of birth fixed effects ω_f so as to compare women whose fathers were born in the same counties.³² I cluster standard errors at the levels of counties of residence and mothers' counties of birth. In this specification, β is identified from variations in the working behaviours of second-generation internal migrant married women of the same cohort, born and residing in the same county, whose fathers were born in the same county, but whose mothers were born in neighbouring counties that experienced different military death rates.

I then explore the role of fathers by estimating a specification analogous to Equation (3) in which variables are assigned at the level of fathers' counties of birth, and where I include mothers' county of birth fixed effects instead of fathers'.

Finally, to assess the relative role of each parent directly, I estimate Equation (3) without parental origin fixed effects, but including historical variables relative to both parents, distinguishing between women whose parents were born in the same county from those whose parents were born in different counties through the indicators same_{mf} and different_{mf} , where m indexes mothers, and f , fathers:

$$\begin{aligned} \text{employed}_{imfr} = & \beta_1 \text{same}_{mf} \cdot \text{death_rate}_m + \beta_2 \text{different}_{mf} \cdot \text{death_rate}_m \\ & + \beta_3 \text{different}_{mf} \cdot \text{death_rate}_f + \alpha_1 \mathbf{X}'_{1911,m} + \alpha_2 \mathbf{X}'_{1911,f} \\ & + \gamma_i + \delta_r + \varepsilon_{imfr}, \end{aligned}$$

where I use three-way clustering and cluster standard errors at the levels of counties of residence and mothers' as well as fathers' counties of birth.

4.1.2. Results

I report results for the mother-to-daughter transmission channel in Table 3.³³ In column (1), the estimate is significant at the 1% level and implies that women whose mothers were born in counties that experienced military death rates of 20% rather than 10% were 11 percentage points more likely to be working, which represents 13% of the outcome mean. Including fathers' county of birth fixed effects, to further isolate the role of mothers' origins, decreases the estimate only slightly, suggesting a primary role of maternal relative to paternal transmission (column (2)). Restricting the identifying variation to mothers born in neighbouring counties through mothers' military region of birth fixed effects similarly has little impact on the estimate (column (3)). Finally, I include controls for husband and household characteristics, parental characteristics

³² Because the sample size for the extended labour surveys is rather limited—15,091 second-generation internal migrant married women aged 30 to 59—I do not include interacted fixed effects. Doing so would generate more than 3,000 comparison groups, leaving too little residual variation for identification.

³³ Results are similar when the outcome is a labour force participant indicator (Online Appendix Table A2).

Table 3. *The Effect of Mothers' Birth County Military Death Rates on Their Daughters' Labour Force Participation.*

Dependent variable	Employed				
	(1)	(2)	(3)	(4)	(5)
Military death rate (mother)	1.05*** [0.36]	0.93** [0.38]	1.11*** [0.40]	1.25*** [0.34]	0.83** [0.37]
Father birth county FEs	No	Yes	Yes	Yes	Yes
Mother birth region FEs	No	No	Yes	Yes	Yes
Other controls	No	No	No	Yes	No
Clusters					
Own birth-residence county	95	95	95	95	95
Mother birth county	95	95	95	95	95
Observations	15,091	15,091	15,091	15,091	32,081
Outcome mean	0.81	0.81	0.81	0.81	0.79
Outcome SD	0.39	0.39	0.39	0.39	0.40

Notes: This table reports the OLS coefficients from estimating Equation (3) on the sample of non-migrant married women aged 30 to 59 whose mothers and fathers were born in another county in columns (1)–(4), and whose mothers (but not necessarily fathers) were born in another county in column (5). They result from regressing an employed indicator on the military death rate of respondents' mothers' counties of birth, including cohort, own county of birth and residence, and survey-year fixed effects, as well as the set of historical controls measured at the level of respondents' mothers' counties of birth. *Other controls* include parental, husband and household characteristics, and the respondent's education and number of children. Standard errors are clustered at the level of counties of residence and mothers' counties of birth. Estimates are computed using sample weights provided in the extended labour force surveys 2005–12. *** Significant at the 1% level. ** Significant at the 5% level.

as well as respondents' education and number of children (column (4)).³⁴ Again, the estimate seldom changes, which increases the credibility that economic factors potentially correlated with mothers' origins are not driving the results.

In column (5), I reproduce the estimate of column (3) using an alternative definition of second-generation internal migrants: I include women whose mothers were born in another county, but not necessarily their fathers. The resulting estimate is slightly lower than the baseline. Given that more than half of women in this sample had fathers born in their own county of birth, these 'half-natives' probably assimilated the local culture of their county, making the impact of their mothers' origins less influential.

Next, I reproduce the above analysis with paternal origins and find that women whose fathers were born in counties that experienced military death rates of 20% rather than 10% were 8 percentage points more likely to be working (Online Appendix Table A3, column 1). Once fixed effects for maternal origins are included along with fathers' military regions of birth fixed effects and parental, individual and household characteristics, the estimate is still positive, but declines substantially and is rather imprecise (columns (2)–(4)). The lesser role of paternal relative to maternal origins can be rationalised by the fact that, while fathers are primary role models for their sons, they are less so for their daughters, especially when it comes to gender roles (Farre and Vella, 2013).

The role of fathers nevertheless obtains to some extent once both parental origins are simultaneously considered (Online Appendix Table A4). Estimates in column (1) imply that women

³⁴ Husband and household characteristics consist of husbands' incomes, age, age squared, education, an indicator for home ownership and the number of rooms in the home; parental characteristics include indicators for parents' occupational status and mothers' labour status; education controls consist of indicators for educational attainment.

with both parents born in a county that experienced military death rates of 20% rather than 10% were 12 percentage points more likely to be working. Coefficients on mothers' and fathers' independent origins add up to the coefficient when both were born in the same county, suggesting that the impact of each parental origin cumulates linearly, with maternal origins being more important. Including parental regions of birth fixed effects along with various controls does not alter the results much (columns (2) and (3)).

The analysis in this section highlights the primary role of mothers in transmitting the legacy of the war to their daughters. Although there are many potential channels through which this mother-to-daughter transmission might have operated, a likely candidate is that mothers who entered the labour force because of the war provided a progressive role model to their daughters when growing up, which in turn altered their working behaviours when adult. Consistent with this idea, and focusing on the same sample of women as before, I find that mothers born in counties that experienced military death rates of 20% rather than 10% were 13–19 percentage points more likely to have been working when their daughters were growing up (Online Appendix Table A5).³⁵

4.2. *Transmission through Marriage*

4.2.1. *Empirical strategy*

After parental transmission, I explore a second channel: transmission through marriage, that is, from husbands to wives, and by extension, from mothers-in-law to daughters-in-law. Fernández *et al.* (2004) argue that sons of working mothers hold more progressive attitudes toward female labour than sons of stay-at-home mothers, making these men less averse to having a working wife, thereby incentivising their wives to enter the labour force. Following this line of reasoning, since men born in counties that experienced greater military death rates are more likely to have grown up with a working mother, they should be more likely to have a working wife. Women whose mothers-in-law were born in counties that experienced greater military death rates should also be more likely to work.

Two alternative mechanisms could account for these implications beyond the transmission from husbands' preferences to their wives' working behaviours through a process of joint labour supply decision. First, assortative matching: women from counties that experienced greater military death rates might be more likely to marry men from similar counties because of their higher initial propensity to work and these men's preference toward having working wives. Although I provide evidence of some degree of homogamy in military death rates in Online Appendix F, in the analysis below I make sure to compare wives who reside and were born in the same county through wives' county of birth fixed effects. This ensures that I am comparing women who carry a similar legacy of the war in their initial propensity to work. Second, these men might assume a larger share of household production, thereby freeing some time for their wives to enter the labour force. However, in Online Appendix G, I use data from the Generation

³⁵ If this was the only channel through which the mother-to-daughter transmission operated, one could use a 2SLS strategy to explore directly whether war-induced changes to mothers' working behaviours transmitted to their daughters—which I do in Online Appendix Table A6, finding that a 10 percentage points increase in mothers' labour force participation rates induced by WWI military fatalities generates a 5–8 percentage points increase in the likelihood of their daughters working. However, the exclusion restriction might be violated, for instance, if the war changed men's and women's beliefs toward female labour more broadly, thereby affecting those of younger women in these counties (see Section 5 for implications of the war on attitudes).

and Gender Survey 2005 to show that married men born in counties that experienced greater military death rates do not perform more household and childcare tasks.³⁶

Concretely, to assess the role of husbands' origins, I regress women's working behaviours on the military death rates exposure of their husbands' counties of birth. Importantly, I isolate the role of husbands' origins from that of their wives' by including wives' county of birth fixed effects so as to compare women born in the same county, but whose husbands were born in different counties. Using the censuses 1962 to 2012, I restrict the sample to women aged 30 to 49 married to first-generation internal migrant men and estimate

$$\text{employed}_{ihbrt} = \beta \text{death_rate}_h + \alpha \mathbf{X}'_{1911,h} + \theta_{i;r;b} + \varepsilon_{ihbrt}, \quad (4)$$

where death_rate_h denotes the military death rates exposure of the county of birth of husband h with wife i . $\theta_{i;r;b}$ are cohort by wives' county of birth by county of residence fixed effects. Here, β is identified from variations in the working behaviours of married women of the same cohort who reside and were born in the same county, but whose husbands were born in counties that experienced different military death rates.

Next, to explore the relative role of a woman's origins vis-à-vis her husband's directly, I distinguish between women whose husbands were born in the same county from those whose husbands were born in a different county through indicator variables same_{bh} and different_{bh} . Because this analysis applies the empirical strategy to both husbands and wives, I estimate the following specification on the sample of internal migrant women aged 30 to 49 married to internal migrant men:

$$\begin{aligned} \text{employed}_{ihbrt} = & \beta_1 \text{same}_{bh} \cdot \text{death_rate}_h + \beta_2 \text{different}_{bh} \cdot \text{death_rate}_b \\ & + \beta_3 \text{different}_{bh} \cdot \text{death_rate}_h + \alpha_1 \mathbf{X}'_{1911,b} + \alpha_2 \mathbf{X}'_{1911,h} \\ & + \theta_{i;r} + \varepsilon_{ihbrt}, \end{aligned} \quad (5)$$

where $\theta_{i;r}$ are cohort by county of residence fixed effects. I use three-way clustering and cluster standard errors at the levels of counties of residence and each spouse's county of birth.

Finally, I explore the role of mothers-in-law in transmitting the legacy of the war to their daughters-in-law through their sons by reproducing the analysis of parental transmission and estimating Equation (3) on the sample of second-generation internal migrant married women aged 30 to 59. To further isolate the role of mothers-in-law from that of own parents and husbands, I augment Equation (3) with fixed effects for both parents' and husbands' counties of birth so as to compare women who reside and were born in the same county, whose husbands and parents were born in the same county, but whose mothers-in-law were born in counties that experienced different military death rates.

4.2.2. *Transmission from husbands to wives*

I report coefficients from estimating Equation (4) in Figure 5.³⁷ First, I do not include wives' county of birth fixed effects. Estimates are significant at the 1%, level and imply that women whose husbands were born in counties that experienced military death rates of 20%, rather than

³⁶ This finding is consistent with those in other contexts. For instance, focusing on West Germany, Schmitz and Spiess (2022) find that, while 'men who grew up with a working mother develop different social norms and attitudes towards working women than other men, [...] these men still do not spend more time on unpaid work, such as housework or childrearing' (p. 283).

³⁷ Results are similar when the outcome is a labour force participant indicator (Online Appendix Figure A6).

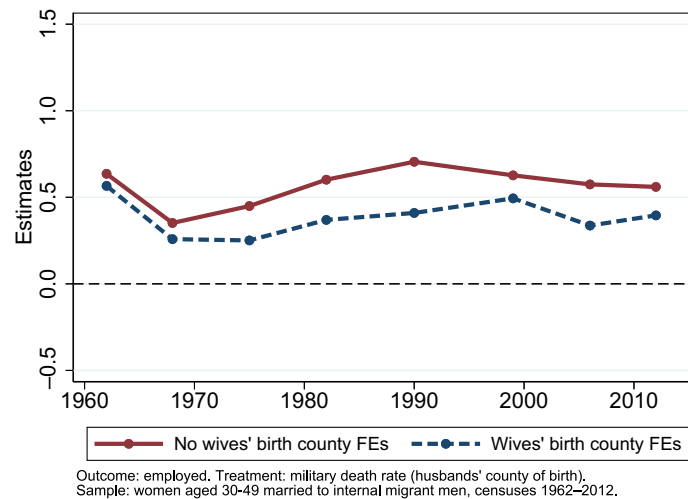


Fig. 5. *The Effect of Husbands' Birth County Military Death Rates on Their Wives' Labour Force Participation.*

Notes: This figure displays OLS coefficients from estimating Equation (4) on the sample of women aged 30 to 49 that are married to internal migrant men in the censuses 1962–2012. They result from regressing an employed indicator on the military death rate of respondents' husbands' counties of birth, along with county-level pre-war controls and cohort by wives' county of birth by county of residence fixed effects. Estimates are computed using sample weights provided in the censuses. Standard errors are clustered at the level of counties of residence and birth. All estimates are significant at the 1% level.

10%, were 4–7 percentage points more likely to be working between 1962 and 2012. Including wives' county of birth fixed effects to further compare women with identical origins decreases the estimates by 11%–44%, suggesting a limited role for homogamy in military death rates and assortative matching. Including husband and household controls as well as fertility and education barely affects the results (Online Appendix Figure A7). Moreover, restricting the sample to internal migrant women married to internal migrant men generates similar estimates (Online Appendix Figure A8).

Next, I explore the relative role of a woman's origins vis-à-vis her husband' and estimate Equation (5). I report estimates in Online Appendix Figure A9. Focusing on estimates when spouses' origins are different reveals that women's own origins are quantitatively twice as important as those of their husbands'.

4.2.3. *Transmission from mothers-in-law to daughters-in-law*

I now explore directly the role of mothers-in-law by reproducing the analysis of maternal transmission on the sample of second-generation internal migrant married women aged 30 to 59 along with fixed effects for both parents' and husbands' counties of birth. I report results in Table 4. The coefficient in column (1) implies that women whose mothers-in-law were born in counties that experienced military death rates of 20% rather than 10% were 13 percentage points more likely to be working. Including own parents' counties of birth fixed effects and restricting the identifying variation to mothers-in-law born in neighbouring counties slightly increases the estimate to 1.5 (columns (2) and (3)). Including controls for husband and household characteristics, parental characteristics as well as respondents' education and the number of children further increases the estimate to 1.6 (column (4)).

Table 4. *The Effect of Mothers-In-Law's Birth County Military Death Rates on Their Daughters-In-Law's Labour Force Participation.*

Dependent variable	Employed			
	(1)	(2)	(3)	(4)
Military death rate (mother-in-law)	1.32* [0.77]	1.24* [0.64]	1.50** [0.68]	1.64** [0.68]
Own parents birth county FEs	No	Yes	Yes	Yes
Mother-in-law birth region FEs	No	No	Yes	Yes
Other controls	No	No	No	Yes
Clusters				
Own birth-residence county	94	94	94	94
Mother-in-law birth county	94	94	94	94
Observations	13,387	13,387	13,387	13,387
Outcome mean	0.81	0.81	0.81	0.81
Outcome SD	0.39	0.39	0.39	0.39

Notes: This table reports the OLS coefficients from estimating Equation (3) on the sample of non-migrant married women aged 30 to 59 whose mothers and fathers were born in another county. They result from regressing an employed indicator on the military death rate of respondents' mothers-in-law's counties of birth, including cohort, own county of birth and residence, husband's county of birth and survey-year fixed effects, as well as the set of historical controls measured at the level of respondents' mothers-in-law's counties of birth. *Other controls* include parental, husband and household characteristics, and the respondent's education and number of children. Standard errors are clustered at the level of counties of residence and mothers-in-law's counties of birth. Estimates are computed using sample weights provided in the extended labour force surveys 2005–12. ** Significant at the 5% level. * Significant at the 10% level.

Estimates in Tables 3 and 4 suggest that mothers and mothers-in-law were equally important agents of intergenerational transmission of the war. Comparing their relative roles more directly by including military death rates of mothers' and mothers-in-law's origins together through a specification analogous to that of Online Appendix Table A4 confirms this interpretation. I report results in Online Appendix Table A7. Estimates in column (1) imply that women with mothers and mothers-in-law born in the same county that experienced military death rates of 20% rather than 10% were 25–31 percentage points more likely to be working. Coefficients on mothers' and mothers-in-law's independent origins add up to the coefficient when both come from the same county, suggesting again that the impact of each maternal origin cumulates linearly. While both maternal origins appear important, estimates for mothers-in-law are slightly larger than those for own mothers, although they are not significantly different.³⁸ This result complements those of Fernández *et al.* (2004), who find that mothers-in-law rather than own mothers explain the impact of WWII mobilisation rates on female labour force participation in the United States.

To conclude, although I use a strategy that isolates the role of husbands' and mothers-in-law's origins, these results remain only suggestive that they are driven by transmission from husbands' preferences to their wives' working behaviours through a process of joint labour supply decision. Indeed, assortative matching could still explain part of these results, whereby women who have a higher baseline propensity to work might be more likely to marry men from counties that experienced greater military death rates because of their preferences toward having working wives. Results in the literature nevertheless suggest that the former mechanism is more likely to

³⁸ The difference between both coefficients in column (1) is 0.08 with a standard error of 0.89; the difference in column (3) is 0.33 with a standard error of 0.75.

be at play.³⁹ It would also be consistent with research showing the predominant role of husbands in household decision making, for instance, for residence location decisions (Løken *et al.*, 2013; Brandén and Haandrikman, 2019).

4.3. Transmission through Local Social Interactions

4.3.1. Empirical strategy

The third set of mechanisms I explore is horizontal and oblique transmission through local social interactions. Culture is a locally embedded process and social interactions in culture might materialise in two ways: first, the cultural composition of a neighbourhood might affect the relative importance of an individual's own culture as individuals surrounded by many with identical origins might find it easier—or feel pressured—to preserve their culture of origin. Conversely, the cultural composition of a neighbourhood might affect individuals' beliefs through a process of cultural assimilation.

I first explore whether the legacy of the war diffused through cultural assimilation at the local level. For each census and municipality, I construct a measure of cultural composition in military death rates.⁴⁰ It is the weighted average military death rates exposure of counties of birth among residents in a municipality: $\text{death_rate}_m = \sum_o \text{sh_res}_{o,m} \times \text{death_rate}_o$, where $\text{sh_res}_{o,m}$ denotes the share of residents in municipality m born in county o . Then, I estimate the following specification on the sample of first-generation internal migrant married women aged 30 to 49:

$$\text{employed}_{ibmlt} = \beta \text{death_rate}_m + \alpha \mathbf{X}'_{1911,m} + \gamma_i + \delta_l + \omega_b + \varepsilon_{ibmlt}, \quad (6)$$

where $\mathbf{X}_{1911,m}$ contains pre-war characteristics and is constructed in the same way as death_rate_m . To compare individuals across neighbouring municipalities and alleviate potential issues of spatial sorting, I include local labour market fixed effects δ_l .⁴¹ I further include county of birth fixed effects ω_b to ensure that correlated differences in cultural origins are not driving the results. I cluster standard errors at the levels of counties of birth and residence. Here, β is identified from variations in the working behaviours of first-generation internal migrant married women of the same cohort, born in the same county, but residing in neighbouring municipalities characterised by different cultural compositions in military death rates.

Next, I analyse whether the legacy of the war was culturally preserved because of geographic clustering of individuals with similar origins. Again, I restrict the identifying variation to neigh-

³⁹ In particular, in the context of West Germany, Schmitz and Spiess (2022) find that husbands' preferences toward working women, due to the labour status of their own mothers, directly affect their wives' working behaviours rather than an assortative matching mechanism. They establish this result in two ways: examining 'whether single women who later have a partner who grew up with a working mother already exhibit different labour market outcomes before matching with their future partner' (p. 295) and observing the same women with different partners over time.

⁴⁰ Censuses contain information on individuals in up to 26,000 municipalities—there are 36,000 municipalities in France, but the remaining 10,000 municipalities are too small to be sampled in the censuses as they contain less than 200 residents.

⁴¹ Because the independent variable varies at the municipality level, I can use a narrow definition of local labour markets, with the identifying variation restricted to neighbouring municipalities. Specifically, I use *canton* fixed effects. Cantons represent the fourth level of subnational government, between *arrondissements* and *communes*. There were about 4,000 cantons before 2013, containing nine municipalities on average.

