Forthcoming *Econometrica*

Taxing Identity
Theory and Evidence from Early Islam

Mohamed Saleh* Jean Tirole*

March 10, 2021

Abstract

A ruler who does not identify with a social group, whether on religious, ethnic, cultural or socioeconomic grounds, is confronted with a trade-off between taking advantage of the out-group population's eagerness to maintain its identity and inducing it to "comply" (conversion, quit, exodus or any other way of accommodating the ruler's own identity). This paper first nests economists' extraction model, in which rulers are revenue-maximizers, within a more general identity-based model, in which rulers care also about inducing people to lose their identity, both in a static and an evolving environment. The paper then constructs novel data sources to test the implications of both models in the context of Egypt's conversion to Islam between 641 and 1170. The evidence comes in support of the identity-based model.

Keywords: Islam, poll tax, identity taxation, Laffer curve, legitimacy.

JEL numbers: D82, H2, N45, Z12.

*Toulouse School of Economics and Institute for Advanced Study in Toulouse, 1 Esplanade de l'Université, Toulouse 31080, France. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 669217 - ERC MARKLIM). Financial support from the ANR-Labex IAST is also gratefully acknowledged. We are grateful to participants at numerous conferences and seminars and to Ran Abramitzky, Robert Barro, Roland Bénabou, Dora Costa, Barry Eichengreen, James Fenske, Roberto Galbiati, Oded Galor, Avner Greif, Walker Hanlon, Richard Hornbeck, Emir Kamenika, Timur Kuran, Alessandro Lizzeri, Joel Mokyr, Thomas Piketty, Giacomo Ponzetto, Emmanuel Saez, Noam Yuchtman, Fabrizio Zilibotti, and four referees for their very helpful comments. We thank Amirreza Ahmadzadeh, Paul-Henri Moisson, Norhan Muhab, Roxana Pozo, and Cyril Thomson, for their excellent research assistantship.

"Muhammad was sent as a prophet and not as a tax collector."

Umar II, the Umayyad Caliph from 717 to 720 CE

1 Introduction

1.1 Motivation and main insights

Hostility toward populations on the ground of their religious, ethnic, linguistic, cultural, economic, political, or sexual-orientation identity, is commonplace. While a voluminous literature covers rulers' violent (non-price) policies against these "unwanted" populations,[1] the non-violent (price) approach of taxing identity has received much less attention. Yet, explicit taxation of identity was frequent; for instance, a poll (head) tax was levied by the early Arab Caliphate and by subsequent Muslim-ruled polities on their non-Muslim subjects up to the mid-19th century. Alternatively, identity taxation may be more subdued, as when local governments discriminate among neighborhoods when locating amenities, or when countries restrict access to public goods to permanent residents or citizens or (in dictatorships) members of the ruling party.[2]

Taxing identity exposes rulers to a tradeoff between extracting agents' willingness to pay for keeping their identity and inducing them to lose it (convert, assimilate, quit the organization or the country...). There are two views on how they solve this dilemma. The economists' typical view is that rulers, especially in pre-modern polities, are revenue-maximizers, perhaps constrained by their state capacity.[3] While the *extraction model* has much merit, a second view of identity taxation is that rulers care not only about money, but also about inducing people to lose their identity, even at the expense of lower tax revenue. This may be due to an ideological mission to win converts or to a political goal to expand a ruler's support base. Accordingly, we study the paradigm in which a ruler has both revenue and identity objectives. In this *identity-based model*,

[1] For example, Voigtländer and Voth (2012) and Anderson et al. (2017) study anti-Semitic persecutions.

[2] Examples of both types of identity taxation abound. Various European polities imposed a tax on Jews up to the 1800s. Romans levied a poll tax from which citizens were exempted, until Roman citizenship became universal under Emperor Caracalla. The Protestant Reformation was characterized by a shift from identity taxation, the tithe imposed by the Catholic church on its adherents, to secular taxation (Dittmar and Meisenzahl 2020). In constitutional countries, taxes can be targeted less explicitly toward unwanted populations. For instance, the 1942 one-off Varlik Vergisi (wealth) tax in Turkey was imposed on all citizens' fixed assets (Artunç and Agir 2017). While on paper a non-discriminatory tax, it affected most severely non-Muslims, who controlled a large portion of the economy, and led to their exodus. Communist countries used Communist Party membership to allocate positions. Local and national governments' policies with respect to the provision of local public goods for migrants (training, housing, low bureaucratic hassle, intolerance toward harassment...) is yet another example (Tabellini 2020).

[3] For example, see De Long and Shleifer (1993) and Besley and Persson (2011).

which nests the extraction model, the ruler optimally levies two taxes: a uniform tax, which mechanically has no impact on conversions and therefore is purely extractive, and a discriminatory one levied on those who maintain their identity, which does affect conversions. A straightforward implication of this more general model is that, absent delegation problems, the discriminatory tax lies on the downward-sloping side of the corresponding Laffer curve.

Our historical context is taxation in the aftermath of the Arab conquest of the then-Coptic Christian Egypt in 641 CE, until the fall of the Fatimid Caliphate in 1170.[4] The Arab Caliphate levied both a discriminatory (poll) tax on religion, imposed on non-Muslims (initially all Egyptians) and removed upon conversion to Islam, and a non-discriminatory (uniform) one on land that was paid regardless of the taxpayer's religion.[5] While this system is consistent with the extraction model and its more general version emphasizing state capacity, the identity-based model was (implicitly) endorsed on empirical grounds by early-20$^{\text{th}}$-century historians such as Wellhausen (1902), Becker (1902), Bell (1910), Grohmann (1932); they postulated that tax-induced conversions led to a loss in poll tax revenue over time, which is only possible under the identity-based model. Indeed, faced with a deteriorating poll tax revenue, Umar II, who was renowned for his piety, called for more conversions at the cost of a lower tax revenue, suggesting an identity-based motive.[6]

The historical context offers a number of advantages to study identity taxation. First, authorities automatically validated conversions to Islam, in contrast with situations in which attempts at adopting the politically dominant identity (e.g., permanent residency or naturalization) can be rejected. Second, there were two forms of taxation: discriminatory and uniform, which generates interesting dynamics of their co-evolution. Finally, conversion to Islam was irreversible because of the death penalty on apostates. While definitive exit from the tax base is a good approximation in many identity taxation contexts, the Caliphate institutionalized the irreversibility of the conversion decision.

Cross-sectional analysis. Section 3 develops the framework. Taxation is delegated by the ruler/central authority (CA) to local authorities (LAs). Districts differ in the identity strengths of the local collector (the LA) and of the population. LAs levy a uniform and

[4]While the Arab Caliphate enforced its tax system throughout all its conquered territories, we limit ourselves in this paper to Egypt, because it is where local-level records on churches and the poll tax rate and revenue, and Egypt-level medieval narratives on poll tax hikes and conversions, survived.

[5]Between 641 and circa 750, non-Muslim landholders also paid a higher land tax rate. By 750, the difference in land tax rate was abolished (see the end of the introduction and Sections 2 and 4.2.8).

[6]Consistent with the identity-based model, a leading modern historian of early Islam emphasizes "the problems [the early Arab Caliphate] faced trying to ensure a continuous source of fiscal income while simultaneously serving the Muslim mission to win converts" Sijpesteijn (2013, p. 189).

a discriminatory tax. The CA has no local presence and can only request a transfer from the LA.

An extractive ruler does tax agents' identity. Maximal extraction requires maximizing separately revenues on the uniform and the discriminatory taxes. In the identity-based paradigm, the CA's fiscal motivation is two-fold: extract as much revenue as possible and induce conversions. The latter motivation alters the discriminatory tax, which induces conversions, but not the uniform tax, which does not and remains purely extractive. This introduces a divergence relative to the extraction model, the consequences of which we investigate theoretically and empirically.