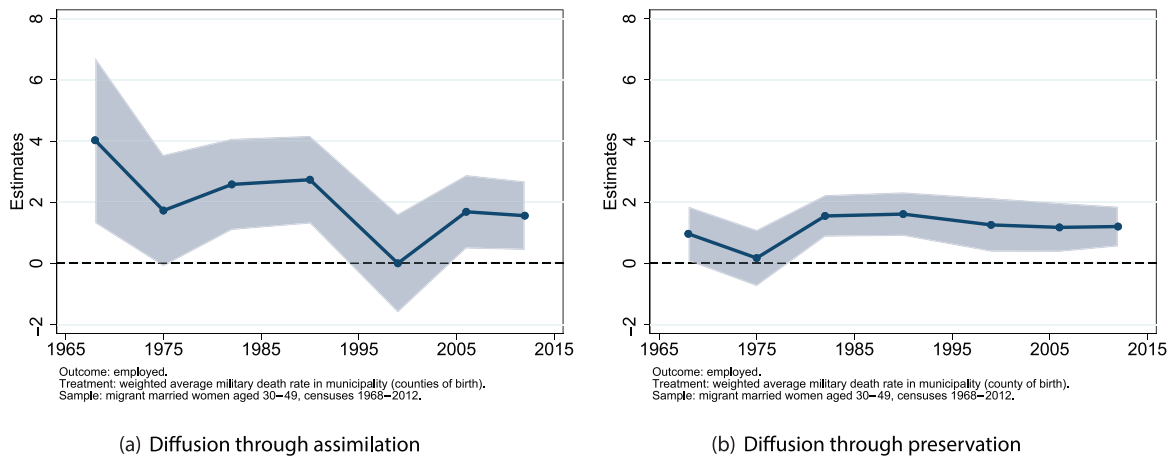


Fig. 6. *The Local Effects of WWI Military Death Rates on Female Labour Force Participation.*
 Notes: Panel (a) reports OLS coefficients on death_rate_m from estimating Equation (6), and panel (b), OLS coefficients on the interaction $\text{sh_res}_{b,m} \times \text{death_rate}_b$ from estimating Equation (7), along with 95% confidence intervals on the sample of internal migrant married women in the censuses 1968–2012. Standard errors are clustered at the level of respondents' counties of birth and residence. Estimates are computed using sample weights provided in the censuses.

bouring municipalities and estimate

$$\text{employed}_{ibmlt} = \beta \text{sh_res}_{b,m} \times \text{death_rate}_b + \alpha \text{sh_res}_{b,m} + \gamma_i + \delta_l + \omega_b + \varepsilon_{ibmlt}, \quad (7)$$

where $\text{sh_res}_{b,m}$ denotes the share of residents in municipality m born in county b , the same as the respondent's. This specification includes county of birth fixed effects ω_b , so that the coefficient on the interaction term β is identified from variations in the working behaviours of first-generation internal migrant married women of the same cohort, born in the same county, but residing in neighbouring municipalities. It indicates whether the impact of WWI military death rates was relatively stronger among women surrounded by more individuals from their own county of birth.

4.3.2. Results

I report results for the cultural diffusion channel in panel (a) of Figure 6. Estimates imply that women who reside in a municipality in which the average military death rates composition was 1 percentage point higher, were 2–4 percentage points more likely to be working between 1968 and 2012.⁴² Including controls for husband and household characteristics as well as respondents' education and the number of children barely affects the estimates (Online Appendix Figure A10, panel a).

I report results for the cultural preservation channel in panel (b) of Figure 6. Estimates imply that women born in the same county who reside in a municipality in which the share of residents from their own counties of birth was 1 percentage point higher, were 1–2 percentage

⁴² The coefficient is close to zero and not significant for the census of 1999, the census that contains the least observations in the sample—it is a 5% sample of the population. Because preferences and beliefs are formed while growing up, I assign the cultural composition using the previous census, so the census of 1962 is not part of the analysis.

points more likely to be working. Again, including controls barely affects the estimates (Online Appendix Figure A10, panel b).

Results in this section highlight the role of local social interactions for long-run historical processes.⁴³ While cultural assimilation affects the diffusion aspect of history, cultural preservation affects its persistence. Both channels empirically work toward the manifestation of the legacy of the war in the long run. But these channels need not oppose each other: merging Equations (6) and (7) results in estimates that change little, suggesting some degree of orthogonality between both processes (Online Appendix Figure A11). This sheds a new light on the dichotomy between culture and institutions: given the local embeddedness of culture and mechanisms of horizontal transmission through local social interactions, estimates from a naïve cross-county approach might partly capture implications of the local embeddedness of culture rather than solely differences in inherited institutional structures.

5. Changes to Preferences and Beliefs toward Female Labour

To account for the above results on working behaviours, I argue that women induced to enter the labour force during the interwar period because of WWI altered preferences and beliefs toward female labour of their daughters, sons and social relationships, and that these changes translated into the working behaviours of women in subsequent generations. Since individuals form preferences and beliefs early in life from learning and socialising with their parents, peers and neighbours (Bisin and Verdier, 2011; 2023), men and women who grew up with a working mother, or in an environment in which many women worked, should form more progressive attitudes toward female labour. To explore the validity of this argument, I analyse the long-run implications of WWI military death rates for attitudes toward gender roles.⁴⁴

5.1. Data and Empirical Strategy

Respondents to the Generation and Gender Survey 2005 were proposed three statements related to the role of women in the labour force and asked whether they ‘agree,’ ‘somewhat agree,’ ‘do not agree nor disagree,’ ‘somewhat disagree’ or ‘disagree’. Statements were (1) ‘If a woman earns more than her partner, it is bad for their relationship’, (2) ‘Women should not be able to decide how to spend the money they earned without asking their partners’ and (3) ‘In an economic crisis, men should keep their jobs in priority’. I assign 0 to ‘agree’ and 1 to ‘disagree’, and use 0.25-point increments for responses in between, so that higher values indicate more progressive attitudes toward gender roles. Then, I aggregate statements using a three-point scale, which I standardise to a one-point scale. Since I study preferences and beliefs rather than behaviours, I do not place age or gender restrictions on the regression sample, though I still focus on internal migrant respondents.⁴⁵ I report average responses for men and women separately in Online

⁴³ One caveat to the interpretation of the results in this section is that, despite comparing women from the same county of birth, it might still be possible that those with greater baseline labour force participation rates might have moved to municipalities of the same labour market with a greater local exposure to WWI military death rates within the resident population. Asserting this possibility would require information on pre-migration labour force participation, which is not available in the censuses.

⁴⁴ Before delving into contemporaneous data, it is worth noting that there were no pre-WWI differences in attitudes toward female labour across counties that experienced different military death rates (see Online Appendix B for details).

⁴⁵ As with the censuses, there is no correlation between migration status and origin county military death rates, employment status or fertility. However, internal migrants are more educated than non-migrants, as well as slightly more

Appendix Table A9. According to the above coding scheme, respondents ‘somewhat disagree’ with the statements on average with an index at 0.75, with no difference between men and women. The survey also contains questions related to preferences and beliefs about religion, marriage and family. I report average responses relative to these statements in Online Appendix Tables A10 to A12.

Applying the same approach as before, I isolate intergenerational cultural transmission by comparing respondents who reside in the same county, but were born in counties that were differentially affected by the war:

$$\text{values}_{ibr} = \beta_1 \text{death_rate}_b + \beta_2 \text{female}_i + \beta_3 \text{death_rate}_b \times \text{female}_i + \alpha \mathbf{X}'_{1911,b} + \gamma_i + \delta_r + \varepsilon_{ibr}, \quad (8)$$

where values_{ibr} denotes the one-point scale gender values index for respondent i , born in county b and residing in county r . Historical controls $\mathbf{X}_{1911,b}$ are assigned at the level of counties of birth. Consistent with the empirical strategy, I include county of residence fixed effects δ_r . I cluster standard errors at the levels of counties of birth and residence. Here, β_1 is identified from variations in attitudes held by respondents of the same cohort, who reside in the same county, but were born in counties that experienced different military death rates.

5.2. Results

I report results in Table 5. When the interaction term is not included, the coefficient of interest is significant at the 1% level and implies that respondents born in counties that experienced military death rates of 20% rather than 10% hold more progressive attitudes toward female labour (column (1)): the index is 10 percentage points higher in this case, which corresponds to 13% of the mean. Adding the interaction term only slightly reduces and reveals a somewhat stronger response of women’s attitudes, though the coefficient is not significant (column (2)). Adding parental and household controls and controlling for respondents’ employment status, education and fertility does not change the results substantially (columns (3)–(5)).

Estimating Equation (8) separately for each component of the gender values index reveals that the effect of military death rates on preferences and beliefs is driven by two of its three components (Online Appendix Table A13): whether respondents disagree with the statements ‘if a woman earns more than her partner, it is detrimental for their relationship’—a strong marker of gender identity norms in relation with female labour supply (Bertrand *et al.*, 2015)—and ‘in an economic crisis, men should keep their jobs in priority’—another important marker of gender norms (Alesina *et al.*, 2013, pp. 524–7).

Finally, to assess whether these results simply reflect more progressive attitudes in general, I reproduce the analysis for attitudes toward religion, marriage and the family (Online Appendix Table A14). Consistent with my argument, WWI military death rates did not impact attitudes beyond those specifically related to the role of women in the labour force.

To take stock of the set of findings in this article and put their magnitudes in perspective, I report the main estimates of the analysis altogether in Table 6, where I display their median whenever there are multiple estimates over time for a given specification. Finding a common

progressive. As before, including an interaction term between the internal migrant indicator and the military death rate of respondents’ counties of birth, reveals that selection along all observable characteristics does not operate differently across counties of origin (Online Appendix Table A8).

Table 5. *The Effect of WWI Military Death Rates on Beliefs About Female Labour.*

Dependent variable	Gender values index (one-point scale)				
	(1)	(2)	(3)	(4)	(5)
Military death rate	1.01*** [0.20]	0.90*** [0.29]	0.83*** [0.31]	0.85*** [0.31]	0.84*** [0.32]
Female	0.02** [0.01]	−0.00 [0.03]	−0.01 [0.03]	−0.01 [0.03]	−0.01 [0.03]
Military death rate × female		0.17 [0.21]	0.21 [0.21]	0.19 [0.21]	0.24 [0.22]
Controls					
Parental	No	No	Yes	Yes	Yes
Household	No	No	No	Yes	Yes
Individual	No	No	No	No	Yes
Clusters					
Residence counties	92	92	92	92	92
Birth counties	88	88	88	88	88
Observations	2,822	2,822	2,822	2,822	2,822
Outcome mean	0.75	0.75	0.75	0.75	0.75
Outcome SD	0.21	0.21	0.21	0.21	0.21

Notes: This table reports the OLS coefficients from estimating Equation (8) on the sample of internal migrant respondents in the Gender and Generation Survey 2005. All regressions contain cohort and county of residence fixed effects, as well as the set of historical controls measured at the level of respondents' counties of birth in 1911. Household controls include an indicator for whether the respondent's home is a house rather than an apartment, the number of rooms in the home, an indicator for whether the respondent owns her housing and an indicator for whether the respondent has a partner present in the household. Parental controls include the labour status of the respondent's mother and educational attainment indicators for both parents. Individual controls include employment status, educational attainment and the number of children in the household. Standard errors are clustered at the level of counties of birth and of residence. Estimates are computed using sample weights provided in the Gender and Generation Survey 2005. *** Significant at the 1% level. ** Significant at the 5% level.

ground for comparison is challenging as samples and sources of identification differ throughout. One possibility is to standardise estimates in terms of residual standard deviations to compute residual magnitudes. That is, rely on the effective variation in the treatment variable used for identification once fixed effects are taken into account, multiply raw estimates with these residual standard deviations, then divide them with the residual standard deviation of the outcome variable calculated in the same way (Mummolo and Peterson, 2018). To make this exercise meaningful, I compare regression coefficients estimated across samples of the same generation: first-generation internal migrants in panel A and second-generation internal migrants in panel B.

Focusing first on results relying on first-generation internal migrants (panel A), marriage appears as an important mechanism of persistence, as the residual magnitude of coefficients identified through husbands' origins is close to the residual magnitude of the overall effect. Both mechanisms of social interactions also appear important for persistence—though their sources of variation are markedly different from that of other analyses in this article, as they rely on variation in cross-city population composition. Finally, the residual magnitude on preferences and beliefs is large, suggesting that changes to attitudes of both men and women are an important precondition for changes to behaviours. Turning to analyses relying on second-generation migrants (panel B), residual magnitudes of estimates on mothers and mothers-in-law confirm that they were equally

Table 6. *Estimates' Magnitudes.*

Analysis	Dataset	Year	Source of variation	Estimate			Reference
				Raw	Standardised	Residual	
Baseline	Censuses	1962–2012	Birth county	0.48***	0.044***	0.011***	Figure 4
Baseline	Labour surveys	1982–2012	Birth county	0.69***	0.065***	0.018***	Table E.3, column (1)
Husbands	Censuses	1962–2012	Husband birth county	0.38***	0.035***	0.011***	Figure 5
Assimilation	Censuses	1968–2012	Residence city	1.73***	0.108***	0.006***	Figure 6(a)
Preservation	Censuses	1968–2012	Residence city	1.20***	0.035***	0.008***	Figure 6(b)
Attitudes	GGs	2005	Birth county	0.90***	0.167***	0.058***	Table 5, column (2)

Analysis	Dataset	Year	Source of variation	Estimate			Reference
				Raw	Standardised	Residual	
Mother	Labour surveys	2005–12	Mother birth county	1.11***	0.120***	0.030***	Table 3, column (3)
Father	Labour surveys	2005–12	Father birth county	0.26	0.029	0.007	Table A.3, column (3)
Mother-in-law	Labour surveys	2005–12	Mother-in-law birth county	1.50**	0.146**	0.029**	Table 4, column (3)

Notes: This table reports the main estimates in this article. *Standardised estimates* correspond to raw estimates multiplied by the sample standard deviation in the independent variable, then divided by the sample standard deviation in the dependent variable. *Residual estimates* correspond to raw estimates multiplied by the residual standard deviation in the independent variable, then divided by the residual standard deviation in the dependent variable. Variables are residualised with respect to specification-specific fixed effects. When multiple estimates are available for a given specification, due to multiple time points, estimates correspond to the median.

*** Significant at the 1% level. ** Significant at the 5% level.

important agents in the transmission of the legacy of the war, echoing results in Fernández *et al.* (2004).⁴⁶ In contrast, fathers appear to play a limited role in this transmission process.

6. Conclusion

One century after the First World War, its legacy on female labour in France is still vivid. Comparing women who reside under similar institutional conditions, but were born in counties that experienced different military death rates, I provide empirical evidence for a persistent and continued impact of World War I on women's working behaviours throughout the second half of the twentieth century. I uncover three channels of intergenerational transmission: transmission through parents, transmission through marriage and transmission through local social interactions. Consistent with models of intergenerational transmission, I find that men and women born in counties that experienced greater military death rates hold more progressive attitudes toward the role of women in the labour force today.

Findings in this article elucidate the intergenerational diffusion pathways of women's labour market involvement over the long run. They suggest that the entrance of women in the labour force in the period directly preceding revolution of female labour paved the way for subsequent generations of women. I interpret them as evidence for a process of cultural diffusion and change, by which women induced to enter the labour force during the interwar period because of the war altered preferences and beliefs about female labour of their daughters, sons and social relationships, and that these changes translated into the working behaviours of women in subsequent generations. Providing direct empirical evidence for this feedback process is challenging as no survey traces back to the interwar period. Finding innovative ways to build measures of preferences and beliefs toward female labour far back in the past to explore further this process of cultural change is a promising avenue of research.

University of Toulouse Capitole, France

Additional Supporting Information can be found in the online version of this article:

Online Appendix Replication Package

References

- Abramitzky, R., Delavande, A. and Vasconcelos, L. (2011). 'Marrying up: The role of sex ratio in assortative matching', *American Economic Journal: Applied Economics*, vol. 3(3), pp. 124–57.
- Acemoglu, D., Autor, D.H. and Lyle, D. (2004). 'Women, war, and wages: The effect of female labor supply on the wage structure at midcentury', *Journal of Political Economy*, vol. 112(3), pp. 497–551.
- Acemoglu, D., Johnson, S. and Robinson, J.A. (2005). 'Institutions as a fundamental cause of long-run growth', in (P. Aghion and S.N. Durlauf, eds.), *Handbook of Economic Growth*, vol. 1, pp. 385–472, Amsterdam: Elsevier.
- Albanesi, S. and Olivetti, C. (2016). 'Gender roles and medical progress', *Journal of Political Economy*, vol. 124(3), pp. 650–95.
- Alesina, A., Giuliano, P. and Nunn, N. (2013). 'On the origins of gender roles: Women and the plough', *The Quarterly Journal of Economics*, vol. 128(2), pp. 469–530.
- Alix-Garcia, J., Schechter, L., Valencia Caicedo, F. and Jessica Zhu, S. (2022). 'Country of women? Repercussions of the triple alliance war in Paraguay', *Journal of Economic Behavior and Organization*, vol. 202, pp. 131–67.

⁴⁶ The fact that these are larger the overall effects might be the result of the smaller amount of measurement error in these analyses, as results in Online Appendix Table C2 attest.

- Attanasio, O., Low, H. and Sánchez-Marcos, V. (2008). 'Explaining changes in female labor supply in a life-cycle model', *The American Economic Review*, vol. 98(4), pp. 1517–52.
- Bailey, M.J. (2006). 'More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply', *The Quarterly Journal of Economics*, vol. 121(1), pp. 289–320.
- Bertrand, M., Kamenica, E. and Pan, J. (2015). 'Gender identity and relative income within households', *The Quarterly Journal of Economics*, vol. 2(130), pp. 571–614.
- Bette, P. (2017). *Veuves de la grande guerre. Itinéraires et combats*, Berine: Peter Lang.
- Bisin, A. and Verdier, T. (2011). 'The economics of cultural transmission and socialization', in (J. Benhabib, A. Bisin and M.O. Jackson, eds.), *Handbook of Social Economics*, vol. 1, pp. 339–416, Amsterdam: North-Holland.
- Bisin, A. and Verdier, T. (2023). 'Advances in the economic theory of cultural transmission', *Annual Review of Economics*, vol. 15, doi: 10.1146/annurev-economics-090622-100258.
- Blau, F.D., Kahn, L.M. and Papps, K.L. (2011). 'Gender, source country characteristics, and labor market assimilation among immigrants', *The Review of Economics and Statistics*, vol. 93(1), pp. 43–58.
- Boehnke, J. and Gay, V. (2020). 'Replication data for: The missing men. World War I and female labor force participation', Harvard Dataverse, V3, <https://doi.org/10.7910/DVN/APIHZ8>.
- Boehnke, J. and Gay, V. (2022). 'The missing men. World War I and female labor force participation', *Journal of Human Resources*, vol. 57(4), pp. 1209–41.
- Boggiano, B. (2020). 'Long-term effects of the Paraguayan war (1864–1870): from male scarcity to intimate partner violence', Discussion Paper 20–024, ZEW.
- Bostrom, A. (2016). 'Supplying the front French artillery production during the first world war', *French Historical Studies*, vol. 39(2), pp. 261–86.
- Boulanger, P. (2001). *La France devant la conscription: Géographie historique d'une institution républicaine, 1914–1922*, Paris: Economica.
- Brandén, M. and Haandrikman, K. (2019). 'Who moves to whom? Gender differences in the distance moved to a shared residence', *European Journal of Population*, vol. 35, pp. 435–58.
- Brilhault, G. and Caron, N. (2016). 'Le passage à une collecte par sondage: Quel impact sur la précision du recensement?', *Economie et Statistique*, vol. 483–5, pp. 23–40.
- Brodeur, A. and Kattan, L. (2022). 'World War II, the baby boom, and employment: County-level evidence', *Journal of Labor Economics*, vol. 40(2), pp. 437–71.
- Bütikofer, A. (2013). 'Revisiting “mothers and sons” preference formation and the female labor force in Switzerland', *Labour Economics*, vol. 20, pp. 82–91.
- Campos-Vazquez, R.M. and Velez-Grajales, R. (2014). 'Female labour supply and intergenerational preference formation: Evidence for Mexico', *Oxford Development Studies*, vol. 42(4), pp. 553–69.
- Cantoni, D., Hagemeister, F. and Westcott, M. (2020). 'Persistence and activation of right-wing political ideology', Discussion Paper 143, ZBW.
- Cardoso, A.R. and Morin, L.P. (2018). 'Can economic pressure overcome social norms? The case of female labor force participation', Working Paper 11822, IZA.
- Chen, X.C. and Ge, S. (2018). 'Social norms and female labor force participation in urban china', *Journal of Comparative Economics*, vol. 46(4), pp. 966–87.
- Clemens, M.A. (2017). 'The meaning of failed replications: A review and proposal', *Journal of Economic Surveys*, vol. 31(1), pp. 326–42.
- Commission Consultative des Dommages et des Réparations. (1951). *Dommages subis par la France et l'Union Française du fait de la guerre et de l'occupation ennemie (1939–1945)*, Paris: Imprimerie Nationale.
- Doepke, M., Hazan, M. and Maoz, Y.D. (2015). 'The baby boom and World War II: A macroeconomic analysis', *The Review of Economic Studies*, vol. 82(3), pp. 1031–73.
- Farre, L. and Vella, F. (2013). 'The intergenerational transmission of gender role attitudes and its implications for female labour force participation', *Economica*, vol. 80(318), pp. 219–47.
- Fernández, R. (2011). 'Does culture matter?', in (J. Benhabib, A. Bisin and M.O. Jackson, eds.), *Handbook of Social Economics*, vol. 1, pp. 481–510, Amsterdam: North-Holland.
- Fernández, R. (2013). 'Cultural change as learning: The evolution of female labor force participation over a century', *The American Economic Review*, vol. 103(1), pp. 472–500.
- Fernández, R. and Fogli, A. (2009). 'Culture: An empirical investigation of beliefs, work, and fertility', *American Economic Journal: Macroeconomics*, vol. 1(1), pp. 146–77.
- Fernández, R., Fogli, A. and Olivetti, C. (2004). 'Mothers and sons: Preference formation and female labor force dynamics', *The Quarterly Journal of Economics*, vol. 119(4), pp. 1249–99.
- Ferrara, A. (2022). 'World War II and black economic progress', *Journal of Labor Economics*, vol. 40(4), pp. 1053–91.
- Flaus, L. (1947). 'Les destructions d'immeubles d'habitation et d'usines en France au cours de la Guerre 1939–1945', *Journal de la Société Française de Statistique*, vol. 88, pp. 353–81.
- Fogli, A. and Veldkamp, L. (2011). 'Nature or nurture? Learning and the geography of female labor force participation', *Econometrica*, vol. 79(4), pp. 1103–38.
- Fouka, V. and Voth, H.J. (2022). 'Collective remembrance and private choice: German-Greek conflict and consumer behavior in times of crisis', *American Political Science Review*, doi:10.1017/S0003055422001095.