Section 3's focus is on agency. The delegation of the collection of taxes to local tax collectors is of no consequence in the extraction model, at least if the ruler has enough information on local conditions: the CA and LAs both aiming at maximizing revenue creates congruence between them. Not so in the identity-based model, as the local authorities may not share the ruler's identity preferences. A case in point is early Islamic Egypt, in which Copts rather than Arabs administered tax collection in many districts. The main theoretical result here is that the discriminatory tax still lies on the downward-sloping side of the Laffer curve, but ceteris paribus an LA with a stronger identity strength levies a higher discriminatory tax, induces more conversions and raises less revenue. We also demonstrate that the population's own identity strength increases the discriminatory tax rate and revenue, but mitigates conversions, an implication that is common across both models.

To test apart the cross-sectional implications of the two models, we exploit the local variation in early Islamic Egypt across *kuras* (Egypt's administrative units in 641–1100) in the identity strength of both the LAs and Coptic taxpayers.[7] We think of the CA as either Egypt's governor or the Caliph (see the time-series part for a discussion). We think of the LAs, not only as *kura* headmen, but as the entire local bureaucracies. We measure the LAs' identity strength by a *kura*-level dummy variable that takes value 1 if an Arab tribe settled permanently in 700–969. This variable arguably captures the level of Arabization of the LAs: In *kuras* where Arabs settled, they were more likely to replace Coptic LAs, whereas in non-Arab-settled *kuras* Coptic LAs remained in power. To account for the potential endogeneity of Arab settlement, we employ an instrumental variable strategy based on the geographic determinants of the spatial distribution of Arab tribes in Egypt in 700–969. We measure Copts' identity strength by a village-level[8] dummy variable =1 if the village is believed, according to pre-641 local Coptic legends (recorded in Coptic narratives), to have been visited by the Holy

[7] We do not have a panel dataset that traces *kuras* over time, though.

[8] There are 1,782 villages located within 42 *kuras*.

Family in its legendary biblical flight to Egypt. We also provide, as a robustness check, an alternative measure: a dummy variable =1 if a pre-641 Coptic saint or martyr is documented to have lived in the village according to another Coptic narrative.

We construct novel data, based on medieval Coptic narratives and papyrological tax records, in order to measure our three outcomes: (1) The proportion of converts is measured by a village-level dummy variable =1 if the village did not have any Coptic churches or monasteries circa 1200. We conduct a number of robustness checks to verify the validity of this measure. (2) The poll tax rate is measured by the individual-level annual poll tax payments in 641–1100, localized at the *kura*-level, based on papyrological poll tax registers and receipts that survived for only 4 out of 42 *kuras* (11% of villages in 1315 and 14% of the population in 1897). However, sample selection appears to be quasi-random.[9] The small number of clusters (*kuras*) does not permit an econometric analysis, though; instead, we rely on statistical comparisons of means. (3) While we do not have local-level data on tax revenue under the Arab Caliphate, we provide two indirect pieces of evidence: (i) We impute the poll tax revenue per capita circa 1200 at the *kura* level for the poll-tax sub-sample, by multiplying the proportion of villages with at least one church or monastery circa 1200 by the average poll tax payment. Imputed poll tax revenue, being at the *kura* level for only 4 *kuras*, does not allow a statistical test. (ii) Building on the tax revenue hysteresis predicted by the theory, we employ village-level Mamluk-period data on the total tax revenue per unit of land in 1375, post the Caliphate period, which permit an econometric analysis.

Our findings on the impact of the identity strength of the LAs come in support of the identity-based model. We find that villages located in Arab-settled *kuras* are less likely to have Coptic churches and monasteries circa 1200 (14% versus 22% in non-Arab-settled *kuras*). We also document that the average individual poll tax payment in 641–1100 is 27% higher, and that the imputed poll tax revenue per capita circa 1200 is halved, in Arab-settled *kuras*. Furthermore, we document that villages in Arab-settled *kuras* generate a 21% lower total tax revenue per unit of land in 1375 than villages in non-Arab-settled *kuras* that are located within the same region.[10]

Our evidence on the identity strength of Coptic taxpayers comes in support of both models. We provide the first evidence that local Coptic identity strength during the late antiquity period had a strong and long-lasting effect on conversions and taxation under the Arab Caliphate: The Holy-Family-visit areas are more likely to have churches and monasteries circa 1200. They also have higher individual poll tax payments, on average,

[9]The poll-tax sub-sample areas do not differ on most observables from out-of-the-sample areas, with the exception of (exogenous) geography: papyri were more likely to survive in hotter and drier areas.

[10]The negative impact of Arab settlement on tax revenue in 1375 is detected only when we include geographic region fixed effects, or province fixed effects according to the administrative division in 1375.

in 641–1100 (p-value = 0.11), higher imputed per-capita poll tax revenue circa 1200, and higher total tax revenue per unit of land in 1375. We obtain similar results when we use the saint-martyr measure. We conduct multiple robustness checks for all our findings regarding measurement, model specification, and estimation of standard errors.

Finally, we argue that the findings are not consistent with four alternative interpretations of the impact of Arab settlement: state extractive capacity, migration, persecutions, and persuasion. Importantly, the negative effect of Arab settlement on tax revenue lends support to the identity-based model, where Arab LAs willingly operate on the downward-sloping side of the Laffer curve; it also distinguishes it from the state capacity model, under which Arab LAs impose a higher tax rate and induce more conversions because of their lower collusion with Coptic taxpayers, and so generate a higher tax revenue.

Time-series analysis. Section 4 explores the time-series implications of the extraction and identity-based models. Rulers and agents are forward-looking, and conversions are permanent (apostasy assumption). We obtain two key theoretical results. First, under the extraction model, all conversions occur early on, and we hardly expect any conversions, poll tax hikes, or revenue busts, to occur thereafter. The identity-based model allows for later poll tax hikes, conversion waves, and revenue reductions; the equilibrium exhibits a sufficient-statistic property: In particular, date-t outcomes are determined by the highest ruler identity so far, a form of ratcheting. Second, a statistical implication of the identity-based model is that earlier rulers are more likely to order discriminatory tax hikes and induce conversion waves than later rulers, due to the ever more religious discriminatory tax base, and that the extraction model's relevance increases over time. We then show that under the binary coding of the ruler's identity strength (see our measures below), the identity-based model implies two testable hypotheses: the probability of tax hikes and conversions is (1) increasing in the current ruler's identity strength, and (2) decreasing in the number of previous strong identity rulers.

To test the two hypotheses, we exploit the time-series variation in the CA's identity strength across Egypt's rulers. Our main analysis is at the Caliph level ($N = 65$) spanning 530 years from 641 to 1170. We employ medieval Coptic narratives to measure the two outcomes: the occurrence of poll tax hikes and of conversion waves. The third outcome, poll tax revenue, is only observed at scattered points in time that do not permit an econometric analysis. We measure the Caliph's identity strength by a dummy variable =1 if the Caliph is not known for drinking alcohol, based on medieval Muslim narratives. We also employ, as a robustness check, a governor-level dataset ($N = 122$),[11]

[11] The two levels are identical when Egypt is not part of the Caliphate in 868–905 and 935–1170.

where we measure governor's identity strength by a dummy variable =1 if the governor is perceived as hostile toward non-converts, based on medieval Coptic narratives. While the Caliph's piety measure, being based on Muslim perceptions of Caliphs' personal religiosity, is plausibly exogenous allowing a causal interpretation, the governor's hostility measure may be subject to reverse causality as Coptic portrayals of governors are possibly shaped by their tax policy. We control for the occurrence of foreign attacks and of Nile shocks (both are inspired by theory), in addition to a linear time trend.

We start by plotting the long-term trends of the three outcomes. We document that poll tax hikes and conversion waves took place after "date 1" (defined as 641–661), but they became less frequent over time. The poll tax revenue fell rapidly early on, and then continued to decline at a slower rate. Taken together, these trends are consistent with the identity-based model. Next, we introduce our Caliph-level econometric evidence, which comes in support of the identity-based model. First, pious Caliphs are 29% more likely to trigger poll tax hikes, and 26% more likely to induce conversions. Second, the probability of poll tax hikes and conversion waves is decreasing in the number of previous pious Caliphs, only when we do not control for a linear time trend (which is systematically negative itself), because of the collinearity between the two variables over long time horizons. Third, when we include both Caliph's piety and the number of previous pious Caliphs in a horserace, Caliph's piety retains its magnitude and statistical significance. This suggests that even though earlier Caliphs left less fiscal leeway to the posterity, later Caliphs were still able to influence tax policy and induce conversions. We obtain similar findings when we use the governor-level measure of identity strength, and when we use the number of previous poll tax hikes instead of the number of previous pious Caliphs.