- Gay, V. (2020a). 'TRF-GIS départements (1870–1940) [database]', Harvard Dataverse, <https://doi.org/10.7910/DVN/ULQYM5>.
- Gay, V. (2020b). 'TRF-GIS military regions (1874–1940) [database]', Harvard Dataverse, <https://doi.org/10.7910/DVN/SQPEUW>.
- Gay, V. (2021). 'Mapping the Third Republic. A geographic information system of France (1870–1940)', *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, vol. 54(4), pp. 189–207.
- Gay, V., Hicks, D.L., Santacreu-Vasut, E. and Shoham, A. (2018). 'Decomposing culture: An analysis of gender, language, and labor supply in the household', *Review of Economics of the Household*, vol. 16(4), pp. 879–909.
- Girard, A., Bastide, H. and Pourcher, G. (1964). 'Mobilité géographique et concentration urbaine en France. Une enquête en province', *Population*, vol. 19(2), pp. 227–66.
- Giuliano, P. (2017). 'Gender: A historical perspective', in (S.L. Averett, L.M. Argys and S.D. Hoffman, eds.), *The Oxford Handbook of Women and the Economy*, pp. 645–72, New York: Oxford University Press.
- Giuliano, P. and Nunn, N. (2021). 'Understanding cultural persistence and change', *Review of Economic Studies*, vol. 88, pp. 1514–81.
- Goldin, C. (1990). *Understanding the Gender Gap: An Economic History of American Women*, New York: Oxford University Press.
- Goldin, C. (2006). 'The quiet revolution that transformed women's employment, education, and family', *American Economic Review*, vol. 96(2), pp. 1–21.
- Goldin, C. (2014). 'A grand gender convergence: Its last chapter', *The American Economic Review*, vol. 104(4), pp. 1091–119.
- Goldin, C. and Olivetti, C. (2013). 'Shocking labor supply: A reassessment of the role of World War II on women's labor supply', *The American Economic Review*, vol. 103(3), pp. 257–62.
- Gonzalez-Feliu, J. and Parent, A. (2016). 'Clio-combinatorics: A novel framework to analyze military logistics choices using operations research techniques', in (A. Ochoa-Zazzatti, J. Sánchez, M.G. Cedillo-Campos and M. de Lourdes, eds.), *Handbook of Research on Military, Aeronautical, and Maritime Logistics and Operations*, pp. 79–101, Hershey, PA: IGI Global.
- Greenwood, J., Seshadri, A. and Yorukoglu, M. (2005). 'Engines of liberation', *The Review of Economic Studies*, vol. 72(1), pp. 109–33.
- Grosjean, P. and Khattar, R. (2019). 'It's raining men! Hallelujah? The long-run consequences of male-biased sex ratios', *The Review of Economic Studies*, vol. 86(2), pp. 723–54.
- Hazan, M. and Maoz, Y.D. (2002). 'Women's labor force participation and the dynamics of tradition', *Economics Letters*, vol. 75(2), pp. 193–8.
- Hiller, V. and Baudin, T. (2016). 'Cultural transmission and the evolution of gender roles', *Mathematical Social Sciences*, vol. 84, pp. 8–23.
- Huber, M. (1931). *La population de la France pendant la guerre*, New Haven, CT: Yale University Press.
- INSEE. (1953). *Résultats statistiques du recensement général de la population du 10 mars 1946: Population présente totale*, Paris: Imprimerie Nationale.
- INSEE. (1962). 'Recensement de la population 1962: fichier détail au 1/20 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (1968). 'Recensement de la population 1968: fichier détail au 1/4 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (1975). 'Recensement de la population 1975: fichier détail au 1/5 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (1982). 'Recensement de la population 1982: fichier détail au 1/4 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (1989). 'Enquêtes emploi 1982–1989 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (1990). 'Recensement de la population 1990: fichier détail au 1/4 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (1999). 'Recensement de la population 1999: fichier détail au 1/20 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (2002). 'Enquêtes emploi 1990–2002 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (2006). 'Recensement de la population: fichier détail 2006 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (2008). 'Table d'appartenance géographique des communes au 1er janvier 2008 [database]', Paris: Institut National de la Statistique et des Études Économiques (INSEE).
- INSEE. (2010). 'Données historiques de la statistique générale de France. SGF 1800–1925. T226 – recensements de 1911 [database]', Paris: Institut National de la Statistique et des Études Économiques (INSEE).
- INSEE. (2012a). 'Enquêtes emploi 2003–2012 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- INSEE. (2012b). 'Recensement de la population: fichier détail 2012 [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).

- INSEE. (2012c). 'Table d'appartenance géographique des communes au 1er janvier 2012 [database]', Paris: Institut National de la Statistique et des Études Économiques (INSEE).
- INSEE. (2018). 'Transformation coefficient from one year's euro or French franc to another year's euro or French franc—base 2015 [database]', Paris: Institut National de la Statistique et des Études Économiques (INSEE).
- INSEE. (2019). 'Structure de la population active de 25 à 54 ans selon la catégorie socioprofessionnelle, selon le niveau de diplôme, selon le sexe ou le secteur d'activité (1968 à 2019) [database]', Paris: Institut National de la Statistique et des Études Économiques (INSEE).
- INSEE and INED. (2005). 'Étude des relations familiales et intergénérationnelles (ERFI) [database]', Paris: Archives de Données Issues de la Statistique Publique (ADISP).
- Johnston, D.W., Schurer, S. and Shields, M.A. (2014). 'Maternal gender role attitudes, human capital investment, and labour supply of sons and daughters', *Oxford Economic Papers*, vol. 66(3), pp. 631–59.
- Kawaguchi, D. and Miyazaki, J. (2009). 'Working "mothers and sons" preferences regarding female labor supply: Direct evidence from stated preferences', *Journal of Population Economics*, vol. 22(1), pp. 115–30.
- Lagrou, P. (2002). 'Les guerres, les morts et le deuil: Bilan chiffré de la Seconde Guerre Mondiale', in (S. Audouin-Rouzeau, A. Becker, C. Ingrao and H. Rousso, eds.), *La Violence de guerre, 1914–1945. Approches comparées des deux conflits mondiaux*, pp. 313–27, Bruxelles: Complexe.
- Li, Z. and Liu, L. (2019). 'Preference or endowment? Intergenerational transmission of women's work behavior and the underlying mechanisms', *Journal of Population Economics*, vol. 32(4), pp. 1401–35.
- Løken, K.V., Lommerud, K.E. and Shelly, L. (2013). 'Your place or mine? On the residence choice of young couples in Norway', *Demography*, vol. 50(1), pp. 285–310.
- Maruani, M. and Meron, M. (2012). *Un siècle de travail des femmes en France: 1901–2011*, Paris: La Découverte.
- Michalopoulos, S. and Papaioannou, E. (2017). *The Long Economic and Political Shadow of History*, London: CEPR Press.
- Michalopoulos, S., Putterman, L. and Weil, D.N. (2019). 'The influence of ancestral lifeways on individual economic outcomes in Sub-Saharan Africa', *Journal of the European Economic Association*, vol. 17(4), pp. 1186–231.
- Miyata, S. and Yamada, H. (2016). 'Do female gender role attitudes affect labour market participation in Egypt?', *The Journal of Development Studies*, vol. 52(6), pp. 876–94.
- Morrill, M.S. and Morrill, T. (2013). 'Intergenerational links in female labor force participation', *Labour Economics*, vol. 20, pp. 38–47.
- Mummolo, J. and Peterson, E. (2018). 'Improving the interpretation of fixed effects regression results', *Political Science Research and Methods*, vol. 6(4), pp. 829–35.
- Ngai, R.L. and Petrongolo, B. (2017). 'Gender gaps and the rise of the service economy', *American Economic Journal: Macroeconomics*, vol. 9(4), pp. 1–44.
- Nunn, N. (2014). 'Historical development', in (P. Aghion and S.N. Durlauf, eds.), *Handbook of Economic Growth*, vol. 2, pp. 347–402, Amsterdam: Elsevier.
- Nunn, N. and Wantchekon, L. (2011). 'The slave trade and the origins of mistrust in Africa', *American Economic Review*, vol. 101(7), pp. 3221–52.
- Ochsner, C. and Roesel, F. (2017). 'Activated history—the case of the Turkish sieges of Vienna', Working Paper 6586, CESifo.
- Olivetti, C., Patacchini, E. and Zenou, Y. (2020). 'Mothers, peers, and gender-role identity', *Journal of the European Economic Association*, vol. 18(1), pp. 266–301.
- Olivetti, C. and Petrongolo, B. (2016). 'The evolution of gender gaps in industrialized countries', *Annual Review of Economics*, vol. 8, pp. 405–34.
- Porte, R. (2005). *La mobilisation industrielle: 'premier front' de la Grande Guerre?*, Paris: 14–18 Éditions.
- Régnier-Loilier, A. (2016). *Parcours de familles: l'enquête Etude des Relations Familiales et Intergénérationnelles*, Paris: INED.
- Rose, E.K. (2018). 'The rise and fall of female labor force participation during World War II in the United States', *The Journal of Economic History*, vol. 78(3), pp. 673–711.
- Schmitz, S. and Spiess, C.K. (2022). 'The intergenerational transmission of gender norms—why and how adolescent males with working mothers matter for female labour market outcomes', *Socio-Economic Review*, vol. 20(1), pp. 281–322.
- Teso, E. (2019). 'The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade', *Journal of the European Economic Association*, vol. 17(2), pp. 497–534.



Contents lists available at ScienceDirect

Explorations in Economic History

journal homepage: www.elsevier.com/locate/eeh

Shorter Article

Morts Pour la France: A database of French fatalities of the Great War[☆]Victor Gay^{a,*}, Pauline Grosjean^b^a Department of Social and Behavioral Sciences, Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France^b University of New South Wales and CEPR, Sydney, Australia

ARTICLE INFO

Dataset link: <https://doi.org/10.7910/DVN/2A PQZ6>

JEL classification:

N44

D74

J11

Keywords:

World war I

Military fatality

France

ABSTRACT

This article describes the construction and content of the *Morts pour la France* database. This database contains individual-level data on the universe of the 1.3 million French fatalities of the Great War who were officially recognized as war victims. It provides information on each soldier's first and last names, dates of birth and death, circumstances of death, recruitment status, military rank and unit, and locations of birth and recruitment. We further provide datasets of military death rates at the level of 87 départements and 35 thousand municipalities along with relevant shapefiles and discuss potential applications of this database.

1. Introduction

World War I was a watershed moment for French society. Over four years of battle, 8.3 million French men served in the army and 1.3 million among them died in combat, representing close to 16 percent of the male population of conscripts (Prost, 2008). The imprint of the conflict on France's demography was profound. The war depleted several cohorts of young men and cut a lasting gash in the French population pyramid (Héran, 2014; Bellis et al., 2022). Fertility plummeted and its modest post-war uptick hardly replaced the million "missing" children that were never born (Rohrbasser, 2014). The gender imbalance resulting from the war disrupted marriage markets (Abramitzky et al., 2011) with dramatic implications on the economy writ large, as an entire generation of war widows and single women was drawn into the labor force (Boehnke and Gay, 2022). This labor supply shock perpetuated across generations until today (Gay, 2023). Politically, the Great War may have set the seeds of France's rapid defeat to Germany in 1940 (Bloch, 1946) and some of its features help explain collaboration with Nazi Germany upon Philippe Pétain's accession to power (Cagé et al., 2023).

This article describes a novel database that will help improve our understanding of the demographic, economic, and political consequences the Great War on French society: the *Morts pour la France* (MPF) database ("Dead for France"). The MPF database includes a host of individual-level information for each of the 1.3 million French individuals who died as a result of the war: their first and last names, dates of birth and death, circumstances of death, recruitment status, military rank and unit at the time of death,

[☆] We are grateful to participants to the virtual workshop "Life and Death in the WW1 Trenches: Causes and Consequences" held on August 6, 2021 at the Australian National University for helpful feedback. We thank Rhea Banerjee, Matthias Besombes, Garance Casset, Morgane Fridlin, Caroline Marchand, Sofia Perez, and Andrea Vargas for excellent research assistance. Pauline Grosjean acknowledges funding from the Australian Research Council (FT190100298). Victor Gay acknowledges funding from the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

* Corresponding author.

E-mail addresses: victor.gay@tse-fr.eu (V. Gay), p.grosjean@unsw.edu.au (P. Grosjean).

<https://doi.org/10.1016/j.eeh.2023.101550>

Received 12 June 2023; Received in revised form 5 September 2023; Accepted 13 September 2023

Available online 23 September 2023

0014-4983/© 2023 Elsevier Inc. All rights reserved.

and locations of birth and recruitment. In particular, location information in the MPF database is available at the municipality level, France's most disaggregated administrative unit.

The unprecedented granularity of the MPF database opens several avenues for research into the characteristics and consequences of the conflict. In particular, we provide a dataset of military death rates at the level of 35 thousand municipalities, which dramatically improves upon the sources of geographic variation that have been used to identify the consequences of the war. Up to now, researchers had nearly exclusively relied on subsamples of the underlying MPF data aggregated at the level of 87 départements or 21 military regions (Gilles, 2010; Abramitzky et al., 2011; Gilles et al., 2014; Beau, 2017; Guironnet and Parent, 2019; Boehnke and Gay, 2022; Gay, 2023).¹

The MPF database will also further improve our understanding of battle-specific violence, as it contains information on circumstances and dates of death as well as the military ranks of deceased soldiers. For instance, we document that the deadliest day of French history was September 25, 1915, when 22,059 French soldiers died on the first day of the (ultimately failed) Second Battle of Champagne. Previous historiography instead supported that the deadliest day of the conflict occurred on August 22, 1914, in the context of the Battle of the Frontiers (Steg, 2013) while other advanced instead the Battle of the Marne in early September 1914 as the deadliest day (Guillot and Parent, 2018). The precision and exhaustiveness of our database will also improve our understanding of the sociodemographic consequences of the conflict, which so far have mostly been studied based on subsamples (see footnote 1). As an example, in Section 4, we extend Guillot and Parent's (2018) analysis of the determinants of soldiers' survival duration – based on a sample of 18 thousand deceased soldiers – to the full database. Finally, combined information on soldiers' death and rank will enable economic historians to study organizational aspects of the conflict, in the spirit of Ager et al. (2022), who show how public recognition induced a shift in the mortality profile of German fighter pilots during World War II, or Dippel and Ferrara (2023), who study the effectiveness of military leadership in the context of the US Civil War.

There has been comparable efforts to study the implications of the war using individual-level military fatalities information in other European belligerent countries. In the case of Weimar Germany, De Juan et al. (2023) study the effect of WWI fatalities on the rise of and support for the Nazi Party. They rely on (partly) newly digitized individual-level data on 8.6 million German soldiers who died, were wounded, captured, or went missing during the war.² In the case of Italy, Fornasin (2017) uses the *Albo d'oro* register of military deaths to construct estimates of war-related fatalities rates, while Acemoglu et al. (2022) explore the implications of war fatalities on the support for Socialism and Fascism. Finally, in the case of Britain, Carozzi et al. (2023) explore the implications of WWI fatalities on civic capital and combat motivation in WWII.

This article also relates to the literature focusing on the characteristics of WWI servicemen to study the selection effects of mortality among soldiers—see, e.g., Fornasin et al. (2019) on Italian soldiers, Kanazawa (2007) and Bailey et al. (2023) on British soldiers, and McCalman et al. (2019) on Australian soldiers. More broadly, this article relates to the literature that uses individual-level military records to study health, wealth, and migration of historical populations (Costa and Kahn, 2008; Cranfield and Inwood, 2015; Bailey et al., 2016; Costa et al., 2018b,a; Cunningham, 2018; Fourie et al., 2020; Salem, 2023). The largest effort to date in this literature concerns the Union Army during the American Civil War (Fogel et al., 2000; Dippel and Heblich, 2021; Weaver, 2022).³ Similar data collection efforts are now being undertaken for the Confederate Army (Hall et al., 2019).

In the rest of this article, we first provide some historical context about the organization of the French Army and the archival source of the MPF data (Section 2). We then describe the content of the database (Section 3) and discuss potential uses of the database (Section 4). Finally, we provide technical information on data availability (Section 5). Section 6 concludes.

2. Historical context

This section first discusses the organization of the French Army during the war (Section 2.1) then provides an overview of the construction process of the archival source of the MPF database (Section 2.2).⁴

¹ Related studies using subsamples of the same data source as that of the MPF database include Rolland (2005) for an analysis of suicides, Le Gall (2013) for the geography of recruitment, Guillot and Parent (2018) for survival rates, Moura (2018) for the characteristics of soldiers who died on battlefields in Meuse and Somme, and Darlu and Chareille (2020) for the disappearance of last names. Vandenbroucke (2014) and Knowles and Vandenbroucke (2019) study the implications of the war on fertility though use a calibration approach based on national-level data.

² In a previous effort, Koenig (2023) approximated the constituency-level presence of veterans by the relative growth of the male population to study WWI veterans' support for right-wing parties.

³ Spanning three decades of data collection effort, the Union Army Data (Fogel et al., 2000) consists in a sample of 40 thousand soldiers – a 1.6 percent random sample of whites mustered into the Union Army – together with additional samples of six thousand Black soldiers, White recruits in six major cities, and prisoners of war who survived until 1900. This database has recently been expanded to the 2.2 million volunteers of the Union Army by Dippel and Heblich (2021) and Weaver (2022). Dippel and Heblich (2021) rely on the *Union Army Registers* reports issued by each US State's Adjutant General's Reports. These reports include soldiers' full name, enlistment and discharge date, age, military unit, rank, promotion, whether they volunteered, were drafted or commissioned, their State of enlistment, and their status at the end of the war. Weaver (2022) uses an alternative dataset, the American Civil War Research Database.