We then introduce an analytic narrative to investigate why the Caliphate removed the cap on the uniform tax only circa 750, and not before. We cannot study the timing of this reform econometrically, because it was a Caliphate-wide one-time policy change. This policy change is important on both historical and theoretical grounds, though. Historically, it created the canonical form of Islamic taxation, where both converts and non-converts are subject to the same land tax rate, but only differ on the poll tax. Theoretically, the removal of the cap on the uniform tax results from the threat of rebellion subsiding over time, even in an otherwise fully stationary environment. The intuition is that converts have less to gain from rebelling as they have already given up on their identity. This holds even though agents realize that by converting they lose their option value of having kept their identity in case of a successful rebellion in the future. While, as predicted by theory, the composition of rebels changed over time to include both converts and non-converts, the suppression of the tax revolts enabled the

Arab Caliphate to preserve the new tax system.

Extensions. Finally, Section 5 discusses two extensions: persecutions and emigration. First, the agency approach, being based on a potential conflict between the LAs and the CA, also rationalizes in well-defined circumstances, the use of inefficient, non-price instruments such as persecutions. Persecutions can be alternatively justified as a signaling device. These insights shed light on the relative role of persecutions (El-Leithy 2005) versus taxation (Frantz-Murphy 2004, Rapoport 2018) in inducing conversions, which has generated a debate among historians of early Islam. Persecutions were rare under the Caliphate which relied more on the fiscal tool, but became more common under the Bahri Mamluks (1250–1354). We think that signaling is a likely explanation of persecutions, as the latter took place under the less legitimate rulers who wanted to signal their piety. Second, we show that emigration is accommodated by our model. While irrelevant to non-convert Copts (who rarely emigrated from Egypt before 1950), emigration has been relevant in many other contexts, and we discuss some of these historical examples.

The rest of the paper is organized as follows. Section 1.2 reviews the related literature. Section 2 provides a historical background. The cross-sectional and time-series analyses are presented in Sections 3 and 4, respectively. We discuss the extensions in Section 5. Section 6 concludes. The paper has two appendices: *Supplemental Appendix* (Sections A and B), that appears on this journal's website, and *Webpage Appendix* (Sections C and D), that appears on the authors' personal webpages.

1.2 Related literature

The paper is related to a few strands of literature. It differs from the optimal taxation literature in at least two ways: the agency problem and the hysteresis effects associated with exit from the tax base. The paper shares with Becker (1957)'s theory of discrimination the feature that decision-makers have a possible distaste for minority membership. Acemoglu (2006)'s ruler taxes a constituency beyond the revenue-maximizing level so as to weaken the latter. The focus of the two papers is markedly different, as are the conclusions; for example, revolts are triggered by a soft tax treatment in Acemoglu and a tough one in our paper.[12] Taxes may also lie on the downward-sloping side of the Laf-

[12] In his model, the ruling elite not only aims at extracting rents from the output of an enterprising middle-class via a tax on its output, but also may try to achieve other goals with the tax, thus exceeding the maximal extraction tax rate. First, the elite may itself own firms and taxing the middle-class output discourages middle-class production and reduces the market wage. As Acemoglu emphasizes, this result hinges on limited tax instruments (for example, a tax on labor hired by the middle-class firms could take care of limiting competition for labor). By contrast, we study optimal taxation. Second, the middle class may rely on its financial power to rebel. That reason is complementary to our section on rebellion, which

fer curve for the taxation of externalities and internalities (e.g. tobacco or pollution). The interaction between Pigouvian taxation and revenue-generating distortive taxation is well known in public finance (Sandmo 1975, Bovenberg and de Mooij 1994, Bovenberg and Goulder 1996). Furthermore, our modeling, properly reinterpreted, also covers the design of "sin taxes" (O'Donoghue and Rabin 2006). For a hyperbolic consumer with present bias, consumption today brings immediate benefit and a delayed cost. The optimal tax may lie on the downward-sloping side of the Laffer curve. This literature however ignores agency in tax collection as well as issues related to the tax structure and to the specific dynamics of taxation and rebellion under ratcheting of compliance (apostasy or costly return); it thereby cannot guide the empirical analysis of this paper.

A large literature studies optimal taxation with non-utilitarian welfare functions (e.g. Fleurbaey and Maniquet 2011, Saez and Stantcheva 2016).[13] Much work has also been devoted to investigate the impact of altruism on optimal taxation (e.g. Diamond 2006, Farhi and Werning 2010, Kaplow 1995). These two literatures investigate neither the taxation of unwanted populations, nor its dynamic evolution as unwanted population members convert or leave the polity or organization. polity or organization. Finally, the time-series section shares with Tirole (2016) the focus on a positively selected tax base. Many theoretical results such as those on external threats and on time-decreasing resistance under multidimensional instruments are new, though.

The paper contributes to the economics of religion. One primary focus of this literature has been to identify the causal impact of religious beliefs on economic outcomes (Barro and McCleary 2003, Botticini and Eckstein 2005, Becker and Woessmann 2009). Instead of emphasizing the causality from religion to economics, our paper documents how economic incentives can alter the religious affiliation. In this respect, our paper contributes to a recent empirical literature that attempts to elicit the willingness to pay to maintain one's identity (or beliefs) (Augenblick et al. 2016, Delavande and Zafar 2018). Another line of this literature explores how political authorities co-opt religious ones to preserve legitimacy under the threat of rebellion (Greif and Tadelis 2010, Chaney 2013, Belloc et al. 2016, Rubin 2017, Cantoni et al. 2018). While less central to our explanation of discriminatory taxation and conversions, the strengthening of ruler's legitimacy via altering the religious composition of taxpayers is our preferred explana-

is based on manpower rather than money; as a consequence, the agents rebel when ill-treated by the ruler in this paper, while they rebel when well-treated and therefore empowered in Acemoglu's contribution. Overall, both the rationales for hurting and the focus differ between the two papers.

[13]The latter derive optimal taxation in an environment that is not necessarily welfarist (in particular, social welfare weights can depend on individual or aggregate characteristics which do not enter individuals' utilities). Their focus is on allowing various considerations, such as counterfactuals (what would have happened in the absence of taxes?), horizontal equity, libertarianism, equality of opportunity concerns, and poverty alleviation, to matter per se, independently of their consequences on taxpayers' utility.

tion of the removal of the cap on the uniform tax rate. Michalopoulos et al. (2018) show that the spread of Islam across countries today is correlated with pre-600 trade routes, lower land productivity, and higher land inequality, as Islam's redistributive institutions mitigated the incentives for predation in highly unequal areas. We focus instead on the role of the identity strengths of local authorities and population, while controlling for land productivity and inequality (trade routes do not vary within Egypt).[14] Finally, our paper extends Saleh (2018)'s empirical analysis of taxation and conversions in medieval Egypt, by endogenizing taxation from the viewpoint of the Caliphate.

2 Historical background

Islamization of Egypt. Arabs conquered Egypt in 641. On the eve of the Conquest, the vast majority of Egyptians were Coptic Christians, while non-Coptic Christians and Jews formed two small (urban) minorities. During the centuries that followed, non-Muslims shrank from 100% of the population in 641 to 16% in 1200 (Saleh 2018).