⁴ Appendix Table B.1 provides the list of abbreviations used in this section. The list of laws discussed therein and links to their original texts are provided in Appendix Table B.2. In particular, to support our assertions, we refer to the relevant legislation, which is generally available in the *Journal Officiel de la République, Lois et Décrets*, abbreviated JOLD.

2.1. The military in World War I France

The territorial organization of the military at the eve of the war ensured consistency between military command and recruitment.⁵ Recruitment was structured into 21 military regions, each further subdivided into eight subdivisions (Figure A.1). In each of these subdivisions of military region, a unique recruitment bureau managed both recruitment and mobilization operations.⁶ Regarding military command, each military region was equipped with a unique army corps, which was composed of two infantry battalions, each in turn composed of one artillery and two infantry regiments—one per subdivision of military region. The army further included several cavalry divisions and additional artillery regiments, among which nine regiments of mobile artillery, two regiments of mountain artillery, and some mobile artillery groups.⁷

French men were subject to universal conscription since the Bertreux Law of March 1905 (Crépin, 2009, 264–329; JOLD, 37(81), 1869–90). Moreover, by August 1913 and the *Law of Three Years*, military obligations had increased to 28 years (Boulanger, 2001, 41–56; JOLD, 46(2), 7138–42): starting at age 20, conscripts were to serve successively for three years in the active army, for eleven years in the reserve of the active army, for seven years in the territorial army, and for seven years in the reserve of the territorial army. As a result, the active army comprised 817 thousand men of the cohorts 1891–3 upon mobilization on August 2, 1914. Within a fortnight, these troops were supplemented with the 3 million men of the reserve of the active army. An additional 3 million men from the territorial army and its reserve were then gradually incorporated through December 1915, and another 1.2 million in 1916, 1917, and 1918 (Huber, 1931, 88–111).⁸ By the end of the war, close to 8 million French men and an additional 500 thousand colonial troops had participated to the conflict.

There were initially relatively few exemptions to universal conscription, as only a few conscripts employed in the military administration were exempted from armed services upon mobilization. However, as the war lingered, the military command soon realized that its plan for supplying troops with weapons and machinery was dramatically insufficient (Porte, 2005; Bostrom, 2016). Indeed, consistent with the cult of the all-out offensive that was then prevalent among belligerent nations (Marill, 2014), the initial plan of military mobilization did not specify an industrial organization that would support a potentially long war.⁹ To cope with industrial warfare together with the shortage of civilian labor and the German occupation of the industrial Northeast, the military command soon started to withdraw soldiers with manufacturing skills from the battlefield and allocate them into war factories and mines.¹⁰ As a result, up to 559 thousand soldiers were withdrawn from the front lines over the course of the war, representing close to 12 percent of army personnel. About 300 thousand soldiers were also mobilized in harvesting from 1917 onward.¹¹

2.2. The *Mort pour la France* records

Mort pour la France individual records constitute the underlying source of the MPF database. This archival material was the outcome of a complex process that involved the *Mort pour la France* recognition and the preparation of the *Livre d'Or* memorial through most of the 1920s.

The Mort pour la France recognition. As early as Winter 1914, the army was facing unprecedented mass mortality (Prost, 2008; Bach, 2010). To foster patriotism and mitigate the risk of demoralization, the government implemented a set of measures to pay tribute to those who had died in its name (Petit, 2016). The idea of a honorary title attributed to fallen soldiers first materialized in December 1914, eventually passing into law in July 1915 (JOLD, 47(184), 4653–4).¹²

The *Mort pour la France* honorary title represented an official recognition that was awarded to all military personnel of the French Army as well as to the French civilians who had died as a result of military operations between August 2, 1914, and October 24, 1919.¹³ While its attribution was systematic for those who had died in the line of duty – provided approval of military authorities – it

⁵ The foundations of this organization had been in place since a series of reforms passed in 1873–4 in the aftermath of the Franco-Prussian war (Boulanger 2001, 15–37; Chanet 2006; Gay 2021, 193–5) and in particular the Law relative to the general organization of the army of July 24, 1873 (JOLD, 5(215), 5281–3). The specific territorial structure of the military prevalent in August 1914 had been further adjusted by the military laws of September 22, 1913 (JOLD, 45(261), 8546–7), December 22, 1913 (JOLD, 45(349), 11090–1), and December 31, 1913 (JOLD, 46(2), 108–9).

⁶ Two exceptions to this territorial organization were the cities of Paris and Lyon, which had their own military government and recruitment bureaus.

⁷ The army further included 55 regiments of colonial infantry, seven regiments of colonial artillery, and several regiments of colonial troops. The colonial infantry and artillery had initially been created to defend France's colonial empire and were composed of French men born in mainland France. In contrast, colonial troops were mostly composed of natives of France's colonies (Antier, 2008).

⁸ Figure A.2 displays the war exposure of each cohort from the mobilization on August 2, 1914 to the official cessation of hostilities on October 24, 1919.

⁹ For instance, the plan of mobilization designed in 1913 only provisioned 50 thousand workers allocated across 30 State factories (Porte, 2005, 26). Likewise, it procured for 14 thousand 75 mm shells per day, although nearly 40 thousand of them were used daily during the Race to the Sea in October 1914. By then, half of the stocks had been depleted (Bostrom, 2016, 264).

¹⁰ This allocation policy was enacted by the Dalbiez Law of August 1915 (JOLD, 47(223), 5795–7), which stipulated that “[t]he Ministry of War is authorized to allocate to corporations, factories, and mines working for the national defense men belonging to a mobilized or mobilizable age class, industrial managers, engineers, production managers, foremen, workers, and who will justify to have held their job for at least a year in those corporations, firms and mines, or in comparable corporations, firms, and mines”.

¹¹ Appendix Table B.3 provides a detailed account of the number of mobilized soldiers outside of armed services throughout the war.

¹² The legislative process was initiated by Lower House member Joseph Thierry in December 1914, who proposed to amend the Civil Code regarding death certificates of military personnel (JOCOD doc. parl., 10, 2185). The bill passed promptly in the Lower House in February 1915 (JOCOD, 11, 163) then in the Upper House in April 1915 (JOS, 11, 159–60). It was adopted in June 1915 (JOCOD, 11, 829). For more details, see BALD (68, 226, footnote 1).

¹³ The circumstances under which a death qualified for the recognition were later extended by the Law of February 28, 1922 (JOLD, 54(59), 2450). The date of October 24, 1919 corresponds to the official date of cessation of hostilities per the Law of October 23, 1919 (JOLD, 51(289), 11790).

had to be explicitly requested by the families of civilians or military personnel whose death was not directly attributable to military operations, such as deaths by suicide, disease, or accident. Soldiers executed on grounds of mutiny or treason were denied the recognition. Once approved, this honorary title was transcribed onto the death certificate of the deceased.¹⁴ This inscription on the death certificate had the force of law and, shortly after the war, opened compensation rights to the spouse, children, and ascendants of the deceased.¹⁵ The time interval during which a death opened the right to be awarded the recognition was indefinitely extended by November 1945 beyond the date of cessation of hostilities of October 24, 1919, so that families of those deceased even decades after the war as a result of injuries contracted during the conflict can claim this recognition ex-post, even today.

The Livre d'Or Memorial. The nation's homage to war victims was structured by the Law of October 25, 1919 (*JOLD*, 51(291), 11910), which stipulated the writing of a *Livre d'Or* (Golden Book) containing the names of all those who had received the *Mort pour la France* recognition (Chabord, 1973; Richard, 2013).¹⁶ The 120 volumes of the *Livre d'Or* were to be deposited at the Panthéon in Paris and excerpts sent to all of France's municipalities with the names of the *Morts pour la France* born or residing therein at mobilization.¹⁷

Several ministries contributed to the preparation of the *Livre d'Or*, each enumerating their own *Morts pour la France*: the ministry of the Navy for maritime armed forces, the ministry of the merchant Navy for merchant navigators, the ministry of the Interior for civilians, and the ministry of Pensions for the Land Army, whose constituents represented the overwhelming majority of *Morts pour la France* (Chabord, 1973, 80).¹⁸ The office of the civil registry of the ministry of Pensions began its operations in July 1922, once the situation of most missing soldiers had been settled.¹⁹ It sent to all regiment operating bases 12-by-17 centimeter cards of three types: pale-havana cards for those who had been granted the *Mort pour la France* recognition, green cards for those who were denied the recognition, and white cards for unsettled cases. Regiments then filled these cards using information contained in their own records of deceased soldiers, their *fichier de position* (position file) (Prost, 2008, 49–50). These files were based on soldiers' military service records held by their recruitment bureaus of origin (Keszenbaum, 2013, 10–7). This process lasted until July 1929. By then, the ministry of Pensions had received 1.5 million cards from regiment operating bases. The ministry of Pensions then sent preliminary lists of *Morts pour la France* to all municipalities for verification. Four-fifth of them responded with various comments in 1929–30, which enabled the Ministry of Pensions to rectify the content of individual cards (Chabord, 1973, 80).²⁰

Despite a colossal enterprise that lasted for nearly a decade and mobilized up to 60 employees of the ministry of Pensions, the multiple requests from the Lower House in the 1930s to print the 120 volumes of the *Livre d'Or* were denied due to the budget crisis faced by the government (Chabord, 1973, 82–9). The *Livre d'Or* never saw the light of day.

The Morts pour la France records. While the *Livre d'Or* never materialized, its source material – 1.5 million individual cards – was deposited at the Bureau of Archives of Victims of Contemporaneous Conflicts (BAVCC) in Caen.²¹ Starting in 1999, these cards were scanned by the BAVCC's Direction of Heritage, Memory, and Archives (DPMA) and made available on the *Mémoire des Hommes* (MdH) web platform on November 11, 2003 (Lemarchand, 2008; Aufray, 2014).²²

Morts pour la France records contain the following entries: last name, first name, military rank, army corps, recruitment and regiment identification numbers, class (recruitment cohort), recruitment bureau, date, location and type of death, date and location of birth, and date and location of transcription of death—we display the MPF record of one of the authors' ascendants in Fig. 1. Scanned records made available through the MdH web platform were initially indexed by the BAVCC through five fields: first name, last name, date of birth, département of birth, and country of birth. To gather information on the remaining fields – representing close to 20 million cells – starting in November 2013, the MdH web platform made it possible to annotate *Morts pour la France* records online (Aufray, 2017). Ten days later, the initiative “1 Jour–1 Poilu” was launched, mobilizing “citizen historians” through social media in order to fulfill this immense task (Gilot et al., 2018).²³ Thanks to the participation of 25 thousand contributors, the indexation of all the fields of the *Morts pour la France* database was achieved on April 30, 2018 (Gilot, 2020). On January 22, 2019, the resulting annotated database was first made downloadable in csv format on the MdH web platform. Its content is periodically updated with corrections made by online contributors.²⁴

¹⁴ The practicalities of the applications of the law are detailed in the circular of the Ministry of Justice of July 8, 1915 (*GGDO*, 5, 152–8).

¹⁵ Conditions of attribution of these compensation rights are defined by the Law of March 31, 1919, art. 15–6 (*JOLD*, 51(90), 3342–3).

¹⁶ The idea of this memorial was initiated as early as June 1916 by Upper House member Louis Martin. However, the bill was only discussed in May 1918, eventually passing into law in October 1919 after several modifications (*BALD*, 72, 495, footnote 1). For a comparable homage through a Golden Book in Italy, see Fornasin (2017).

¹⁷ Homages to fallen soldiers were pervasive throughout the country, generally emanating from decentralized initiatives. The most iconic initiative was the construction of *Monuments aux Morts* (War Memorials) in almost all of France's municipalities (Prost, 1992; Aubry and de Oliveira, 2014).

¹⁸ See Chabord (1973, 80–9) for details on how ministries other than the ministry of Pensions managed this process.

¹⁹ The precise guidelines of these operations are detailed in the Instruction of January 11, 1922 of the ministry of Pensions.

²⁰ Annotated lists sent by municipalities to the ministry of Pensions were deposited at the National Archives in 1963. These were digitized in 2008–9 (Richard, 2013, 151–2) and are now available on the website of the National Archives (Richard, 2023 [2010]).

²¹ Three types of individual cards are deposited at the BAVCC: the 1.4 million pale-havana cards of the *Morts pour la France*, among which 2 thousand civilians, the 60 thousand green cards of those who were denied the *Mort pour la France* recognition, and the 30 thousand white cards of initially unsettled cases (Rolland, 2005, 270).

²² The original MPF records are available at <https://www.memoiredeshommes.sga.defense.gouv.fr/fr/article.php?larub=24&titre=morts-pour-la-france-de-la-premiere-guerre-mondiale> (accessed on September 25, 2023).

²³ Primary tools for this mobilization campaign were the Twitter hashtag #1J1P together with its associated account @1J1Poilu and web page <http://www.1jour1poilu.com> (accessed on September 25, 2023).

²⁴ These raw data files are available at <https://www.memoiredeshommes.sga.defense.gouv.fr/fr/article.php?laref=1380&titre=telechargement> (accessed on September 25, 2023).

PARTIE À REMPLIR PAR LE CORPS.

Nom GALINIER

Prénoms Joseph Henri

Grade Soldat

Corps 22^e Régiment d'Infanterie Coloniale

N^o 018686 au Corps. — Cl. 1901

Matricule. 349 au Recrutement Carbonna

Mort pour la France le 23 Septembre 1914

H^{at} au site de Brienne le château

Genre de mort Suite Blessure de guerre

Né le 6 novembre 1881

St-Laurent de la Cabrerisse département Aude

Arr^s municipal (p^r Paris et Lyon), }
a défaut rue et N^o.

Jugement rendu le 20 Janvier 1915

par le Tribunal de St-Victor

acte ou jugement transcrit le Aude

N^o du registre d'état civil

101-708-1022. [26434]

Fig. 1. Example of a Mort pour la France record.

Notes: Example of the MPF card of Joseph Henri Galinier, 22nd colonial infantry regiment, Mort pour la France on September 23, 1914, in Brienne-le-Château due to war injuries. He was born on November 6, 1881, in Saint-Laurent-de-la-Cabrerisse, in the département of Aude.

Limits of the Mort pour la France records. As rich as the content of the Mort pour la France records may be, the process through which this source was constructed conditions the reality of military fatalities the economic historian can capture.

First, these records remain a byproduct of a broader enterprise devised by the State with two primary objectives: provide a national homage to war victims and budget compensations to their families to be distributed by the ministry of Pensions, as the Mort pour la France recognition opened rights to compensations. An accurate census of deceased soldiers was never an explicit objective and, in fact, has always remained out of reach (Loez, 2008; Prost, 2008). Indeed, these records do not account for soldiers who died outside of the line of duty and whose families did not claim the Mort pour la France recognition. Conversely, they account for some civilian victims, among which some employees of the merchant Navy. This source nevertheless remains the closest to a census of deceased soldiers ever attempted.

Second, these records are subject to potential inaccuracies in their information content. They remain a third-level source: their content is based on soldiers' military service records held by recruitment bureaus, which were initially transcribed onto regiments' position files during the war, then onto Morts pour la France records between 1922 and 1929. Regiments' personnel, generally composed of fresh conscripts, were not necessarily trained for these tasks. Moreover, annotations made by the BAVCC between 1999 and 2003 and by citizen historians between 2013 and 2018 might contain inaccuracies not only due to user error, but also because the Mort pour la France records may sometimes be unreadable, exhibit crossed-off content, or contain multiple information for the same field. And contrary to other comparable crowdsourcing projects, no consensus method was used during the annotation process (see, e.g., Van Galen, 2019). Finally, annotators transcribed the information content of the Mort pour la France records as they appeared. As a result, except for the five fields annotated by the BAVCC, annotations remain unstandardized, heterogeneous, and oftentimes replete with redundant information, making them very challenging to use directly for empirical analysis.

One of our contributions is to process the raw data files available on the MdH web platform and create a curated database with standardized fields and interoperable nomenclatures compatible with, e.g., both historic and contemporaneous geographic information systems (Gay, 2021; IGN, 2015), so that the MPF database we provide is directly amenable to empirical analysis by quantitative historians.

Table 1
Variables in the MPF database (1/2).

Category	Variable	Description	Values
Identifiers	ark	Mémoire des Hommes url	String
	id	Record unique identifier	Numeric
General	mpf	MPF status	See table notes
	gender	Gender	Male (1), Female (2)
	civilian	Civilian status	No (0), Yes (1)
Names	lname	Last name	String
	fname	First name	String
	oname_type	Other name type	See table notes
	oname	Other name	String
Birth date	birth_date_flag	Birth date flag	See table notes
	birth_date	Birth date	Date
	birth_year	Birth year	Numeric
	birth_month	Birth month	Numeric
	birth_day	Birth day	Numeric
Death date	death_date_flag	Death date flag	See table notes
	death_date	Death date	Date
	death_year	Death year	Numeric
	death_month	Death month	Numeric
	death_day	Death day	Numeric
Death type	death_type_flag	Death type flag	See table notes
	death_type	Death type	12 categories
Recruitment class	class_flag	Recruitment class flag	See table notes
	class	Recruitment class	Numeric
	engage	Volunteer status	No (0), Yes (1)
Military rank	rank_flag	Military rank flag	See table notes
	rank_gen	Military rank [general]	See table notes
	rank_det	Military rank [detailed]	52 categories
Military unit	unit_flag	Military unit flag	See table notes
	unit_type_gen	Military unit type [general]	See table notes
	unit_type_det	Military unit type [detailed]	16 categories
	unit_id	Military unit id	Numeric
	unit	Military unit acronym	String
	unit_name	Military unit name	String
	unit_colony	Colonial military unit	Mainland (0), Colonial (1)

Notes: Values for mpf: Non MPF (0), MPF (1), Unsettled (2). Values for oname_type: No other name (1), Ou (2), Alias (3), Dit (4), Née (5), Épouse (6). Values for birth_date_flag: Birth date available (1), Birth day missing (4), Birth month and day missing (7), Birth date entirely missing (8). Values for death_date_flag: Death date available (1), Death day missing (4), Death month and day missing (7), Death date entirely missing (8). Values for death_type_flag: Values for class_flag: Missing (1), Location of death (2), Type of death (3). Class available and in bounds pm 4 years (1), Class missing and inputted from birth year (2), Class out of bound (pm 3 years), inputted from birth year (4), Class missing (5). Values for rank_flag: Rank available (1), Rank missing (2), Civilian (3). Values for rank_gen: Enlisted (2), NCO (3), CO (4), Rank missing (.a), Civilian (.b). Values for unit_flag: Unit available (1), Unit missing (2), Civilian (3). Values for unit_type_gen: Infantry (1), Artillery (2), Engineering (3), Cavalry (4), Other (5), Unit missing (.a), Civilian (.b). The MPF database is available in the Harvard Dataverse at <https://doi.org/10.7910/DVN/2APQZ6> (Gay and Grosjean, 2023).

3. A database of French fatalities of the Great War

In this section, we provide an overview of the variables in the MPF database: soldier's first and last name, dates of birth and death, circumstances of death, recruitment class, military rank, military unit, location of birth, and recruitment bureau—see Tables 1 and 2 for a summary of the content of the MPF data files.

Raw data files available on the MdH web platform as of September 2023 contain 1,825,593 observations corresponding to 1,422,003 unique MPF records. Starting from these files, we drop duplicate records and soldiers born outside of mainland France.²⁵

²⁵ There are close to 400 thousand duplicate records in the raw data files because one additional observation is created whenever a given field on an MPF record contains multiple annotations.

Table 2

Variables in the MPF Database (2/2).

Category	Variable	Description	Values
Birth place	birth_place_flag	Birth place flag	See table notes
	birth_cassini	Birth municipality identifier [Cassini]	Numeric
	birth_insee_2015	Birth municipality identifier [INSEE]	String
	birth_geofla_2015	Birth municipality identifier [GEOFLA]	String
	birth_x_geofla	Birth municipality latitude [GEOFLA]	Numeric
	birth_y_geofla	Birth municipality longitude [GEOFLA]	Numeric
	birth_dep_YYYY	Birth département identifier [INSEE, TRF-GIS]	String
	birth_dep_name_YYYY	Birth département name [INSEE, TRF-GIS]	String
Recruitment bureau	birth_com_name_YYYY	Birth municipality name [INSEE, TRF-GIS]	String
	bureau_flag	Recruitment bureau flag	See table notes
	bureau_id	Recruitment bureau identifier	Numeric
	bureau_name	Recruitment bureau name [TRF-GIS]	String
	bureau_mreg	Recruitment bureau military region identifier [TRF-GIS]	String
	bureau_mreg_name	Recruitment bureau military region name [TRF-GIS]	String
	bureau_mregsub	Recruitment bureau military region-subdivision identifier [TRF-GIS]	String
	bureau_msub	Recruitment bureau military subdivision identifier [TRF-GIS]	String
	bureau_msub_name	Recruitment bureau military subdivision name [TRF-GIS]	String
	bureau_cassini	Recruitment bureau municipality identifier [Cassini]	Numeric
	bureau_insee_2015	Recruitment bureau municipality identifier [INSEE]	String
	bureau_geofla_2015	Recruitment bureau municipality identifier [GEOFLA]	String
	bureau_x_geofla	Recruitment bureau municipality latitude [GEOFLA]	Numeric
	bureau_y_geofla	Recruitment bureau municipality longitude [GEOFLA]	Numeric
	bureau_dep_YYYY	Recruitment bureau département identifier [INSEE, TRF-GIS]	String
	bureau_dep_name_YYYY	Recruitment bureau département name [INSEE, TRF-GIS]	String
	bureau_com_name_YYYY	Recruitment bureau municipality name [INSEE, TRF-GIS]	String

Notes: Values for birth_place_flag: Commune available (1), Commune available, Alsace-Lorraine (2), Département available only (3), Birth place missing (4). Values for bureau_flag: Current recruitment bureau (1), Former recruitment bureau (2), Split recruitment bureau (3), Recruitment bureau département only (4), Recruitment bureau city only (5), Former recruitment bureau, Alsace-Lorraine (6), Recruitment bureau out of France (7). YYYY: information relative to the geographic information systems of 2015 (INSEE, 2015; IGN, 2015), 1911, and 1921 (Gay, 2020a,d, 2021). The MPF database is available in the Harvard Dataverse at <https://doi.org/10.7910/DVN/2APQZ6> (Gay and Grosjean, 2023).

The final MPF database contains 1,270,943 records, among which 198 women (mostly nurses and civilians), 1,450 civilian victims, and 53,735 soldiers who were denied the MPF recognition (158 cases are still currently unsettled).²⁶

First and last name. The MPF database contains 182,434 unique last names. Their distribution is highly skewed: the most frequent last name, MARTIN, counts 5,281 records (0.42 percent of MPF records), while 83,180 last names (45.59 percent of all unique last names) count only one entry.²⁷ Comparing the distribution of last names in the MPF database to that of all contemporaneous last names reveals that the last names composition of MPFs is representative of the general population (Appendix Table B.6).²⁸ Likewise, the distribution of first names is highly skewed. Over 227,701 unique first names, the most frequent, JEAN, counts 31,574 records (2.48 percent of MPF records), while 170,535 first names (74.89 percent of all unique first names) count only one entry.²⁹ Both the last name and first name variables have nearly full coverage as all MPF records have a last name and only 176 have no first name mentioned.³⁰

Dates of birth and death. Birth dates have nearly full coverage: 99.70 percent of MPF records have a complete birth date, while 3,679 records (0.29 percent of MPF records) only contain a birth year with no month or day information, and 75 records (0.01 percent of MPF records) have no birth date at all. While the conscription directly concerned the cohorts 1866–1900, birth dates range from 1824 to 1917, though only 3,240 MPFs were born between 1824 and 1865 and 253 between 1901 and 1917 (0.27 percent of MPF records). We display the distribution of birth years among MPFs born between 1866 and 1900 in Fig. 2. The most affected cohorts

²⁶ See Guillot and Parent (2021, 2022) for a quantitative study of the profile of soldiers who were denied the Mort pour la France recognition based on MPF records.

²⁷ Appendix Table B.4 provides the list of the 50 most frequent last names along with their frequencies and Appendix Table B.5 provides the distribution of unique last names.

²⁸ To construct this comparison, we use a subset of INSEE's (2012) file of last names, which provides the total number of last names among births registered in France over the cohorts 1891–1900. More precisely, in Appendix Table B.6, we compare the 10 most frequent last names in the MPF database among all MPF records and among the MPFs born between 1891 and 1900 to the 10 most frequent last names in INSEE's (2012) database. We find that the ordering of these 10 most frequent last names is almost identical in all three samples. For more details on the representativeness of last names in the MPF database, see Darlu and Chareille (2020, 99–102).

²⁹ Appendix Table B.7 provides the list of the 50 most frequent first names along with their frequencies and Appendix Table B.8 provides the distribution of unique first names. Because the INSEE's (2012) first names file only contains first names given after 1900, we cannot make a comparison to gauge the representativeness of first names in the MPF database.

³⁰ MPF records further contain aliases as well as maiden names for women in 5,405 cases (0.43 percent of MPF records). We display their distribution in Appendix Table B.9.

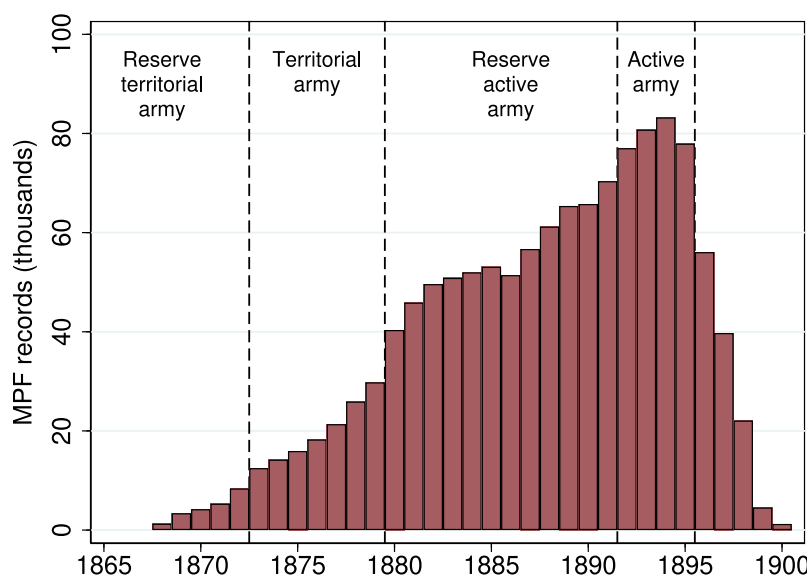


Fig. 2. Distribution of years of birth among MPF records.

Notes: Each bar represents the number of MPF records who were born in a given year. The figure is restricted to the cohorts 1865–1900, which represent 1,267,852 MPF records. It further indicates the armies to which MPF cohorts belonged in 1914 (post-1895 cohorts were drafted after 1914).

were those in the active army (1891–5), who were sent *en masse* to the front lines. In particular, the cohorts 1893–4 each count more than 80 thousand MPFs.

We display the distributions of months and days of birth among MPF records in Appendix Figures A.3 and A.4, respectively. The median MPF was born on February 22, 1889—they would have been 25.5 years old in August 1914. Dates of birth exhibit no cyclical pattern nor any spike in January or on the first day of the month. This contrasts with age heaping often found in censuses and other data from similar time periods—see, e.g., Abramitzky et al. (2021) for age heaping in US census data. A likely explanation is that this information was not provided by soldiers themselves but was based on the information available on soldiers' military service records held by recruitment bureaus, which were filled at the time of recruitment by military officials based on information from civil birth records.

The coverage of death dates is also comprehensive since it is complete for 98.79 percent of records.³¹ 12,615 records (0.99 percent of MPF records) only contain the death year with no month or day information, and 2,693 records (0.21 percent of MPF records) have no death date at all. While the official date of cessation of hostilities was October 24, 1919, 3,530 records indicate a death date after that time (0.28 percent of MPF records), as the MPF recognition could be claimed past the end of the war. Moreover, 37,812 records (2.98 percent of MPF records) indicate a death date between the Armistice of November 11, 1918, and the cessation of hostilities of October 24, 1919. The median MPF died on January 20, 1916. We display the distribution of monthly death dates in Appendix Figure A.5.

Circumstances of death. While a field exists to record the circumstances of death in the original MPF records and is filled in most cases with the mention “killed in action” (*tué à l'ennemi*), this field was (the only one) not indexed by the BAVCC on the MdH web platform. However, the field that indicates the location of death sometimes includes information that we can use to infer soldiers' circumstances of death. This field contains heterogeneous information, from the actual location of the death, such as a given battlefield or hospital, to the circumstances of the death, such as a disease or an accident. In nearly half of cases, however, this field is empty in the MdH data files. In these cases, we assume that the soldier was killed in action. From there, we create two variables: one that provides the type of death across 12 categories and one that qualifies this typology, indicating when the circumstance of death was missing in the MdH data file so as to warn users to exert caution when using this variable. Overall, nearly all MPF records indicate a death in action (972,019 records or 76.48 percent of MPF records—and up to 81.27 percent for infantrymen), while the rest generally died in a hospital (14.42 percent) or an ambulance (6.66 percent) (see Appendix Table B.10).

Recruitment class. Soldiers' recruitment class (*classe*) refers to the year in which they were first drafted, i.e., when they first served their compulsory military service. Ever since the Bertreaux Law of 1905, a soldiers' class normally corresponded to the year when he turned 20. In fact, the distribution of recruitment classes mirrors that of birth years plus 20 years (Appendix Figure A.6).

There were three types of exceptions to this enlistment rule (Boulanger, 2002). First, soldiers could voluntarily enlist by anticipation under the status of war volunteer (*engagé volontaire*) as early as 17 years old. Beyond sheer patriotism, early enrollment

³¹ The recorded date of death on MPF records refers to the actual date when the soldier died, not the date when the death was transcribed onto a death certificate. This implies that for soldiers who died of war injuries, the date of death corresponds to the day of death, not the day when injuries occurred.

enabled volunteers to choose their weapon so as to avoid the infantry, which had higher fatality rates than the artillery or the cavalry (Boulanger, 2001, 128–40). MPF records contain information on the volunteer status of deceased soldiers. They reveal that while only 0.2 percent of the 1870–93 cohorts enrolled as volunteers, cohorts that were to be recruited during the war enrolled as volunteers at much higher rates (Appendix Figure A.7): rates of voluntary enrollment increased from 0.4 percent for the 1894 cohort to 2.6 percent for the 1899 cohort. The second exception contended that soldiers could be forcibly enlisted before age 20 through a procedure of recruitment by anticipation. This procedure was pervasive during the war in order to face unprecedented losses. For instance, men of the 1916–7 classes were drafted as early as 1915 (Boulanger, 2001, 57–67). Comparing years of birth to recruitment classes among non-volunteer soldiers reveals that while only 2.3 percent of the 1870–94 cohorts were recruited by anticipation, this proportion was much higher among younger cohorts, reaching 14.5 percent for the 1899 cohort (Appendix Figure A.8). A third and final exception was deferments. Deferments were possible under various conditions: to complete one's studies (e.g., medical training), for fathers of large families, or on grounds of poor health. Nevertheless, the military command granted much fewer exemptions during the war (Boulanger, 2001, 67–75). MPF data reveal that while 1.4 percent of the 1870–94 cohorts were granted deferments, this was only the case for 0.9 percent for the 1895 cohort and 0.03 percent for the 1899 cohort (Appendix Figure A.9).

Information for recruitment class is available for 98.53 percent of MPF records.³² For 14,093 records, the original recruitment class is more than three years apart from the soldier's birth year plus 20, which is likely erroneous, and for 18,208 records, the recruitment class is missing (1.43 percent of MPF records). In these cases, we impute the recruitment class with the theoretical value of birth year plus 20. Appendix Table B.11 provides a breakdown of these cases, which are flagged in the MPF database.

Military rank (at time of death). We normalize MPFs' military ranks at time of death into 51 categories so that each category contains at least 10 observations—we classify residual categories into an OTHER category, which contains 53 records. Rank information is missing for 4,506 records (0.35 percent of MPF records), not counting civilians. As shown in Appendix Table B.12, the most frequent rank is SOLDIER with 870,382 records (68.81 percent of MPF records). The second most frequent is CORPORAL with 104,956 records (8.30 percent). 79.32 percent of MPFs are low-ranked enlisted soldiers, 18.91 percent are non-commissioned officers, and 1.76 percent are commissioned officers, including 90 generals.

Military unit (at time of death). We normalize MPFs' military units at time of death into 250 entities further divided into 16 categories (Appendix Table B.13). The overwhelming majority of MPFs were part of the infantry, with close to 1.1 million records (86.08 percent of MPF records), and generally part of an infantry regiment (64.90 percent).³³ Other units include artillery (6.59 percent), engineering (2.23 percent), cavalry (1.62 percent), and miscellaneous categories (3.48 percent).³⁴ Despite being born in metropolitan France, nearly 10 percent of MPFs belonged to a colonial military unit (see Footnote 2.1 and the example in Fig. 1). Again, the data coverage for this variable is substantial as only 7,223 records (0.57 percent of MPF records) do not contain information on military unit, not counting civilians.

Location of birth. Information on MPFs' locations of birth is available in nearly all cases at the level of the municipality (99.33 percent of MPF records), the most granular level of administration—France counted close to 35 thousand municipalities before the war. In 8,400 cases (0.66 percent of MPF records), only the département of birth is provided. The location of birth is missing entirely for only 61 records.

To facilitate usage of this variable for quantitative analysis, we match birth location strings provided in original MPF records to their corresponding Cassini identifiers, which uniquely identify each of the 41,410 municipalities that ever existed between 1801 and 2005 (Motte et al., 2003; Motte and Vouloir, 2007; Gay, 2021).³⁵ The availability of this unique identifier makes it possible to match MPFs' birth locations to any posterior geographic information system. In particular, we match these locations to the nomenclatures of municipalities as of 1911 and 1921 (Gay, 2020a) as well as to the nomenclature of municipalities as of 2015 (INSEE, 2015), which spatial counterpart corresponds to IGN's (2015) Geofla communes shapefile.

The most frequent municipality of birth is PARIS with 64,988 records (5.15 percent of MPF records), before LYON with 8,851 records (0.70 percent), and MARSEILLE with 8,188 records (0.65 percent). At the other end, 666 municipalities count only one entry (0.05 percent of MPF records).³⁶ In Appendix Figure A.11, we display the distribution of MPF records across municipalities of birth in 2015 geography.³⁷

³² By construction, this information is missing for civilians and women.

³³ We also provide the military unit number, which is available for 99.01 percent of MPF records for which the military unit information is available.

³⁴ Reflecting the evolution of the strategic importance of the infantry relative to the artillery over the course of the war, the share of infantry among MPFs declined from 94 percent in 1914 to 74 percent in 1918 and 54 percent in 1919 (Appendix Figure A.10).

³⁵ Cassini identifiers are from the *Histoire Administrative des Communes* (HAC) database, initially distributed through <http://cassini.ehess.fr>, now available as a csv file on EHESS' Didoména repository (Cristofoli et al., 2021).

³⁶ Appendix Table B.14 provides the list of the 40 most frequent municipalities of birth along with their frequencies and Appendix Table B.15 provides the distribution of unique municipalities. While Paris appears as a single municipality in these tables, MPF information is broken down across Paris' 20 municipal arrondissements in the database.

³⁷ Close to 4 thousand MPFs were born in Alsace-Lorraine. These numbers are relatively low compared to other municipalities, since those regions were part of German territory in 1914. Still, it is estimated that 18 thousand men from Alsace-Lorraine volunteered for the French Army, including 12 thousand who were already in France at the start of the war and 3 thousand who would have escaped German mobilization (Georges, 2015).

Recruitment bureau. MPF records indicate the name of the recruitment bureau where each soldier was first drafted at age 20. As explained in Section 2.1, each subdivision of military region had one recruitment bureau, i.e., 154 bureaus in total in 1914 (Appendix Figure A.1).³⁸ However, while 97.23 percent of MPFs were recruited by a bureau that existed in 1914, a few MPFs (7,276 cases or 0.57 percent of MPF records) were recruited by former bureaus that did not exist anymore by 1914. Additionally, while born in France, some MPFs were recruited outside of France (2,240 cases or 0.18 of MPF records). Nevertheless, the availability of this variable is nearly complete as only 1.60 percent of MPF records do not have a recruitment bureau information, not counting civilians.

To facilitate usage of this variable for quantitative analysis, we match MPFs' recruitment bureaus to two types of geographic information. First, we match each recruitment bureau to its corresponding military region and subdivision of military region (Gay, 2020c,d)—as explained in Section 2.1, the geography of military command and recruitment overlapped consistently. Second, we match each recruitment bureau to the nomenclatures of municipalities as of 1911, 1921, and 2015, in the same way as for MPFs' locations of birth (IGN, 2015; INSEE, 2015; Gay, 2020a).

4. Applications

This section provides three examples of how the MPF database might be used for empirical analysis: first, by aggregating MPF records over space to create military death rates and approximate the demographic shock of the war at different geographic levels (Section 4.1), second, by aggregating MPF records over time to shed light on the most deadly battles of the war (Section 4.2), and third, to leverage the individual-level data for an analysis of survival duration during the conflict (Section 4.3).

4.1. Aggregate-level military death rates

An important application of the MPF database is the aggregation of MPF records at various geographic levels to measure the demographic shock of the war. Indeed, the sex ratio among adults aged 15–50 dropped from 98 men per 100 women at the onset of the war to 88 by the end of the war, and it was not until the early 1950s that it reverted back to balance (Appendix Figure A.12).³⁹ This type of analysis has been done by previous research both at the département (Gilles, 2010; Gilles et al., 2014; Beau, 2017; Guironnet and Parent, 2019; Boehnke and Gay, 2022; Gay, 2023) and at the municipality levels (Cagé et al., 2023). We discuss these applications below and provide some insight on how to properly use these military death rates variables.

4.1.1. Département-level military death rates

We first create a dataset of département-level military death rates by calculating the ratio of MPFs to the size of a département's drafted population during the war, which we approximate by the number of men aged 15–44 in the census of 1911, the last census before the war.⁴⁰ We display the distribution of military death rates across départements in Fig. 3. They range from 6.7 percent in Alpes-Maritimes to 29.3 percent in Lozère, with an average of 16.1 percent and standard deviation of 3.8 percent (Panel A of Table 3).⁴¹

As highlighted in Section 2.1, the systematic part of the distribution of military fatalities was determined by the territorial organization of military recruitment and the industrial war effort. Soldiers recruited in the same military region were initially grouped into the same units and sent to the same battlefields. However, as the war lingered and fatalities accumulated, soldiers were allocated to depleted units depending on the most urgent needs of the military starting in the Fall of 1914, resulting in variation both within and across military regions. Moreover, to cope with industrial warfare, many soldiers from industrial areas were sent to war factories, effectively protecting them from death in the trenches. These patterns are apparent when analyzing the pre-war determinants of military death rates across départements through linear regressions (Appendix Table B.18), as rural départements experienced greater military death rates—rurality explains 75 percent of the variation in military death rates. Other characteristics capturing the pre-war economic structure of départements are not significantly correlated with military death rates and do not explain much more of their variation, as does further including military region fixed effects.⁴² For these reasons, we advise potential users of the département-level military death rates data to rely on the residual variation in military death rates once rurality is accounted for.⁴³

³⁸ This includes the five recruitment bureaus of the military government of Paris, the three of the military government of Lyon, and the recruitment bureau of Versailles, which were not attached to a specific subdivision of military region.

³⁹ The adult sex ratio was only marginally affected by military fatalities of World War II as 115 thousand French soldiers died over that period, which is an order of magnitude below the fatalities from World War I (Lagrou, 2002).

⁴⁰ More precisely, the numerator consists in the sample of non-civilian male MPFs born in mainland France outside of Alsace-Lorraine and who died between August 2, 1914 and October 24, 1919.

⁴¹ Département-level military death rates were previously made available by Boehnke and Gay (2020) using an early version of the MPF database. In Appendix Tables B.16 and B.17, we compare the raw département-level data used in Boehnke and Gay (2022) to those in the current study. Numbers are fairly close and military death rates are on average larger by 0.51 percentage points when using the current updated database. This difference can be explained by the fact that Boehnke and Gay (2022) did not include soldiers who were not awarded the MPF mention, resulting in a sample smaller by about 40 thousand records.