Egypt's Islamization was driven by Copts' conversions to Islam: Webpage Appendix Section D suggests that (1) the number of Arab settlers was small relative to Egypt's population, and non-convert Copts rarely emigrated from Egypt, (2) Muslims (both Arabs and converts) did not have higher fertility or lower mortality than non-converts, and (3) inter-marriages between Muslim males and Coptic females (which result by Islamic law in Muslim offspring) were rare. Being Muslim was an "absorbing state" owing to three Islamic laws on (1) the death penalty on apostates, (2) the offspring of a Muslim male being automatically Muslim, and (3) Muslim females' obligation to marry only Muslim males.[15]

Islamic tax system. Arabs taxed religion and land. Upon the Conquest, they imposed on Egyptians a discriminatory tax (τ) that was removed upon conversion to Islam, and a uniform tax (λ) on land that was levied regardless of religious affiliation. Up to circa 750, τ was made up of two components: (1) a poll tax, an annual per head cash tax on free adult males, and (2) the positive difference in land tax rate between the rate on Copts (*kharaj*) and the rate on Muslims (variously called *ushr*, *zakat*, *sadaqa*). λ

[14] Kuran (2012) traces the Middle East's relative stagnation to the increased hedging against state confiscation by investing in unproductive *waqfs* (tax-exempt religious endowments). Although this theory does not address the taxation and conversion question, we may under-estimate the discriminatory tax if state confiscation were an additional tax that targeted Copts (who were banned from forming *waqfs*). We think that this is unlikely for two reasons: (1) it targeted Coptic elites, unlike the poll tax which was levied on every Copt, and (2) confiscations are less relevant for the Caliphate than for the Mamluk period.

[15] Because Egyptians were mostly Copts in 641, and Egypt's Muslims are mostly descendants of Copts who converted to Islam, we use the terms "Copt" and "non-Muslim," and "convert" and "Muslim," interchangeably.

was thus equal to the *ushr*. Importantly, the *ushr* was capped at 10% of land's yield according to *hadith* (prophet's sayings).[16]

Circa 750 (date uncertain), Caliphs, backed by jurists, raised the land tax on Muslims from the *ushr* to the *kharaj* rate. They further removed any preexisting country-specific treaty-based upper bound on *kharaj*, by denying the historical existence of peace treaties in most of the conquered territories, including Egypt. Consequently, from that date on, τ equated the poll tax, and λ equated the *kharaj* land tax. As a result, the land tax increased sharply circa 750. It then fluctuated over time in response to economic shocks, but never went back to its pre-750 level. This change marks the establishment of the *canonical* Islamic tax system that remained in place, until the abolition of the poll tax in 1856.[17]

Tax administration. Caliphs delegated tax collection to Egypt's (fiscal) governors, who decided on the total budget that was used to pay the tribute to the Caliphate,[18] and to finance the salaries of Egypt's top officials, the army, the police, the judiciary, and the bureaucracy. Importantly, poll and land taxes were *not* raised to finance local public goods, which were financed instead by ad hoc levies and corvée labor.

Governors/Caliphs delegated tax administration to the local authorities of each *kura*. Transfer demands were issued to each *kura*. Poll and land tax rates varied locally, because local authorities had discretionary power on tax rates, or at least, on the level of tax enforcement.

3 Cross-sectional analysis

3.1 Theory

Model. There is a continuum of districts, indexed by $i \in [0,1]$, each with a mass 1 of *agents*. Agents are initially endowed with the same identity (say, the Coptic religion). There is a second identity (say, Islam) that the agents can embrace, abandoning the initial one.

[16]Due to the lack of papyrological evidence on the *ushr* tax before 750, it has been argued that Muslim landholders actually paid no land tax before 750.

[17]We abstract from other discriminatory taxes in our definition of τ: (1) special taxes on non-Muslims (up to 750) levied for specific uses (e.g., military expenses, lodging for officials, governor's expenses, local public projects), because they were irregular ad hoc levies, (2) military conscription on Muslims (up to 833), because it was in return for a state stipend, and it was *not* widespread in Egypt, (3) (non-state) taxes/subsidies administered by religious organizations (churches, monasteries, mosques), because we do not have evidence on their magnitudes, and they were not enforced by the state. Similarly, we abstract from the expansion in λ tax base after 858 to include pasture, weir, and various crops and products.

[18]When Egypt was a province of the Caliphate in 641–868 and 905–935. Caliphs generally appointed two governors: military and fiscal; the two positions could have been held by the same person. In 868–905 and 935–1170, governors and Caliphs were identical, because Egypt was independent then.

Each district i is run by a *local authority* (LA), which optimally collects discriminatory and uniform taxes $\{\tau_i, \lambda_i\}$. The discriminatory tax rate $\tau_i \geq 0$ is levied solely on agents who choose to maintain their identity. The associated tax revenue is thus equal to τ_i times the fraction of such agents. In contrast, the uniform tax λ_i will for notational simplicity denote the tax revenue or effective tax rate paid by all agents. Our preferred interpretation of λ_i is the extractive one: λ_i stands for the maximum revenue that can be obtained through a uniform tax (say, a tax on land); but we are agnostic about the determination of the district-i uniform tax λ_i (indeed, in the case of Egypt, Islamic law capped the uniform/land tax until 750) and therefore keep its formulation as general as possible.

An individual agent is characterized by a parameter $\theta \in \mathbb{R}$, measuring his willingness to pay for keeping his initial identity (his "identity strength"). Hence, a type-θ agent in district i keeps his identity if and only if $\theta \geq \tau_i$. Willingnesses to pay in district i are distributed according to cdf $F(\theta - r_i)$. Thus, identity is more pregnant in a district with a higher r_i . We assume that the distribution F is smooth, has density f, and satisfies the monotone hazard rate property: $f(\theta)/[1-F(\theta)]$ is strictly increasing.[19]

The departure from the extraction model is that taxes embody an identity-related motive: The district-i LA incurs (dis)utility $c_i \in \mathbb{R}$ times the fraction of agents preserving their identity; depending on the district, the cost c_i can be positive or negative. The parameters r_i and c_i thus measure the agents' and the LA's *identity strengths* in district i. There is a smooth joint distribution $G(c_i, r_i)$ on district characteristics.

A *central authority (CA)* has an indirect extraction motive: It cares about transfers from local authorities. And, like the local authorities, it also has a nonfinancial objective: It incurs disutility $c > 0$ per agent preserving his identity. The CA relies on local authorities to collect taxes. It can only demand a district-specific transfer T_i; that is, it does not have any local presence that would allow it to interfere with local tax collection, and the LA is residual claimant for its revenue. The LA therefore has a direct extraction motive. For simplicity, there is no asymmetry of information between the CA and the LA regarding $\{r_i, c_i\}$.

Objective functions. An *agent* of type θ living in district i has objective function:

$$U_i(\theta) = -\lambda_i - \min\{\tau_i, \theta\}.$$

Because *LA* i has a (dis)taste c_i (or taste if c_i <0) per agent preserving his identity

[19] This assumption will in particular guarantee that objective functions are strictly quasi-concave. The parameter θ should be thought of as the net WTP for keeping identity. In our application, embracing Islam may create an option value for the convert (or his lineage) from possibly adhering to the new, not-yet-experienced religion. This option value is to be subtracted from the gross benefit from remaining Copt. In particular, the support of F may include negative values of θ, i.e. agents who would convert even in the absence of discriminatory tax.

and is residual claimant for its revenue once it has transferred T_i, its utility is:

$$V_i = [\lambda_i + R_i(\tau_i) - T_i] - c_i[1 - F(\tau_i - r_i)] = \lambda_i + R_i^a(\tau_i) - T_i$$

where $R_i(\tau_i) \equiv \tau_i[1 - F(\tau_i - r_i)]$ denote the discriminatory tax revenue, and $R_i^a(\tau_i) \equiv (\tau_i - c_i)[1 - F(\tau_i - r_i)]$ is an "adjusted tax revenue" that accounts for the LA's (dis)taste for the agents' keeping their identity.