⁴² We capture rurality through two aspects: the share of the population that resides in municipalities with fewer than 2,000 inhabitants and the share of the population born in the département. Other characteristics we consider include pre-war female labor force participation rates, the share of girls aged 5–19 in school, total fertility rates, personal wealth (in 10,000 francs), the share of active population in the industrial sector, average height of conscripts, total population, the share of literate population, and average direct taxes per inhabitant. See Boehnke and Gay (2022, 1217–9) and Gay (2023, 2308–9) for more details.

⁴³ See also Boehnke and Gay (2022, 1225–6) for an instrumental variables approach that exploits discontinuities in the timing of military service across the cohorts 1891–4. It is also worth mentioning that we assign military fatalities to a département through MPFs' départements of birth. However, these might

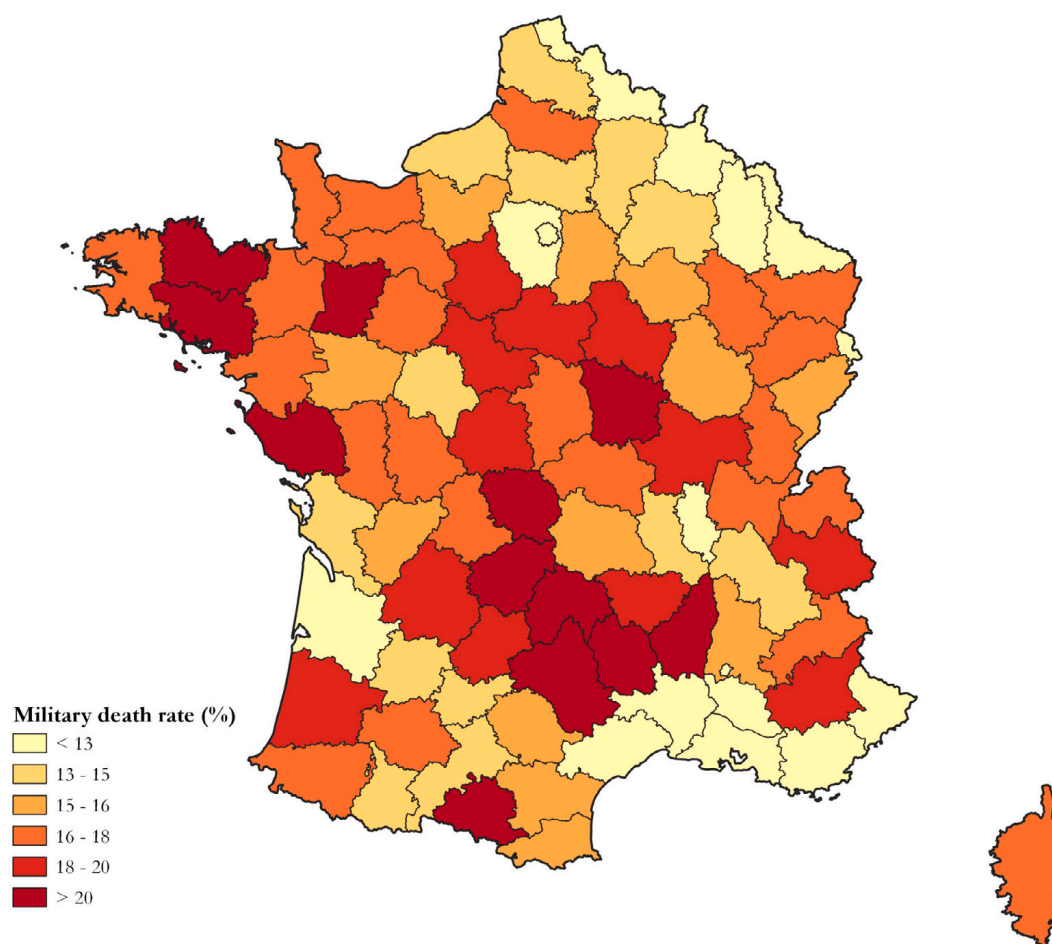


Fig. 3. Military death rates across départements.

Notes: This figure displays the distribution of military death rates across départements. It is the ratio of non-civilian male MPFs who died between August 2, 1914 and October 24, 1919 to the number of men aged 15–44 in the census of 1911. The data are mapped in 1914 geography using the département-level shapefile of [Gay \(2020b\)](#).

4.1.2. Municipality-level military death rates

To analyze the implications of the war at a more disaggregated level, we create a dataset of municipality-level military death rates based on our MPF database. Municipalities are the most disaggregated administrative level in France, as the country counts about 35 thousand municipalities. Unfortunately, the census of 1911 does not provide municipality-level population by age and gender, so we are constrained to rely on total population in the denominator. For consistency, the numerator therefore further includes civilian and female MPFs who were born in mainland France outside of Alsace-Lorraine and died between August 2, 1914, and October 24, 1919. We display the resulting distribution of military death rates in [Fig. 4](#).⁴⁴ Among the 34,893 municipalities of France, only 92 had no MPF—99.7 percent of municipalities had at least one military death from the war. These 92 municipalities were relatively small, as 68 of them had less than 100 inhabitants in 1911, and the largest had 227 (Appendix Table B.19). Other municipalities experienced very high death rates, with 11 municipalities experiencing death rates above 15 percent of their population. But again, these municipalities were relatively small as only one among the 50 most affected by the war had more than a thousand inhabitants in 1911 (Appendix Table B.20). Overall, municipality-level death rates have an average of 3.9 percent and standard deviation of 1.6 percent (Panel B of [Table 3](#)).⁴⁵ Given the modest size of French municipalities, this means that the average municipality lost 36 men during the war.

have differed from their départements of residence at the onset of the war, as 19 percent of men aged 15–44 resided outside their département of birth in 1911—see [Boehnke and Gay \(2022, Appendix E\)](#) on this potential issue.

⁴⁴ Note that we are constrained to map this variable in 2015 geography because municipality-level shapefiles for 1914 do not exist to date. For consistency in the matching process between the pre-war geography and that of 2015, we correct the 1911 population data for municipalities created after the war by attributing the population of their parent municipality in proportion to the population ratio between them and their parent municipality at the time of their creation. We proceed in the same way for the MPF data except that we correct figures for municipalities created after 1891.

⁴⁵ Municipality-level death rates were previously made available by [Cagé et al. \(2023\)](#) using an early version of the MPF database. However, they did not correct for changes in municipalities' geographies, so that we advise users to rely on this updated database of municipality-level death rates.

Table 3
MPFs and military death rates.

	Obs.	Mean	S.d.	Min.	Max.
A. Level of aggregation: Départements					
MPFs	87	14,331	10,253	2,574	81,912
Male population 15–44 in 1911	87	100,855	122,464	21,620	1,090,751
Military death rate	87	16.07	3.84	6.69	29.33
B. Level of aggregation: Municipalities					
MPFs	34,893	36	361	0	64,219
Population in 1911	34,893	1,217	31,465	18	5,776,220
Military death rate	34,893	3.94	1.56	0.00	22.50

Notes: This table provides summary statistics for MPFs and military death rates at two levels of aggregation: départements and municipalities. Information for the male population aged 15–44 in 1911 is from the census of 1911. France's territory excludes Alsace-Lorraine and the communes of La Brigue and Tende. MPFs include non-civilian male MPFs born in France outside of Alsace-Lorraine and who died between August 1, 1914 and October 24, 1919.

Data at the municipality level reveal several new facts compared to previous estimates that were only available at the département level. First, one of the most striking features of Fig. 4 is the extent to which municipalities in the regions of Brittany and Vendée were most consistently and hardly hit. Other regions tend to exhibit much more variation in their municipality-level military death rates, for instance the case in Corsica. Meanwhile, other départements with lower average death rates display pockets of very high death rates, for instance in Doubs and Jura, where respectively 6 and 4 of the 50 municipalities with the highest death rates are located (see Appendix Table B.20). In fact, département fixed effects explain little of the variation in military death rates across municipalities, as the residuals from regressing municipality-level military death rates on département fixed effects have a standard deviation of 1.48, relative to 1.56 for the raw death rates variable.

4.2. The deadliest days of the war

A second potential application of the MPF database is in the study of the dynamics of the conflict. For example, Fig. 5 plots the time series of aggregated daily dates of death in order to shed light on the most deadly battles of the war. It reveals that the deadliest day of the conflict was the first day of the Second Battle of Champagne on September 25, 1915, an ultimately failed offensive for the French Army. On that day, 22,059 soldiers died. Many historians assert instead that the date of August 22, 1914, during the Battle of the Frontiers, was the deadliest day of the conflict (see, e.g., Steg 2013). Our data record this day as the second deadliest day of the conflict with 19,964 fatalities.⁴⁶

Overall, the Battle of the Frontiers of 1914 remains a period of very high mortality. Several days of late August 1914 feature among the days with highest total deaths: 9,671 men died on August 20, 1914, 6,233 on August 25, 1914, and 7,370 on August 28, 1914. A few days later during the First Battle of the Marne on September 8, 1914, 6,624 men would die. Another notable day is the first day of the Chemin des Dames offensive on April 16, 1917, when 9,426 men died during another failed French attempt to break through the German lines. This failed offensive led to the onset of massive mutinies in the French Army (Loez and Mariot, 2008).

4.3. An analysis of survival inequalities on the front

A third potential application of the MPF database is in the study of the determinants of soldiers' survival rates during the war, an analysis previously done by Guillot and Parent (2018) on a non-random sample of 18 thousand deceased soldiers – a 1/52nd sample of the MPF database – born during the first seven days of the month of October, regardless of the year. In this section, we reproduce their analysis based on the full MPF database.

We focus on non-civilian male MPFs of the cohorts 1866–94 born in mainland France outside of Alsace-Lorraine who died between August 2, 1914 and November 11, 1918, for whom the death date is available.⁴⁷ We define survival duration (in days) as the difference between the date of death and 2 August 1914, the first day of the war—conditional on dying during the war, soldiers survived on average for 528 days. We then estimate through linear regressions the relationship between survival duration and the age of soldiers as of August 2, 1914 (in days), their military unit (where the excluded category is Infantry), their military rank

⁴⁶ Note that the MPF database only includes the subset of colonial troops that were born in mainland France and thus does not count colonial troops born in the colonies. The database also counts men who died of diseases on a given day, but these men may have been injured on another day. It therefore remains to be shown whether these differences would make up for the difference in fatalities between what is commonly accepted as the deadliest day of WWI for the French Army and what the MPF database reveals. Another possibility is that some of the previous quantitative studies of the conflict, e.g., Guillot and Parent (2018), relied on the 1866–94 cohorts that were already drafted at the start of the conflict, excluding the new cohorts that were incorporated as the conflict dragged on. We find that August 22, 1914 was indeed the deadliest day for these cohorts. In any case, this point illustrates the potential uses of the MPF database for the historiography of WWI in France.

⁴⁷ This selection is to avoid measurement error due to the fact that we do not know when soldiers first joined the front. From 1,270,943 records, this selection reduces the sample to 976,885 observations.

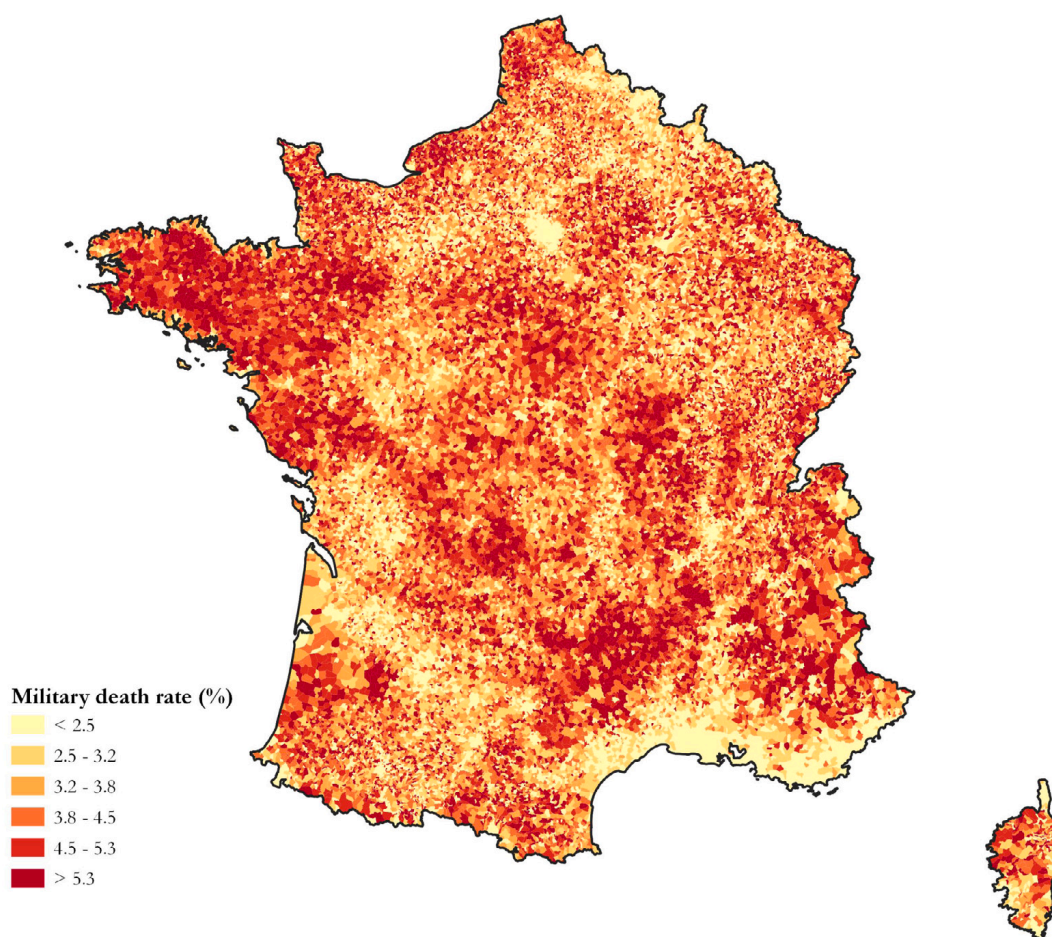


Fig. 4. Military death rates across municipalities.

Notes: This figure displays the distribution of military death rates across municipalities within the boundaries of France as of 1914. It is the ratio of all MPFs who died between August 2, 1914 and October 24, 1919 to the total population in the census of 1911. The data are mapped in 2015 geography using the shapefile of [IGN \(2015\)](#).

(where the excluded category is *Enlisted*), whether they were in a colonial unit, and whether they enrolled voluntarily. We then progressively include a set of fixed effects for soldiers' military region of recruitment, département of birth, and municipality of birth.⁴⁸ We report results in Appendix Table B.21.

Soldiers from the infantry had considerably shorter survival times. On average, soldiers in the artillery survived one more year (350 days) compared to infantrymen, while those in engineering corps survived 231 more days, and those in the cavalry, 314 more days. Soldiers in other units, such as train personnel or marines, survived 434 more days. These findings are qualitatively similar to those in [Guillot and Parent \(2018\)](#), although the magnitudes of differences between the infantry and other units are larger when using the full MPF database. Additionally, officers survived longer than enlisted soldiers: 108 days for non-commissioned officers and 134 days for commissioned officers. Younger soldiers died faster: even conditional on rank and unit, (deceased) soldiers a year older could expect to survive 4 more days. Finally, soldiers from colonial units survived longer, though only by about a month, while the (few) volunteers died faster, by about three months. Including fixed effects for soldiers' spatial origins does not significantly alter estimated coefficients.

5. Data availability

The dissemination of MPF records through the MdH web platform was authorized by the decrees of October 3, 2003 and July 24, 2014 ([Aufray, 2014](#)).⁴⁹ The collection, digitization, indexation, conservation, and online diffusion of the personal data contained

⁴⁸ Selecting on cases for which all these control variables are available further drops close to 23 thousand records, resulting in a sample of 953,382 observations.

⁴⁹ The texts of these two decrees are accessible at <https://www.legifrance.gouv.fr/eli/arrete/2003/10/3/DEFS0302096A/jo/texte> and <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000029361926>.

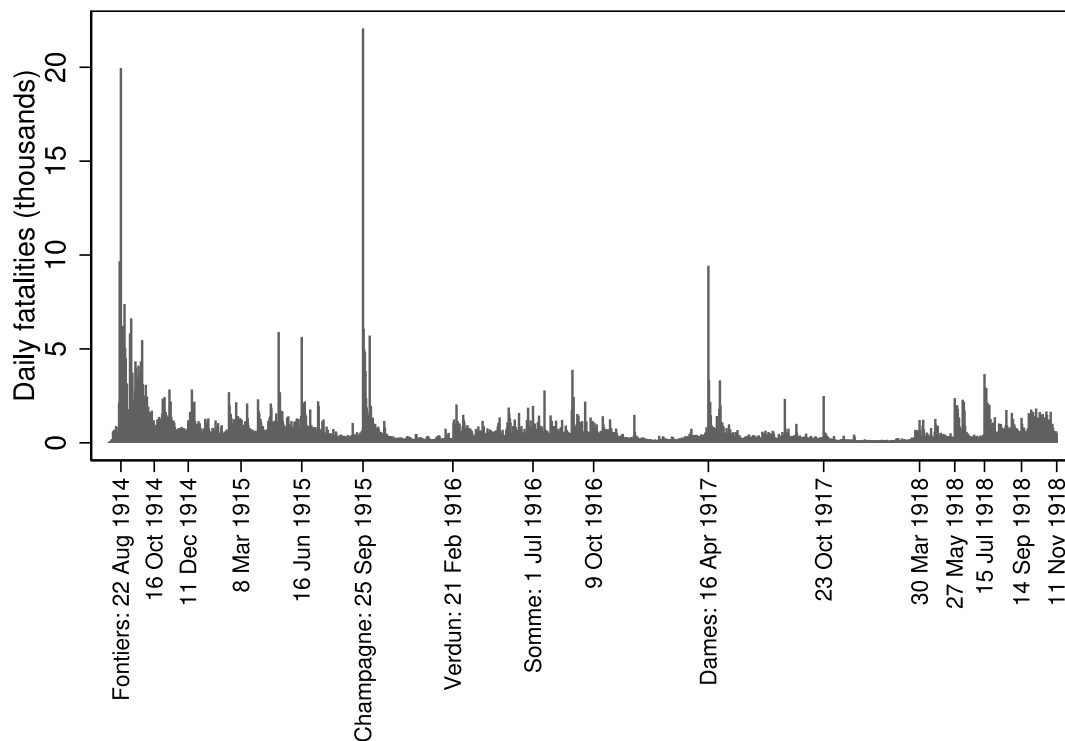


Fig. 5. Daily number of military fatalities.

Notes: This figure displays the number of daily fatalities in the MPF database between August 2, 1914 and November 11, 1918 among non-civilian MPFs born in France and for which the date of death is available (1,213,086 MPFs in total). Some notable, particularly deadly dates are highlighted with accompanying events on the front. Other dates are provided for reference.

in the MPF records was further authorized by the National Commission on Informatics and Liberty (CNIL) on July 10, 2014, for historical, statistical, and scientific purposes.⁵⁰

We disseminate the MPF database under the CC-BY 4.0 license on the Harvard Dataverse at <https://doi.org/10.7910/DVN/2APQZ6> in both Stata data format (dta) and text delimited format (txt) under the heading mpf_database (Gay and Grosjean, 2023). We also disseminate two additional datasets discussed in Section 4: the département- and municipality-level military death rates datasets in both Stata data format (dta) and text delimited format (txt) under the headings mpf_deathrate_dep and mpf_deathrate_municip, respectively, together with relevant shapefiles, under the headings DEATHRATE_DEP and DEATHRATE_MUNICIP, respectively (see Table 4 for a summary of the content of these data files).

6. Conclusion

This article describes the *Morts pour la France* (MPF) database. It includes a host of individual-level information for each of the 1.3 million French individuals who died as a result of the war: their first and last names, dates of birth and death, circumstances of death, recruitment status, military rank and unit at death, and locations of birth and recruitment. We also provide datasets of military death rates at the level of 87 départements and 35 thousand municipalities along with relevant shapefiles.

Matching records from the MPF database to auxiliary individual-level information constitutes a promising avenue of research. Given the availability of soldiers' names, dates of birth, and locations of birth and recruitment, it would be possible to match these records to military service records, which include information on height, weight, occupation, and other physiological characteristics as well as migration and military unit histories (Kesztenbaum, 2013). It would also be possible to match these records to the pre-war censuses of 1901, 1906, and 1911 to assess MPFs' family composition and living arrangements.

Data availability

The data described in this article is available at <https://doi.org/10.7910/DVN/2APQZ6>.