The *CA* receives taxes in amount $\int_0^1 T_i di$, and has welfare $W \equiv \int_0^1 [T_i - c[1 - F(\tau_i - r_i)]]di$. We similarly define the CA's adjusted tax revenue in district i as $R_i^c(\tau_i) \equiv (\tau_i - c)[1 - F(\tau_i - r_i)]$. If each district's collected tax is equal to the transferred revenue (which will be a non-trivial implication of symmetric information[20]), this expression can alternatively be rewritten as:

$$W = \int_0^1 [\lambda_i + R_i^c(\tau_i)]di.$$

Equilibrium tax and revenue. Faced with transfer demand T_i, the LA in district i solves $\max_{\{\tau_i\}} R_i^a(\tau_i) - T_i$ subject to the revenue-collection constraint: $\lambda_i + R_i(\tau_i) \geq T_i$. We define three tax rates that will play a central role in what follows. $\tau_i^m \equiv \arg\max\{R_i(\tau_i)\}$ denotes the extraction-maximizing (or monopoly) discriminatory tax in district i, yielding the maximally extractive revenue R_i^m. Next, $\tau_i^a(c_i) \equiv \arg\max\{R_i^a(\tau_i)\}$ denotes the preferred discriminatory tax of an LA in district i with identity strength c_i. Finally, $\tau_i^c \equiv \tau_i^a(c)$ denotes the CA's preferred discriminatory tax. Trivially, $\tau_i^a(c_i) > \tau_i^m$ if and only if $c_i > 0$. In particular, the CA's preferred tax τ_i^c is on the decreasing-revenue side of the Laffer curve.

Proposition 1 ***(equilibrium discriminatory tax and revenue)***.

For any given district i:

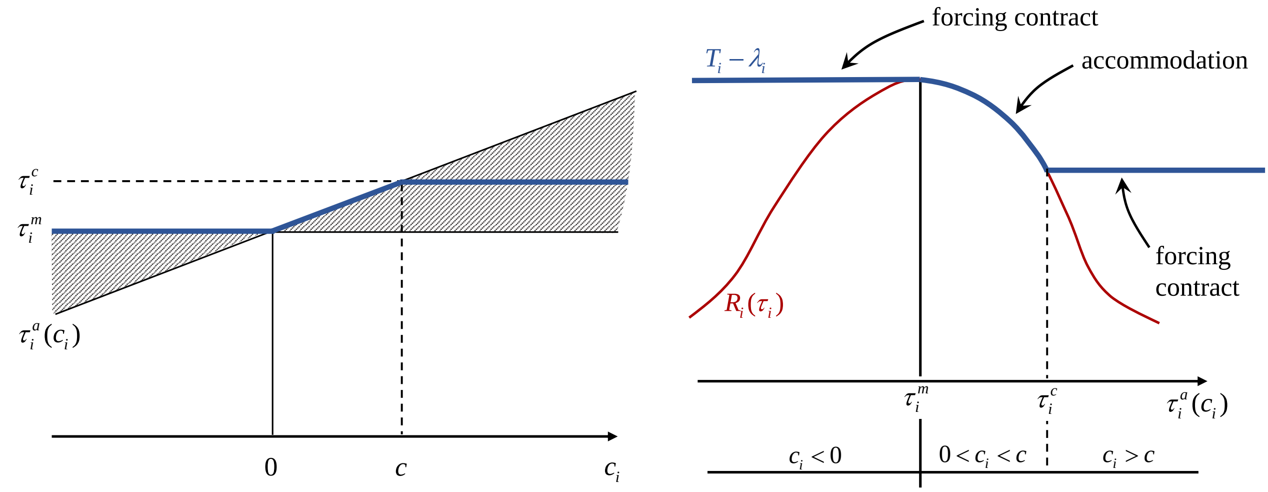
(i) Regardless of the sign of c_i, the district-specific equilibrium discriminatory tax satisfies $\tau_i \geq \tau_i^m$ (that is, it is weakly on the decreasing-revenue side of the district's Laffer curve), is almost everywhere differentiable, and where so, satisfies $\frac{\partial \tau_i}{\partial r_i} \in (0,1)$ and $\frac{\partial \tau_i}{\partial c_i} \in [0,1)$. The conversion rate, $F(\tau_i - r_i)$, is weakly increasing in the LA's strength of identity, c_i, and decreasing in the agents' identity strength r_i.

(ii) District i's transfer T_i is equal to its revenue, $\lambda_i + R_i(\tau_i)$. It is invariant to c_i and is equal to $\lambda_i + R_i^m$ for $c_i < 0$, decreases with c_i in $(0,c)$, and is invariant to c_i for $c_i > c$ (see Figure 1(b)).[21] *It is strictly increasing in the agents' identity strength r_i.*

(iii) There is no delegation cost if and only if $c_i \geq c$.

[20]Under moral hazard, symmetric information may not preclude the agent from enjoying rents, as in the efficiency wage model.

[21]A corollary is that an increase in the CA's identity strength, c, leads to more conversions and a lower total revenue.



(a) Implementable (hatched area) and optimal delegated tax (thick line)

(b) Revenue and local authority identity strength

Figure 1: **Delegated tax collection and Laffer curve**

The intuition behind Proposition 1 can be grasped from Figure 1(a) (the formal proof is in Supplemental Appendix Section A.1). Let us separate local authorities' types into three categories: *zealous* when $c_i > c$, *soft* when $c_i \in (0, c)$, and *counterattitudinal* when $c_i < 0$. From CA's viewpoint, a zealous authority puts too much weight on inducing agents to surrender their identity and too little on revenue; it is easy to control them, as a revenue requirement at the level of the revenue the CA would raise itself forces the local authority both to lower its discriminatory tax and to raise more revenue; there is no agency/delegation cost.

In contrast, the other local authorities are not fervent enough. The CA would like them to raise their discriminatory tax, but faces an agency problem. In the case of a soft LA, an increase in its discriminatory tax leads to less revenue; however, a lower transfer requirement does not induce it to increase its tax; rather, it chooses to pocket the difference between revenue and transfer. Technically, the set of implementable discriminatory taxes is $[\tau_i^m, \tau_i^a(c_i)]$. In sum, the outcome for soft local authorities is dictated by the latter's preferences. In the case of counterattitudinal local authorities, the implementable set is symmetrically $[\tau_i^a(c_i), \tau_i^m]$. The CA prefers the highest tax that can be implemented, the extractive tax, to any below it, because it brings both more revenue and more agents to surrender their identity.

The comparative statics with respect to agent religiosity are straightforward. District i's demand for maintaining identity is $1 - F(\tau_i - r_i)$. Under the monotone-hazard-rate condition, the demand elasticity, $\tau_i f(\tau_i - r_i)/[1 - F(\tau_i - r_i)]$, is decreasing and so the discriminatory tax is increasing in identity strength r_i, while overall conversions, $F(\tau_i - r_i)$, decrease. Those properties are not affected by the agency problem.[22]

Extraction model. We compare outcomes in the identity-based and the extraction-only models. The latter corresponds to $c_i = 0$ for all i.[23]

Corollary 1 ***(horserace with extraction model).***
(i) Both the extraction model and the identity-based one predict that the discriminatory tax, the revenue and the transfer all grow with the district's agent identity strength r_i, and that the conversion rate decreases with r_i.
(ii) In contrast with the identity-based model, the discriminatory tax and revenue and the overall transfer in the extraction model do not vary with either local or central authority's strength of identity.

3.2 Empirics

Overview. To test the identity-based and extraction models apart, we exploit the local variation in early Islamic Egypt in the identity strength of both the LAs and the Coptic taxpayers. We define the CA as Egypt's governors or Caliphs, and the LAs as *kura* headmen and the local bureaucracies.[24] Section 3.2.1 first explains how we measure our two main regressors: While we measure the LA's identity strength, c_i, by the incidence of Arab settlement in 700–969 in *kura* i, we measure local Coptic religiosity, r_{ji}, at the village level j located within *kura* i by two measures that are based on local Coptic beliefs during the late antiquity period.

Sections 3.2.2–3.2.4 introduce the analysis of the three outcomes: (1) the proportion of converts, F_{ji}, proxied by the non-existence of Coptic churches and monasteries at the village level j circa 1200; (2) the poll tax rate, τ_{hi}, measured by the annual poll tax payments at the individual taxpayer level h in the Egyptian papyri in 641–1100

[22] The threat of rebellion will be studied in much detail in the time-series analysis, where it is most interesting. Webpage Appendix Section C however builds the possibility of rebellion into the cross-sectional analysis. A rebellion, if successful, kicks out the rulers and results in (say) the absence of taxes. It requires cooperation among the various districts, as each district is too small to challenge the CA's rule. The CA can change the agents' incentive to rebel by reducing the revenue demands it imposes on local authorities. For example, and referring to Figure 1(b), it can moderate its revenue demands in districts run by counterattitudinal tax collectors. Webpage Appendix Section C shows that under the threat of rebellion, the discriminatory tax τ_i is still increasing in the LA's identity strength c_i. The discriminatory tax revenue (R_i) is now inverted-U shaped in the LA's identity strength (c_i), with a peak for a secular LA ($c_i = 0$).