⁵⁰ The text of this deliberation is available at <https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000029312428>. The MdH web platform specifies that “[t]he public information visible on the Mémoire des Hommes website may be freely reused for purposes other than those of the public service mission for the needs of which the documents were produced or received according to the Etalab 2 open license model”. (<https://www.memoiredeshommes.sga.defense.gouv.fr/fr/article.php?laref=7>, accessed on September 25, 2023).

Table 4
Variables in the military death rates databases.

Dataset	Variable	Description	Values
Département level	dep_1911	Département identifier [TRF-GIS 1911]	String
	dep_name_1911	Département name [TRF-GIS 1911]	String
	mpf	Number of non-civilian male MPFs	Numeric
	pop_m_15t44	Male population 15–44 in 1911	Numeric
	deathrate	Military death rate (%) [MPF/male population 15–44 in 1911]	Numeric
Municipality level	dep_2015	Département identifier [INSEE 2015]	String
	dep_name_2015	Département name [INSEE 2015]	String
	geofla_2015	Municipality identifier [GEOFLA 2015]	String
	insee_2015	Municipality identifier [INSEE 2015]	String
	com_name_2015	Municipality name [INSEE 2015]	String
	mpf	Number of non-civilian MPFs	Numeric
	pop_1911	Municipality population in 1911 in 2015 geography	Numeric
	deathrate	Military death rate (%) [MPF/population in 1911]	Numeric

Notes: The military death rates databases are available in the Harvard Dataverse at <https://doi.org/10.7910/DVN/2APQZ6> (Gay and Grosjean, 2023).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.eeh.2023.101550>.

References

- Abramitzky, Ran, Boustan, Leah, Eriksson, Katherine, Feigenbaum, James, Pérez, Santiago, 2021. Automated linking of historical data. *J. Econ. Lit.* 59 (3), 865–918.
- Abramitzky, Ran, Delavande, Adeline, Vasconcelos, Luis, 2011. Marrying up: The role of sex ratio in assortative matching. *Am. Econ. J. Appl. Econ.* 3 (3), 124–157.
- Acemoglu, Daron, De Feo, Giuseppe, De Luca, Giacomo, Russo, Gianluca, 2022. War, socialism, and the rise of fascism: An empirical exploration. *Q. J. Econ.* 137 (2), 1233–1296.
- Ager, Philipp, Bursztyn, Leonardo, Leucht, Lukas, Voth, Hans-Joachim, 2022. Killer incentives: Rivalry, performance and risk-taking among German Fighter Pilots, 1939–45. *Rev. Econ. Stud.* 89 (5), 2257–2292.
- Antier, Chantal, 2008. Le recrutement dans l'empire colonial français, 1914–1918. *Guerr. Mond. Conflits Contemp.* 230 (2), 23–36.
- Aubry, Martine, de Oliveira, Mathieu, 2014. Une base de données sur les monuments aux morts: Histoire concrète et valorisation numérique. *In Situ* 25, 1–27.
- Aufay, Sandrine, 2014. Le site Mémoire des hommes et ses évolutions: Entre mémoire et histoire. *Gaz. Arch.* 236 (4), 71–83.
- Aufay, Sandrine, 2017. Le programme d'indexation collaborative du site Mémoire des hommes. *Rev. Hist. Armées* 3 (288), 93–98.
- Bach, André, 2010. La mort en 1914–1918. *Rev. Hist. Armées* 259, 23–32.
- Bailey, Roy E., Hatton, Timothy J., Inwood, Kris, 2016. Health, height and the household at the turn of the 20th century. *Econ. History Rev.* 69 (1), 35–53.
- Bailey, Roy E., Hatton, Timothy J., Inwood, Kris, 2023. Surviving the Deluge: British servicemen in World War I. *Econ. Hum. Biol.* 49, 101216.
- Beau, Laurent, 2017. From the local to the national: A new approach to casualties per département, 1914–1918. *Mouv. Soc.* 259 (2), 59–77.
- Bellis, Gil, Léger, Jean-François, Parant, Alain, 2022. La guerre de 1914–1918: un cataclysme démographique. effets immédiats et conséquences à long terme de la guerre de 1914–1918 sur la démographie française. *Espace Popul. Soc.* 2–3.
- Bloch, Marc, 1946. *L'Étrange Défaite*. Société des Éditions Franc-Tireur, Paris.
- Boehnke, Jörn, Gay, Victor, 2020. Replication data for: The missing men. World War I and female labor force participation [database]. <http://dx.doi.org/10.7910/DVN/AP1HZ8>, Harvard Dataverse.
- Boehnke, Jörn, Gay, Victor, 2022. The missing men. World War I and female labor force participation. *J. Hum. Resour.* 57 (4), 1209–1241.
- Bostrom, Alex, 2016. Supplying the front: French artillery production during the First World War. *Fr. Hist. Stud.* 39 (2), 261–286.
- Boulanger, Philippe, 2001. *La France Devant la Conscription: Géographie Historique d'une Institution Républicaine, 1914–1922*. Economica, Paris.
- Boulanger, Philippe, 2002. The 1914 conscripts: contribution of the French youth in forming a mass army. *Ann. Demogr. Hist.* 103 (1).
- Cagé, Julia, Dagorret, Anna, Grosjean, Pauline A., Jha, Saumitra, 2023. Heroes and villains: The effects of heroism on autocratic values and nazi collaboration in France. *Amer. Econ. Rev.* 113 (7), 1888–1932.
- Carozzi, Felipe, Pinchbeck, Edward, Repetto, Luca, 2023. Scars of war: The Legacy of WWI deaths on civic capital and combat motivation. *CESifo Working Paper*, (10298), CESifo.
- Chabord, Marie-Thérèse, 1973. Le Livre d'or de la Première Guerre mondiale: Un projet sans suite. *Rev. Hist. Armée* 2, 76–89.
- Chanel, Jean-François, 2006. *Vers l'Armée Nouvelle: République Conservatrice et Réforme Militaire, 1871–1879*. Presses Universitaires de Rennes, Rennes.
- Costa, Dora L., Kahn, Matthew, 2008. Heroes and Cowards: The Social Face of War. Princeton University Press, Princeton, NJ.
- Costa, Dora L., Kahn, Matthew E., Roudiez, Christopher, Wilson, Sven, 2018a. Persistent social networks: Civil war veterans who fought together co-locate in later life. *Reg. Sci. Urban Econ.* 70, 289–299.
- Costa, Dora L., Kahn, Matthew E., Roudiez, Christopher, Wilson, Sven, 2018b. Data set from the Union Army samples to study locational choice and social networks. *Data Brief* 17, 226–233.
- Cranfield, John, Inwood, Kris, 2015. A tale of two armies: Comparative growth in the mirror of the First World War. *Aust. Econ. Hist. Rev.* 55 (2), 212–233.
- Crépin, Annie, 2009. *Histoire de la Conscription*. Gallimard, Paris.
- Cristofoli, Pascal, Vouloir, Marie-Christine, Duménieu, Bertrand, Motte, Claude, 2021. Des chefs-lieux de Cassini aux communes de France (1756–1999) [database]. Didoména, Paris, https://didomena.ehess.fr/concern/data_sets/6395wb092.
- Cunningham, Angela R., 2018. After “it's over over there”: Using record linkage to enable the reconstruction of World War I veterans' demography from soldiers' experiences to civilian populations. *Hist. Methods* 51 (4), 203–229.
- Darlu, Pierre, Chareille, Pascal, 2020. The First World War and the disappearance of surnames in France: A trial estimation based on the Galton–Watson model. *Population* 75 (1), 97–128.

- De Juan, Alexander, Haass, Felix, Koos, Carlo, Riaz, Sascha, Tichelbaecker, Thomas, 2023. War and nationalism: How WWI battle deaths fueled civilians' support for the Nazi Party. *Am. Polit. Sci. Rev.* 1–19.
- Dippel, Christian, Ferrara, Andreas, 2023. Leadership and team cohesion under flying bullets: Captains and their companies in the Civil War. Working Paper.
- Dippel, Christian, Heblich, Stephan, 2021. Leadership in social networks: Evidence from the “Forty-Eighters” in the Civil War. *Amer. Econ. Rev.* 111 (2), 472–505.
- Fogel, Robert W., Costa, Dora L., Haines, Michael, Lee, Chulhee, Nguyen, Louis, Pope, Clayne, Rosenberg, Irvin, Scrimshaw, Nevin, Trussell, James, Wilson, Sven, Wimmer, Larry T., Kim, John, Bassett, Julene, Burton, Joseph, Yetter, Noelle, 2000. Aging of veterans of the Union Army: Version M-5 [database]. Center for Population Economics, University of Chicago Graduate School of Business, Department of Economics, Brigham Young University, and The National Bureau of Economic Research, Chicago.
- Fornasin, Alessio, 2017. The Italian Army's losses in the First World War. *Population* 72 (1), 39–62.
- Fornasin, Alessio, Breschi, Marco, Manfredini, Matteo, 2019. Deaths and survivors in war: The Italian soldiers in WWI. *Demogr. Res.* 40, 599–626.
- Fourie, Johan, Inwood, Kris, Mariotti, Martine, 2020. Military technology and sample selection bias. *Soc. Sci. Hist.* 44 (3), 485–500.
- Gay, Victor, 2020a. TRF-GIS communes (1870–1940)[database]. <http://dx.doi.org/10.7910/DVN/LZTZWE>, Harvard Dataverse.
- Gay, Victor, 2020b. TRF-GIS départements (1870–1940) [database]. <http://dx.doi.org/10.7910/DVN/ULQYM5>, Harvard Dataverse.
- Gay, Victor, 2020c. TRF-GIS military regions (1874–1940) [database]. <http://dx.doi.org/10.7910/DVN/SQPEUW>, Harvard Dataverse.
- Gay, Victor, 2020d. TRF-GIS military subdivisions (1874–1940) [database]. <http://dx.doi.org/10.7910/DVN/TAMQAG>, Harvard Dataverse.
- Gay, Victor, 2021. Mapping the Third Republic. A geographic information system of France (1870–1940). *Hist. Methods* 54 (4), 189–207.
- Gay, Victor, 2023. The intergenerational transmission of World War I on female labor. *Econ. J.* 133 (654), 2303–2333.
- Gay, Victor, Grosjean, Pauline, 2023. The Morts Pour la France database [database]. <http://dx.doi.org/10.7910/DVN/2APQZ6>, Harvard Dataverse.
- Georges, Raphaël, 2015. Quelle mémoire pour les soldats alsaciens-lorrains de la Grande Guerre? *Mouv. Soc.* 251 (2), 59–74.
- Gilles, Henri, 2010. La guerre 1914–1918: Les morts par départements et régions. In: Gwendal, Denis (Ed.), *Mémoire et Trauma de la Grande Guerre: Bretagne, Catalogne, Corse, Euskadi, Occitanie*. TIR, Rennes, pp. 275–323.
- Gilles, Henri, Guironnet, Jean-Pascal, Parent, Antoine, 2014. Geography of the deaths of WWI in France. *Rev. Écon.* 35 (3), 519–532.
- Gilot, Jean-Michel, 2020. “1 Jour- 1 Poilu”, une épopée commémorative citoyenne (2013–2018). *Gaz. Arch.* 258 (2), 121–133.
- Gilot, Jean-Michel, Grandjean, Martin, Clavert, Frédéric, 2018. 1914–1918: Quand la commémoration devient participative. *Le Temps des Médias* 31 (2), 219–229.
- Guillot, Olivier, Parent, Antoine, 2018. “Farewell life, farewell love”: Analysis of survival inequalities among soldiers who “died for France” during World War I. *Population* 73 (3), 413–444.
- Guillot, Olivier, Parent, Antoine, 2021. Les fusillés de la Grande Guerre sont-ils morts au nom de leurs idées pacifistes? *Revue de l'OFCE* 171 (1), 135–159.
- Guillot, Olivier, Parent, Antoine, 2022. ‘On vient de fusiller un des nôtres’: A quantitative study of military executions in the French army during WWI. *Def. Peace Econ.* (Forthcoming).
- Guironnet, Jean-Pascal, Parent, Antoine, 2019. Morts pour la France: Demographic or economic factors? *Def. Peace Econ.* 30 (2), 197–212.
- Hall, Andrew B., Huff, Connor, Kuriwaki, Shiro, 2019. Wealth, slaveownership, and fighting for the Confederacy: An empirical study of the American Civil War. *Am. Polit. Sci. Rev.* 113 (3), 658–673.
- Héran, François, 2014. Lost generations: The demographic impact of the Great War. *Popul. Soc.* 510 (4), 1–4.
- Huber, Michel, 1931. *La Population de la France Pendant la Guerre*. Yale University Press, New Haven.
- IGN, 2015. *GeoFla® Communes édition 2015 France Métropolitaine* [database]. IGN, Paris, <https://geoservices.ign.fr/geofla>.
- INSEE, 2012. *Fichiers des noms de famille de 1891 à 2000, édition 2008* [database]. <http://dx.doi.org/10.13144/lil-0697>, INSEE [producer], ADISP [distributor].
- INSEE, 2015. *Code officiel géographique au 1er janvier 2015* [database]. INSEE, Paris, <https://www.insee.fr/fr/information/2560698>.
- Kanazawa, Satoshi, 2007. Big and tall soldiers are more likely to survive battle: A possible explanation for the ‘returning soldier effect’ on the secondary sex ratio. *Hum. Reprod.* 22 (11), 3002–3008.
- Kesztenbaum, Lionel, 2013. L'étude des migrations grâce aux registres matricules militaires. *Popolazione e Storia* 14 (2), 9–38.
- Knowles, John, Vandenbroucke, Guillaume, 2019. Fertility shocks and equilibrium marriage-rate dynamics. *Internat. Econom. Rev.* 60 (4), 1505–1537.
- Koenig, Christoph, 2023. Loose cannons: War veterans and the erosion of democracy in Weimar Germany. *J. Econ. History* 83 (1), 167–202.
- Lagrou, Pieter, 2002. Les guerres, les morts et le deuil: Bilan chiffré de la Seconde Guerre Mondiale. In: Audouin-Rouzeau, Stéphane, Becker, Annette, Ingrao, Christian, Roussou, Henry (Eds.), *La Violence de Guerre, 1914–1945. Approches Comparées des Deux Conflits Mondiaux*. IHTP-CNRS, Bruxelles.
- Le Gall, Erwan, 2013. Saint-Malo, la France : Des multiples inscriptions territoriales du 47e régiment d'infanterie. In: Lagadec, Yann, Le Gall, Erwan, Bourlet, Michaël (Eds.), *Petites Patries dans la Grande Guerre*. Presses Universitaires de Rennes, Rennes, pp. 71–87.
- Lemarchand, Christian, 2008. Le fichier général des militaires de l'armée française décédés au cours de la Première Guerre Mondiale. *Rev. Histor. des Armées* 252, 1–4.
- Loez, André, 2008. Autour d'un angle mort historiographique: La composition sociale de l'armée française en 1914–1918. *Matériaux Pour l'Histoire de Notre Temps* 91 (3), 32–41.
- Loez, André, Mariot, Nicolas, 2008. Obéir, Désobéir. Les Mutineries de 1917 en Perspective. La Découverte, Paris.
- Marril, Jean-Marc, 2014. L'offensive à outrance: Une doctrine unanimement partagée par les grandes puissances militaires en 1914. *Rev. Hist. Armes* 49–64.
- McCalman, Janet, Kippen, Rebecca, McMeeken, Joan, Hopper, John, Reade, Michael, 2019. Early results from the ‘Diggers to Veterans’ longitudinal study of Australian men who served in the First World War. Short- and long-term mortality of early enlistees. *Hist. Life Course Stud.* 8, 52–72.
- Motte, Claude, Séguis, Isabelle, Théré, Christine, 2003. *Communes d'Hier, Communes d'Aujourd'hui. Les Communes de la France Métropolitaine, 1801–2001*. Dictionnaire d'Histoire Administrative. INED, Paris.
- Motte, Claude, Vouloir, Marie-Christine, 2007. Le site cassini.ehess.fr: Un instrument d'observation pour une analyse du peuplement. *Bull. du Comité Français de Cartograph.* 68–84.
- Moura, Sylvain, 2018. Ce que les statistiques sur les soldats morts révèlent des combats en Meuse et en Somme pendant la Première Guerre mondiale. *EcoDef Stat.* 116, 1–8.
- Petit, Vincent, 2016. “Mort pour la France” ou le triomphe d'une théologie nationale (1914–1915). *Hist. Monde Cult. Rel.* 37 (1), 175–179.
- Porte, Rémy, 2005. *La Mobilisation Industrielle: Premier Front de la Grande Guerre?*. Paris, 14–18 Éditions.
- Prost, Antoine, 1992. Mémoires locales et mémoires nationales: Les monuments de 1914–1918 en France. *Guerr. Mond. Conflits Contemp.* 167, 41–50.
- Prost, Antoine, 2008. Compter les vivants et les morts: L'évaluation des pertes françaises de 1914–1918. *Mouv. Soc.* 222 (1), 41–60.
- Richard, Damien, 2013. Le Livre d'or des Morts pour la France: Construction et usages d'un monument de papier. *Actes des Congrès Nationaux des Soc. Histor. Sci.* 136, 143–153.
- Richard, Damien, 2023 [2010]. Les Livres d'or des Morts pour la France de la Première Guerre Mondiale [accessed on september 25, 2023]. <https://www.siv.archives-nationales.culture.gouv.fr/siv/cms/content/helpGuide.action?preview=false&uuid=16e841bb-ffaa-4359-bf09-9714cb25c1ac>.
- Rohrbasser, Jean-Marc, 2014. *Bouleversements Démographiques de la Grande Guerre*. INED, Paris.
- Rolland, Denis, 2005. Le suicide aux armées en 1914–1918: Une première approche quantitative globale. In: Cazals, Rémy, Picard, Emmanuelle, Rolland, Denis (Eds.), *La Grande Guerre: Pratiques et Expériences*. Éditions Privat, Paris, pp. 269–279.
- Salem, Ariane, 2023. The negative impacts of colonization on the local population: Evidence from Morocco. Working Paper.
- Steg, Jean-Michel, 2013. Le Jour le Plus Meurtier de l'Histoire de France: 22 Août 1914. Fayard, Paris.
- Van Galen, Cornelis W., 2019. Creating an audience: Experiences from the surinamese slave registers crowdsourcing project. *Hist. Methods* 52 (3), 178–194.
- Vandenbroucke, Guillaume, 2014. Fertility and wars: The case of World War I in France. *Am. Econ. J.: Macroecon.* 6 (2), 108–136.
- Weaver, Michael, 2022. “Let our ballots secure what our bullets have won”: Union veterans and the making of radical reconstruction. *Am. Polit. Sci. Rev.* 116 (4), 1309–1324.

VICTOR GAY

1 Esplanade de l'Université, T.458
31080, Toulouse, France



victorgay.me
victor.gay@tse-fr.eu
+ (33) 6 22 22 51 70

CURRENT ACADEMIC POSITIONS

Assistant Professor of Economics, Department of Social and Behavioral Sciences, Toulouse School of Economics (TSE)	Since 2018
Research Fellow, Institute for Advanced Study in Toulouse (IAST)	Since 2018

PAST ACADEMIC POSITION

Visiting Assistant Professor, Department of Economics, University of Michigan	2023–24
---	---------

EDUCATION

<i>PhD</i> , Economics, The University of Chicago	2018
<i>Master</i> , Economics, The University of Chicago	2013
<i>Master</i> , Economics, University of Cergy (THEMA)	2012
<i>MBA</i> , ESSEC Business School	2012

RESEARCH FIELDS

Economic history, political economy, labor economics.