[23] When $c_i = 0$ for all i, the CA's identity strength is irrelevant, as districts maximally extract regardless of the revenue request.

[24] Our focus in the cross-sectional analysis is on the agency problem between the Caliphs/governors and the LAs. The Caliph-governor agency will be relevant in interpreting the time-series evidence.

(localized at the *kura* and not the village level); (3) the tax revenue, measured by the imputed poll tax revenue, R_i, at the *kura* level circa 1200, and by the observed total tax revenue per unit of land, T_{ji}, at the village level in 1375.

Throughout the analysis, we test whether our findings are consistent with either model. Recall that both models imply that local Coptic religiosity reduces the probability of conversion to Islam, and increases the poll tax rate and revenue. The two models differ though on the effect of the LAs' identity strength. While the extraction-only model predicts that the three outcomes do not vary with the identity strength of LAs, the identity-based model implies that strong identity LAs will levy a higher tax rate, induce more conversions, and raise a lower revenue; they deliberately operate on the downward-sloping side of the Laffer curve. Finally, Section 3.2.5 discusses alternative interpretations of our findings, including most importantly, the state capacity model.

3.2.1 Measuring the identity strength of LAs and Coptic taxpayers

Identity strength of LAs (c_i). We measure c_i by a dummy variable =1 if at least one Arab tribe settled permanently in *kura i* between 700 and 969. Before the Arab Conquest in 641, LAs were all Coptic Christians. While Arabs initially kept the status quo, they soon attempted to Arabize the LAs. Arabization was constrained, though, by the number of Arabs with enough highly specialized human capital to replace, not only *kura* headmen, but the entire local bureaucracies.[25] This is supported by administrative evidence. Individual-level (non-localized) data on occupational titles and religious affiliation, constructed from the Egyptian papyri dating between 641 and 969, reveal that LAs were partially Arabized during this period: Muslims (Arabs and converts) came to occupy jobs at all bureaucracy levels, and to be over-represented among high bureaucrats.[26] Non-convert Copts kept being over-represented among mid-low bureaucrats (tax collectors, scribes, land surveyors), though (Saleh 2018). We argue that Arab settlement measures the Arabization of LAs. Between 641 and 969, Arab tribes settled in certain *kuras* but not others, first temporarily during the spring season (in 641–700), and then permanently (in 700–969). In *kuras* where Arabs settled, they were more likely to replace Copts in the LA. Consequently, these *kuras* faced a larger share of Arabs in the LA, compared to *kuras* where Arabs did not settle.

We think of Arab-settled *kuras* as characterized by $c_i > 0$, and of non-Arab-settled

[25]Coptic mid-low bureaucrats were difficult to replace, because they had highly specialized human capital in land surveying, measuring the Nile level, and agriculture in rural Egypt. Many Copts remained as *kura* headmen, too. Basilios, the Coptic head of *Aphrodito* circa 710, is a well-known example.

[26]While Muslim high-level bureaucrats in this dataset are almost certainly Arabs, we cannot separate Arabs from converts at mid-low levels of the bureaucracy, because converts had to adopt an Arabic name and became clients of Arab tribes. This pooling will over-estimate the share of Arabs in mid-low LAs.

kuras as characterized by $c_i < 0$, with *counterattitudinal* Coptic administrators. While c_i is continuous in theory, our empirical measure is dichotomous, as we do not have a measure of c_i among Coptic and Arab administrators. This is not a concern for the theoretical predictions, which are all monotonic in c_i.

Panel (A) of Supplemental Appendix Figure B.1 shows the location of *kuras* where Arab tribes settled in 700–969. Arabs were more likely to settle in the eastern and western Delta than in the central Delta, and in the northern Nile Valley than in the southern Valley.

Identity strength of Copts (r_{ji}). We measure r_{ji} at the village level by a dummy variable =1 if it is believed, according to pre-641 local Coptic legends, that a village j, located within *kura i*, had been visited by the Holy Family (henceforth, HF) during its legendary biblical flight to Egypt. The HF visit legend has been an important element of popular Coptic Christianity until today. Villages on this list are mentioned in an apocryphal book that is attributed to Theophilus, the patriarch of Alexandria in 384–412. We think of the HF-visit villages as having higher r_{ji}, on average. These villages contained "miraculous" sites that Jesus and/or Mary were believed (among locals) to have created or touched, such as hand-prints, footprints, trees, and wells, which might have instilled a strong sense of Coptic identity among the local population before 641.[27] Panel (B) of Supplemental Appendix Figure B.1 shows the location of villages for which r_{ji} =1, which is the case for 24 villages (1.3%) of 1,782 villages in 1315.

To address the limitations of the HF measure, we employ, as a robustness check, an alternative measure of r_{ji}: a dummy variable =1 if a Coptic saint or martyr spent (part of) their lives in village j between 49 CE, the customary date of establishing the Coptic Church of Alexandria by Saint Mark, and 641 CE. We constructed this measure from the Coptic Synaxarium, the major medieval liturgical Coptic book that compiles biographies of saints and martyrs arranged according to days of the Coptic calendar year. Like the HF measure, we think of the saint-martyr villages (30 villages, 1.7% of villages in 1315) as having higher r_{ji}, on average. According to local Coptic beliefs, these local saints and martyrs performed miracles, and were (mostly) tortured to death by either Roman (pagan) or Byzantine (non-Coptic Christian) governors, in defense of their Coptic Christian faith. This may have instilled a stronger Coptic identity in these villages. The two measures are weakly positively correlated ($\rho = 0.14$).[28]

[27] Although the book's authorship and date are both doubtful, with some scholars attributing the book to an unknown author in the post-641 period, many of the local legends that the book's author compiled likely date from the pre-641 period, as the HF visit was mentioned in Coptic sources during the Roman and Byzantine eras.

[28] Among 1,782 villages, there are 1,730 villages for which both measures are equal to 0, and 4 villages for which they are both equal to 1.

3.2.2 Impact on conversions (F_{ji})

Measuring the proportion of converts (F_{ji}). We employ a dummy variable =1 if village j in *kura* i did not have any operating Coptic church or monastery circa 1200, based on Abul-Makarim's medieval Coptic chronicle. We argue that this measure captures the proportion of converts between 641, when all villages were 100% Copt, and 1200, the end of the Arab Caliphate period. The presence of an operating Coptic church or monastery is an indicator of the local presence of a sizable non-convert population. Supplemental Appendix Section B.2 discusses the assumptions under which our measure is valid, and the robustness checks that we conducted in order to verify its validity.

Panel (C) of Supplemental Appendix Figure B.1 shows the spatial distribution of villages that did not have any Coptic church or monastery in 1200. According to this measure, converts were already in the majority by 1200: 84% of Egyptian villages did not have any Coptic church or monastery by then. But there was considerable spatial heterogeneity; for example, Coptic churches and monasteries were more likely to survive in the central Delta and the middle and southern Valley.