WORKING PAPERS

- WP3 Eva Davoine, Joseph Enguehard, Victor Gay, and Igor Kolesnikov. February 2026. “Mapping the Grandes Gabelles in Early Modern France.” TSE Working Paper 26–1716. HAL: hal-05501504
- WP2 Raphaël Franck and Victor Gay. August 2025. “Urbanization and Electoral Success: Lawyers and Workers in Interwar France.” CEPR Discussion Paper 18737. HAL: hal-04364519
- WP1 Nick Huntington-Klein, Claus C. Pörtner, and the Many-Economists Research Collective [146 researchers, including Victor Gay]. May 2025. “The Sources of Researcher Variation in Economics.” NBER Working Paper 33729. Revision requested at the *Journal of Economic Literature*. DOI: [10.3386/w33729](https://doi.org/10.3386/w33729)

JOURNAL ARTICLES

Analysis Articles

- A9 Michael Albertus and Victor Gay. 2025. “State-Building and Rebellion in the Run-Up to the French Revolution.” Forthcoming at the *American Political Science Review*. DOI: [10.1017/S0003055425101342](https://doi.org/10.1017/S0003055425101342)
- A8 Victor Gay, Paula E. Gobbi, and Marc Goñi. 2025. “Revolutionary Transition: Inheritance Change and Fertility Decline.” Forthcoming at the *Journal of Political Economy*. DOI: [10.1086/739821](https://doi.org/10.1086/739821)
- A7 Margot Dazey and Victor Gay. 2025. “The Mosque Nearby. Visible Minorities and Far-Right Support in France.” *Comparative Political Studies*, 58(9): 2000–48. DOI: [10.1177/00104140241283015](https://doi.org/10.1177/00104140241283015)
- A6 Victor Gay. 2023. “The Intergenerational Transmission of World War I on Female Labor.” *The Economic Journal*, 133(654): 2303–33. DOI: [10.1093/ej/uead029](https://doi.org/10.1093/ej/uead029)
- A5 Jörn Boehnke and Victor Gay. 2022. “The Missing Men: World War I and Female Labor Force Participation.” *Journal of Human Resources*, 57(4): 1209–41. DOI: [10.3368/jhr.57.4.0419-10151R1](https://doi.org/10.3368/jhr.57.4.0419-10151R1)
- A4 Michael Albertus and Victor Gay. 2019. “No Better Time Than Now: Future Uncertainty and Private Investment Under Dictatorship.” *Economics & Politics*, 31(1): 71–96. DOI: [10.1111/ecpo.12120](https://doi.org/10.1111/ecpo.12120)
- A3 Victor Gay, Daniel L. Hicks, Estefania Santacreu-Vasut, and Amir Shoham. 2018. “Decomposing Culture: An Analysis of Gender, Language, and Labor Supply in the Household.” *Review of Economics of the Household*, 16(4): 897–909. DOI: [10.1007/s11150-017-9369-x](https://doi.org/10.1007/s11150-017-9369-x)
- A2 Michael Albertus and Victor Gay. 2017. “Unlikely Democrats: Economic Elite Uncertainty Under Dictatorship and Support for Democratization.” *American Journal of Political Science*, 61(3): 624–41. DOI: [10.1111/ajps.12277](https://doi.org/10.1111/ajps.12277)
- A1 Estefania Santacreu-Vasut, Amir Shoham, and Victor Gay. 2013. “Do Female/Male Distinctions in Language Matter? Evidence from Gender Political Quotas.” *Applied Economics Letters*, 20(5): 495–8. DOI: [10.1080/13504851.2012.714062](https://doi.org/10.1080/13504851.2012.714062)

Data Articles

- D6 Sandra Brée, Victor Gay, Marion Leturcq, Baptiste Coulmont, Yoann Doignon, Thomas Constum, Thierry Paquet, and Pierrick Tranouez. “POPP. An OCR-Generated Database of the Population Censuses of Paris (1926–1936).” *Historical Life Course Studies*, 16:3–28. DOI: [10.52024/hlcs18627](https://doi.org/10.52024/hlcs18627)
- D5 Victor Gay. 2025. “The Jean Nicolas Database. The French Rebellion, 1661–1789.” *Data & Corpus*, 1: 15892. DOI: [10.46298/dc.15892](https://doi.org/10.46298/dc.15892)
- D4 Victor Gay, Paula E. Gobbi, and Marc Goñi. 2024. “The Atlas of Local Jurisdictions of Ancien Régime France.” *Journal of Historical Geography*, 84: 49–60. DOI: [10.1016/j.jhgc.2024.03.008](https://doi.org/10.1016/j.jhgc.2024.03.008)

- D3 Victor Gay, Paula E. Gobbi, and Marc Goñi. 2024. “The Customary Atlas of Ancien Régime France.” *Explorations in Economic History*, 93: 101588. DOI: [10.1016/j.eeh.2024.101588](https://doi.org/10.1016/j.eeh.2024.101588)
- D2 Victor Gay and Pauline Grosjean. 2023. “Morts Pour la France: A Database of French Fatalities of the Great War.” *Explorations in Economic History*, 90: 101550. DOI: [10.1016/j.eeh.2023.101550](https://doi.org/10.1016/j.eeh.2023.101550)
- D1 Victor Gay. 2021. “Mapping the Third Republic. A Geographic Information System of France (1870–1940).” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 54(4): 189–207. DOI: [10.1080/01615440.2021.1937421](https://doi.org/10.1080/01615440.2021.1937421)

Replication Articles

- R3 Olivia Miske, Brian A. Nosek, Timothy M. Errington, and the SCORE Research Collective [124 researchers, including Victor Gay]. 2026. “Investigating the Reproducibility of the Social and Behavioral Sciences.” Forthcoming at *Nature*.
- R2 Victor Gay. 2023. “Culture: An Empirical Investigation of Beliefs, Work, and Fertility. A Verification and Reproduction of Fernández and Fogli (2009).” *Journal of Comments and Replications in Economics*, 2: 1–15. DOI: [10.18718/81781.28](https://doi.org/10.18718/81781.28)
- R1 David Le Bris and Victor Gay. 2023. “Distance to Innovations, Kinship Intensity, and Psychological Traits.” *PLoS ONE*, 18(1): e0279864. DOI: [10.1371/journal.pone.0279864](https://doi.org/10.1371/journal.pone.0279864)

BOOK CHAPTERS

- BC3 Victor Gay. 2025. “Un data paper en SHS : pourquoi, pour qui, comment ?” [A data paper in the social sciences: why, for whom, and how?]. In Christine Kosmopoulos and Joachim Schopfel (Eds.), *Publier, partager, réutiliser les données de la recherche: les data papers et leurs enjeux*, 66–74. Lille: Presses universitaires du Septentrion. HAL: [hal-03434216](https://hal.archives-ouvertes.fr/hal-03434216)
- BC2 Victor Gay. 2025. “La mesure du travail des femmes dans les recensements de la population au tournant du 20^e siècle” [Measuring women’s labor in the population censuses at the turn of the twentieth century]. In Alain Trannoy and Arundhati Virmani (Eds.), *Économistes et historiens, dialogue de sourds?*, 65–78. Paris: Odile Jacob. HAL: [hal-04737809](https://hal.archives-ouvertes.fr/hal-04737809)
- BC1 Victor Gay, Daniel L. Hicks, and Estefania Santacreu-Vasut. 2016. “Language and Gender Roles among Immigrants to the U.S.: A Historical Perspective.” In Paola Paolini (Ed.), *I Mondi Delle Donne*, 165–79. Roma: Edi Cusano. HAL: [hal-02523125](https://hal.archives-ouvertes.fr/hal-02523125)

CONFERENCE PROCEEDINGS

- CP3 Victor Gay and Sébastien Poublanc. 2023. “Espaces construits par la puissance publique et espaces vécus par les justiciables dans le royaume de France (1661–1789)” [Constructed spaces by public authorities and experienced spaces by litigants in the Kingdom of France (1661–1789)]. In *Actes du congrès de la Fédération historique de la région Occitanie*. HAL: [hal-04301060](https://hal.archives-ouvertes.fr/hal-04301060)

- CP2 Lucas Bourel, Nathalie Hernandez, Nathalie Aussenac-Gilles, Victor Gay, and Sébastien Poublanc. 2022. “Graphes de connaissances pour représenter et analyser l’évolution des territoires en histoire” [Knowledge graphs for representing and analyzing the evolution of territories in history]. In Nicolas Lasolle, Oliver Bruneau, and Jean Liever (Eds), *Actes des journées humanités numériques et Web sémantique*, 23–37. Nancy, France. HAL: [hal-03683727](#)
- CP1 Victor Gay, Daniel L. Hicks, and Estefania Santacreu-Vasut. 2016. “Migration as a Window into the Coevolution between Language and Behavior.” In Sean G. Roberts, Christine Cuskley, Luke McCrohon, Lluís Barceló-Coblijn, Olga Fehér and Tessa Verhoef (Eds.), *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*. HAL: [hal-02523115](#)

NON-PEER REVIEWED PUBLICATIONS

- NPR4 Sandra Brée and the POPP Research Collective [11 researchers, including Victor Gay]. 2025. “Paris 100 Years Ago: More People than Today—and Mostly Born Elsewhere.” *Population & Sociétés*, 636, 1–4. DOI: [10.3917/popsoc.636.0001](#)
- NPR3 Victor Gay and Sébastien Poublanc. 2022. “ObARDI : un projet interdisciplinaire pour comprendre la construction de l’État moderne (1661–1789)” [ObARDI: an interdisciplinary project to understand the construction of the Modern State (1661–1789)]. *Les Cahiers de Framespa*, 41, 1–12. DOI: [10.4000/framespa.13379](#)
- NPR2 Victor Gay. 2019. Review of Roger L. Ransom’s (2018) *Gambling on War: Confidence, Fear and the Tragedy of the First World War*. Cambridge: Cambridge University Press. Reviewed for *EH.net*. HAL: [hal-02524006](#)
- NPR1 Victor Gay, Estefania Santacreu-Vasut, and Amir Shoham. 2013. “The Grammatical Origins of Gender Roles.” Berkeley Economic History Laboratory Working Paper 2013–03. HAL: [hal-04046900](#)

ORIGINAL DATASETS

- OD8 Eva Davoine, Joseph Enguehard, Victor Gay, Igor Kolesnikov. 2026. *The Grandes Gabelles in 1665 France* (Version 1.0). Recherche Data Gouv. DOI: [10.57745/DO39SH](#) (Access available upon request).
- OD7 Sandra Brée, Victor Gay, Marion Leturcq, Thierry Paquet, Thomas Constum, Pierrick Tranouez, and Nicolas Kempf. 2025. *Parisian Population Censuses - 1926–1936* (Version 0). CNRS [Producer], Progedo-Adisp [Distributor]. DOI: [10.13144/lil-1719](#) (Access available upon request).
- OD6 Victor Gay. 2025. *The Jean Nicolas Database* (Version 2.0). Harvard Dataverse [Distributor]. DOI: [10.7910/DVN/ANQXMQ](#)
- OD5 Victor Gay, Paula E. Gobbi, and Marc Goñi. 2023. *Bailliages in 1789 France* (Version 2.0). Harvard Dataverse [Distributor]. DOI: [10.7910/DVN/T8UXHK](#)

- P4 EXO-POPP (“Optical Extraction of Handwritten Named Entities for the Marriage Certificates of the Population of Paris (1880–1940)”), [ANR-21-CE38-0004](#) (PI: Sandra Brée), 2022–26.
- P3 IDED (“Inheritance, Demographics, and Economic Development”), [ERC-947654](#) (PI: Paula E. Gobbi), 2021–25.
- P2 POPP (“Océreriser les recensements de la population parisienne de l’entre-deux-guerres”), [Collex-Persée POPP](#) (PI: Sandra Brée), 2020–22.
- P1 SCORE (“Systematizing Confidence in Open Research and Evidence”), [DARPA-OSF](#) (PIs: Tim Errington and Andrew Tyner), 2019–26.

PROFESSIONAL ACTIVITIES

External Seminars (42)

2026: Archives Henri-Poincaré, IÉSEG School of Management, Paris Dauphine University. **2025:** Bologna University, University Pablo de Olavide de Sevilla, Yonsei University. **2024:** Bocconi University, Chapman University, Northwestern University, UC Irvine, University of Toronto. **2023:** Boston University / Harvard University (joint), George Mason University, Paris Cité University (LIRAES), The University of Texas at El Paso, University of Michigan (2), University of Pittsburgh. **2022:** Aix-Marseille School of Economics, IÉSEG School of Management, UCLouvain, University Toulouse 2 (Framespa). **2021:** ENS Lyon, Observatoire Midi-Pyrénées, Université Libre de Bruxelles, University of Bordeaux, Sciences Po (CDSP), Sciences Po (CHSP). **2020:** Paris School of Economics. **2019:** Aix-Marseille School of Economics, Bocconi University, Paris School of Economics, Rutgers University. **2018:** Toulouse School of Economics, University of Southern Denmark, University of Toronto. **2017:** Chapman University, Northwestern University, Stanford University. **2015:** Cergy Paris University, Harris School of Public Policy, Paris School of Economics.

Conferences and workshops (36)

2026: ANR SocFace Workshop (INED), Workshop “Les 30 ans de Cybergeog: 1996–2026” (CNRS). **2025:** Conference “Fascism at the Ballot Box (Europe, 1918–45)” (University of Bordeaux), Workshop “Méthodes et outils des humanités numériques. Conception et exploration statistique de bases de données quantitatives” (University of Toulouse 2), Conference “Proyectos integrales de HGIS” (University of Santiago de Compostela), Conference Data & Corpus “Les articles de données en sciences humaines et sociales” (University of Lorraine). **2024:** ASREC annual conference (Chapman University), ASSA annual meeting (San Antonio), Symposium ANR-COMMUNES (Paris), EEA-ESEM annual meeting (Erasmus University), Rencontres d’histoire et d’histoire économique d’Agen (BSE), Workshop “Méthodes et outils des Humanités numériques” (University Toulouse 2), Webinar “Réutiliser des données” (GTSO-Données). **2023:** AFSE annual congress (Sciences Po), CNEH annual conference (University of Toronto), Webinar “Des Sources aux systèmes d’information géographiques: des outils pour la cartographie dans les humanités numériques” (DYPAC UVSQ / Dlab SHS Paris-Saclay), SSHA annual conference (Washington), VonGrep Workshop in Economic and Entrepreneurial History (UCLA). **2022:** Rencontres d’histoire et d’histoire économique d’Agen (BSE). **2021:** AFSE annual congress (IÉSEG

School of Management), Conference #dhnord2021 “Publier, partager, réutiliser les données de la recherche: les data papers et leurs enjeux” (MESHS), Rencontres d’histoire et d’histoire économique d’Agen (BSE), Seminar “Des problèmes en partage, des méthodologies divergentes, le regard croisé d’historiens et d’économistes” (Fondation des Treilles), Webinar “Des Sources aux systèmes d’information géographiques: des outils pour la cartographie dans les humanités numériques” (DYPAC UVSQ / Dlab SHS Paris-Saclay), Virtual workshop “Life and Death in the WWI Trenches: Causes and Consequences” (Center for Economic History at the Australian National University). **2020:** Workshop “Data paper: une incitation à la qualification et à la réutilisation des jeux de données” (MITI, CNRS), ASSA annual meeting (San Diego), EHS annual conference (Oxford). **2019:** EHS annual conference (Queen’s University Belfast), SSHA annual conference (Chicago), SYME annual meeting (Université Libre de Bruxelles). **2018:** Galatina Summer meetings, HEDG Mini-Workshop (University of Southern Denmark). **2017:** Development Day at Northwestern, EHA annual meeting (San Jose). **2016:** IEA annual meeting (Chicago).

Conference discussant (7)

2025: Digital Historical Political Economy Conference (CBS), TSE digital economics conference (TSE), Virtual Workshop in Historical Political Economy (VWHPE). **2024:** ASSA annual meeting (San Antonio). **2022:** IAST workshop “Leadership: Does it Matter?” (Toulouse). **2018:** EHA annual meeting (Montreal). **2016:** IEA annual meeting (Chicago).

Academic journals refereeing activity (104)

Economics (66): *American Economic Review* (3), *Journal of Political Economy*, *The Quarterly Journal of Economics* (2), *The Review of Economic Studies* (4), *American Economic Journal: Applied Economics* (2), *American Economic Review: Insights*, *Journal of the European Economic Association* (4), *The Economic Journal* (5), *Applied Economics* (2), *Applied Economics Letters*, *Eastern Economic Journal*, *Economic Inquiry* (2), *Economics & Human Biology* (5), *Economics & Politics* (3), *European Economic Review* (3), *Financial History Review*, *Journal of Comparative Economics* (2), *Journal of Demographic Economics* (2), *Journal of Development Economics*, *Journal of Economic Behavior & Organization* (4), *Journal of Labor Economics*, *Journal of Labor Research*, *Journal of Population Economics* (2), *Journal of Public Economics* (2), *Labour Economics*, *Public Choice* (2), *The Economics of Transition*, *The Journal of Economic History* (5), *The World Bank Economic Review*, *World Development*.

Political science (16): *American Political Science Review* (5), *The Journal of Politics*, *Comparative Political Studies* (4), *Democratization* (2), *Electoral Studies*, *European Journal of Political Research*, *European Politics and Society*, *Quarterly Journal of Political Science*.

Other (22): *Children and Youth Services Review*, *Historical Methods: A Journal of Quantitative and Interdisciplinary History* (3), *Humanités Numériques* (4), *Geographies*, *International Social Science Journal*, *International Journal of Geo-Information*, *Journal of International Migration and Integration*, *Journal of Open Humanities Data* (6), *Land* (3), *Strategic Management Journal*.

Grant refereeing activity

Agence Nationale de la Recherche (ANR).

Conference refereeing activity

2025–26: AFÉPOP conference. **2019–26:** SYME annual meeting. **2021:** IUSSP international population conference.

Departmental and professional services

Member of the executive committee of EH.net	Since 2024
Member of the scientific committee of <i>Data & Corpus</i>	Since 2024
Co-pilot of the UT1 Committee for open science	Since 2024
Scientific advisor of the University Data Platform of Toulouse (PROGEDO)	Since 2021
Member of the TSE Research Committee	2023
Member of the IAST Executive Committee, researchers' representative	2019–21

Co-organizer of the Virtual Workshop on French Economic History	2025–26
Co-organizer of the Toulouse Summer School in Quantitative Social Sciences	2020–25
Co-organizer of the TSE BID Seminar	2020–23
Co-organizer of the TSE BID Workshop	2018–23
Co-organizer of the IAST Epistemology Workshop	2021–22
Co-organizer of the IAST Political Economy Workshop	2022
Co-organizer of the Toulouse History Workshop	2019–21
Co-organizer of the 7th Toulouse Economics & Biology Workshop	2019

TEACHING**Toulouse School of Economics**

Economic History (M1)	2020–25
Institutions and Long-Run Development (M2 PPD)	2020–25
Applied Econometrics (M1)	2020–25
Development: Theory, Public Policy, and Historical Perspectives (M2 ETE)	2021–23
Natural Experiments in History (PhD)	2021–22
Advanced Public Policy and Applied Economics (PhD)	2019–20
TSE Inaugural Lecture	2020
Microeconomics (Undergraduate)	2019

PhD thesis supervision (3)

Capucine Ferré (external committee, Paris Dauphine University-PSL)	2024–26
Yifan Zhang	2021–25
Luisa Carrer	2020–24

M1 (12) and M2 (6) theses supervision	2019–24
---------------------------------------	---------

University of Michigan

Global Economic History (Undergraduate)	2024
---	------

ESSEC Business School

Advanced Macroeconomics (MBA)

2015–16

The University of Chicago

Elements of Economic Analysis (Undergraduate)

2014–15

MEDIA AND OUTREACH

2024: “Témoignage d’un auteur de *data paper*.” *Quéro mag*, numéro spécial. **2023:** “In the Shadows of War.” *TSE Mag* 25, 17; “Learning from the past.” *IAST Magazine* 20, 12–3. **2022:** “What can political leaders learn from history?” Crossing Channels Podcast. **2021:** “Female labor throughout the twentieth century.” Coffee talk (Tseconomist); “Covid-19 : Comme si nous vivions la grippe espagnole de 1918 et la Grande dépression de 1929.” *L’Opinion*; “Que peut nous apprendre l’histoire sur la Covid-19?” *TSE Mag* 21, 30–1; “La Grande Guerre, une aubaine pour l’industrie textile chinoise.” *Le Point, Phébé*. **2020:** TSE-IAST Public Webinar (Covid and climate); “Comment la grippe espagnole a augmenté la méfiance.” *Le Point, Phébé*. **2019:** “Gender and language.” Coffee talk (Tseconomist); Prendre un café avec un.e chercheur.e. (CNRS); “War and the sexes.” *IAST Magazine* 14, 17. **2018:** “Guerre et genre.” *TSE Mag* 18, 7.

DISSERTATION

The Legacy of the Missing Men: World War I and Female Labor in France Over a Century. 2018.

Granted by the Department of Economics at The University of Chicago.

Committee: Richard Hornbeck (chair), Steven D. Levitt, Derek A. Neal, and James A. Robinson.