Empirical specification. Our objective is to examine whether the effect of the LAs' identity strength on the proportion of converts is consistent with the extraction model or with the identity-based model. Recall that the implications of both models are the same with respect to Coptic identity strength. We first estimate the following OLS model:

$$F_{ji,1200} = \beta_r + \beta_1 c_{i,700\text{–}969} + \beta_2 r_{ji} + X_i \beta_3 + M_{ji} \beta_4 + \varepsilon_{ji} \qquad (1)$$

where $F_{ji,1200}$ =1 if there is not any Coptic church or monastery in village j, located within *kura* i, circa 1200. The main regressor is $c_{i,700\text{–}969}$ =1 if an Arab tribe settled in *kura* i in 700–969. The second regressor is r_{ji} =1 if village j is believed, according to pre-641 local Coptic legends, to have been visited by the Holy Family. To account for regional unobserved heterogeneity in baseline characteristics that may affect both Arab settlement and conversions, we control for a full set of region fixed effects (β_r), where we divide Egypt into four roughly equal regions by latitude range: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley.[29]

We include two sets of control variables. The vector X_i includes Byzantine-period *kura*-level controls: (1) the logarithm of urban population of *kura* i circa 300; using urbanization as a proxy for income is standard in history, as urban populations were richer on average, and were thus less likely to convert (Saleh 2018),[30] and (2) a dummy

[29]The latitude ranges are: (1) Delta $\in [30.11, 31.55]$, (2) northern Valley $\in [28.73, 30.33]$, (3) middle Valley $\in [26.40, 28.70]$, (4) southern Valley $\in [24.58, 26.20]$.

[30]Urban population is defined as the simple sum of the population of Greek cities (metropolis) and of *nome* capitals (Egypt's administrative units during the Roman period).

variable =1 if there was a Byzantine garrison in *kura* i circa 600, which captures military resistance to the Arab Conquest. The vector M_{ji} includes geographic village-level controls: (3) FAO-GAEZ cereals suitability index, which is the maximum value of the suitability indices of barley, wheat, beans, and maize, under irrigation and intermediate input level,[31] (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Standard errors are clustered at the *kura* level, the level of aggregation of Arab settlement. As a robustness check, we estimate a spatial-autoregressive model with spatial-autoregressive standard errors (SARAR). The results of the robustness checks for conversions are summarized in Supplemental Appendix Table B.3.[32]

We estimate two alternative specifications as robustness checks. First, one concern that arises is the potential correlation between c_i and r_{ji}: Arab tribes may have chosen where to settle based on the identity strength of Copts. We thus estimate an alternative specification in which we interact Arab settlement with the HF visit status. We find that the main variables retain their magnitudes while the interaction term itself is statistically insignificant. Second, to account for the impact of land inequality on conversions, suggested by Michalopoulos et al. (2018), we include a *kura*-level dummy variable =1 if there is at least one *autopract* estate circa 600; the *autopragia* was a privilege granted to large landholders in late Byzantine Egypt allowing them to pay taxes directly to the capital city and to collect taxes in their constituencies. It can be thus used to measure land concentration in each *kura*. We fail to detect an effect of the *autopragia* status.[33]

The null hypothesis (H_0) on β_1 is the extraction model, which implies that $\frac{\partial F_{ji}}{\partial c_i} = 0$. The alternative hypothesis (H_1) is the identity-based model, which implies that $\frac{\partial F_{ji}}{\partial c_i} > 0$. H_0 on β_2 is that it is equal to 0. H_1 is that $\frac{\partial F_{ji}}{\partial r_{ji}} < 0$, which is consistent with both models.

Instrumental Variable (IV) strategy. The identification assumption in equation (1) is that the cross-*kura* variation in Arab settlement is exogenous to baseline characteristics of *kuras*, which may be driving conversions. This assumption may be violated due to (1) reverse causality: Arab settlers may have settled in *kuras* with larger convert populations, and (2) omitted variables: Arab settlement choice may have hinged on other unobservable pre-641 characteristics of *kuras* that also account for the variation in conversions. To deal with the potential endogeneity of Arab settlement, we employ

[31] FAO-GAEZ does not provide crop suitability measures under irrigation and *low* input level.

[32] Webpage Appendix Section D shows the full results of all the robustness checks in the paper.

[33] We do not include this variable in the basic set of controls because it is missing for half of the *kuras*. Also, because of this reason, we only estimate this specification without region FE.

an IV strategy, where we estimate the following first-stage regression:

$$c_{i,700\text{–}969} = \alpha_r + \alpha_1 DistancetoArish_i + \alpha_2 BorderDesert_i + \alpha_3 (DistancetoArish_i \times BorderDesert_i) + X_i\alpha_4 + M_i\alpha_5 + v_i \quad (2)$$

where $DistancetoArish_i$ is the distance between the capital of *kura i* and *Arish*, a small town close to Egypt's northeastern border, that was the first to be captured by Arabs in 639 due to its proximity to the Arab peninsula (Conquest was by land from the northeast), $BorderDesert_i$ =1 if *kura i* borders desert land. All *kuras* border hinterland, except these in central Delta.

We argue that the IVs are relevant. Column (1) of Supplemental Appendix Table B.1 suggests that Arabs were more likely to settle in *kuras* that are both closer to *Arish* and bordering desert. This is confirmed by historical evidence. For one, proximity to *Arish* determined the extent to which Arabs were willing to travel, although there were exceptions.[34] For another, Arabs preferred *kuras* that bordered desert, where they first settled temporarily during the spring season in 641–700, to practice hunting and horse riding in a similar environment to that of the Arab peninsula. Starting from circa 700, they settled in these *kuras* permanently.[35]

Furthermore, we argue that the IVs are exogenous, as they are determined by geography. They arguably satisfy the exclusion restriction, conditional on controls. Columns (2)-(9) of Supplemental Appendix Table B.1 reveal that the IVs are not correlated with most Byzantine-period and geographic characteristics, with the exception of urban population circa 300 and temperature.

Findings. The findings are in Table 1. We relegate the results without the region fixed effects (which are almost identical) to Supplemental Appendix Table B.2a for conciseness. Column (1) reveals that the probability of conversion to Islam in 641–1200 is higher in Arab-settled *kuras*: whereas 22% of villages located in *kuras* where Coptic LAs remained in power ($c_i < 0$) had Coptic churches or monasteries in 1200, the proportion is only 14% in *kuras* where Arab tribes settled in 700–969 ($c_i > 0$). Because all *kuras* were 100% Copt before 641, this finding suggests that *kuras* where Arabs settled witnessed relatively more conversions to Islam in 641–1200. Column (2) shows that HF-visit (higher r_{ji}) villages were more likely to have Coptic churches or monasteries in 1200 (75%, compared to only 15% in non-HF-visit villages). Including both regressors and control variables in columns (3)-(5) yields similar results to those in columns (1) and (2). The IV results in columns (6)-(9) indicate that the coefficients of

[34] Regardless of the distance to *'Arish*, Arabs were more likely to settle closer to frontier cities such as *Aswan* in the south and Alexandria in the north. Also, Arabs were more likely to settle in western Delta than in central Delta (that is actually closer to *'Arish*), due to the former's proximity to desert.

[35] Arabs were 58% more likely to settle permanently after 700 in *kuras* where they settled temporarily before 700.

Arab settlement and of the HF visit retain their magnitudes and statistical significance. We obtain similar results when we use the saint-martyr measure as a proxy of Copt religiosity (Supplemental Appendix Table B.3).

To conclude, the positive effect of Arab settlement on conversions is consistent with the identity-based model, but not with the extraction model. The negative effect of the HF visit is consistent with both models.

Table 1: **Local determinants of conversions to Islam in 641–1200**

Dependent variable: =1 if no Coptic church or monastery in village j circa 1200

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.08		0.08	0.08	0.07	0.13	0.12	0.13	0.09
	(0.04)**		(0.04)**	(0.03)**	(0.04)*	(0.07)*	(0.07)*	(0.06)**	(0.05)*
=1 if HF visit (r_{ji})		-0.59	-0.59	-0.59	-0.62		-0.59	-0.59	-0.62
		(0.08)***	(0.08)***	(0.08)***	(0.09)***		(0.08)***	(0.08)***	(0.09)***
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1782	1782	1782	1782	1751	1782	1782	1782	1751
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.01	0.03	0.04	0.04	0.06	0.01	0.04	0.04	0.06
Mean dep. var. in control	0.78	0.85	0.78	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat						16.94	16.93	15.43	19.32

Notes: Standard errors clustered at the *kura* level are in parentheses. Regions are: (1) Delta, (2) northern Valley, (3) middle Valley, (4) southern Valley. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura* i circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura* i circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Sources: See Supplemental Appendix Section B.1.

3.2.3 Impact on poll tax rate (τ_{hi})

Measuring τ_{hi}. We employ Egypt's papyrological tax registers and receipts in 641–1100 to measure the annual poll tax payment (τ_{hi}) made by taxpayer h, located in *kura* i, in 641–1100 (N= 408 individual taxpayers).[36] Poll tax records survived in only 4 out of 42 *kuras*, all located in the Nile Valley: *Hermopolis* (N= 77), *Aphrodito* (N= 314),

[36]Webpage Appendix D shows example pictures of the secondary sources that we used to construct our dataset. Initially, the Arabic term (*jizya*) meant "tax in cash" that included both the poll tax and the cash land tax. The term was later confined to the poll tax during the 8th century. Poll tax payments are clearly identified in our tax papyri sample, though.

Fayum (*N*= 7), and *Ihnas* (*N*= 10) (Supplemental Appendix Figure B.1, panel (D)).[37] Furthermore, 95% of the sample comes from *Hermopolis* and *Aphrodito*. These four *kuras* include 11% of the total number of villages in the 1315 cadastre, and 14% of the population in the 1897 census.[38]

The poll tax sub-sample is subject to two caveats. First, both the small number of clusters (*kuras*), and their concentration in the Nile Valley, raise a natural concern about the national representativeness of the sample. Two remarks bolster our confidence in the sample, though: (1) Sample selection appears to be quasi-random. Supplemental Appendix Table B.4 reveals that villages/*kuras* in the poll-tax sub-sample are not statistically different with respect to most observables, in comparison to out-of-sample villages/*kuras*. The main exception is (exogenous) geographic characteristics: villages in the poll-tax sub-sample have higher temperature, greater temperature range, less rainfall, higher slope, and higher likelihood of bordering desert, than out-of-sample villages. This confirms a long-known fact in Greco-Roman, Coptic, and Arabic papyrology: papyri are more likely to survive in the Nile Valley due to its dry and hot climate.[39] (2) We re-estimate equation (1) for villages in the poll-tax sub-sample, and the results are qualitatively similar to those for the full sample (Supplemental Appendix Table B.4).

The second caveat about the poll-tax sub-sample is that the papyri are typically dated within a range (century or longer).[40] We thus pool the papyri from all four locations and date them between 641 and 1100. This raises a concern that we may be confounding the cross-sectional effect of c_i on τ_{hi} with its time-series effect. The latter effect may arise due to *kura*-specific changes in c_i, or Egypt-level changes in the CA's identity strength, over time. In the absence of panel data on c_i and τ_i over time, we cannot rule out this concern. However, to mitigate the concern that the CA and LAs may have changed under the Fatimids (after 969), we note that our finding that the poll tax payment is higher, on average, in Arab-settled *kuras* holds if we limit our sample to the pre-Fatimid (pre-969) period (the difference is not statistically significant, though).

Evidence. We do not estimate a regression model for τ_{hi}, because of the small number of clusters (*kuras*) in the poll-tax sub-sample. However, we provide suggestive evidence by examining the difference in mean τ_{hi} (1) between Arab-settled and non-Arab-settled

[37] We exclude 143 individual poll tax records with missing location.

[38] Webpage Appendix D shows the histogram of τ_{hi} in each *kura*.

[39] Random events further uncovered papyri in specific locations within the Valley. *Aphrodito*'s tax papyri were discovered in 1901 by local farmers while digging a well. The papyri were then distributed among farmers, and the remaining documents ended up in the British Museum (Bell 1910).

[40] *Hermopolis*'s sample is from 731–1100, *Aphrodito* 703–733, *Fayum* 641–1005, and *Ihnas* 701–900.

kuras in 700–969, and (2) between HF-visit and non-HF-visit *kuras*.[41] The downside is that we are not able to control for potentially confounding variables. The standard error of the difference in means is clustered at the *kura* level.[42]

The null hypothesis on the difference in τ_{hi} between Arab-settled and non-Arab-settled *kuras* is the extraction model, which implies that it is equal to 0. The alternative hypothesis is the identity-based model, which implies that it is greater than 0. The null hypothesis on the difference in average poll tax payment between HF-visit and non-HF-visit *kuras* is that it is equal to 0, whereas the alternative hypothesis is that it is greater than 0, which is consistent with both models.

The findings are shown in Table 2. Taxpayers in the *kuras* of *Hermopolis*, *Fayum*, and *Ihnas*, where Arabs settled in 700–969, paid on average a higher poll tax rate in 641–1100 by 0.29 dinar, 27% more than the average poll tax in *Aphrodito*, where Arabs did not settle and the LA thus remained Coptic. The difference is statistically significant, and the magnitude is economically sizable: It amounts to 3% of the annual wage of manual low-skilled workers in 661–969, who constituted the low-income poll tax bracket, and to 29% of the *de jure* annual poll tax on this bracket (=1 dinar). This result is consistent with the identity-based model, and not the extraction model. Furthermore, taxpayers in the two *kuras* that are believed to have been visited by the Holy Family, *Hermopolis* and *Ihnas*, and thus had higher r_i, paid a higher poll tax (as implied by both models), yet the difference is not statistically significant.[43] When we use the saint-martyr measure, we obtain similar results: taxpayers in the saint-martyr *kuras* ($N = 324$) paid a higher poll tax rate than in non-saint-martyr *kuras* ($N = 84$): 1.36 ($sd = 1.25$) versus 1.08 ($sd = 1.15$) [p-value = 0.16].

3.2.4 Impact on poll and total tax revenue (R_i and T_{ji})

The null hypothesis (H_0) on the effect of LAs' identity strength on the poll tax revenue (R_i) and on the total tax revenue (T_{ji}) is that it is equal to 0, as implied by the extraction model. The alternative hypothesis (H_1), implied by the identity-based model, is that the poll and total tax revenues are both decreasing in LAs' identity strength. H_0 on the effect of Copt identity strength is that it is equal to 0, whereas H_1 is that it is

[41] Because τ_{hi} is only localized at the *kura* level, we aggregate r_{ji} to the *kura* level; we define r_i as a dummy variable =1 if $\bar{r}_i = \sum r_{ji}/n_i > \text{median}(\bar{r}_i)$, where n_i is the number of villages in *kura i*.

[42] As the small number of clusters may bias the standard errors downwards (Cameron et al. 2008), we estimate the p-value using the Wild Cluster Restricted (WCR) bootstrap (Roodman et al. 2018).

[43] We also estimate Lee's bounds of the effects of c_i and r_i under non-random selection of the poll-tax sub-sample. To do this, we first aggregate τ_{hi} to the *kura* level as $\bar{\tau}_i = \sum \tau_{hi}/n_i$, where n_i is the number of individual taxpayers in *kura i*. We define the sample selection variable =1 if *kura i* is in the poll-tax sub-sample, and =0 if not. We then weight each *kura* by a frequency weight that is equal to its population size in 1897. The estimated Lee's bounds of the effects of Arab settlement and of the HF-visit are [0.27,0.29] and [0.08,0.29], respectively.

Table 2: **Local determinants of the poll tax rate in 641–1100**

	=0			=1			
	N	Mean	SD	N	Mean	SD	Diff
Arab settlement in 700–969	314	1.07	1.27	94	1.36	1.09	0.29
							(0.009)
Holy Family visit	321	1.08	1.26	87	1.36	1.12	0.29
							(0.111)

Notes: p-value of the difference in means in parentheses. It is estimated using Wild Cluster Restricted (WCR) bootstrap, with clustering at the *kura* level, Webb weights, and 999,999 replications.
Source: See Supplemental Appendix Section B.1.

greater than 0, which is consistent with both models. There are no local-level data on R_i or T_{ji} under the Arab Caliphate. We thus provide two (indirect) pieces of evidence, based on (1) imputing R_i circa 1200, and (2) observing T_{ji} in 1375.

Caliphate-period evidence. We impute R_i at the *kura*-level i for the four poll tax sub-sample *kuras*: $R^{imp}_{i,1200} = \bar{\tau}_{i,641-1100} \times (1 - \bar{F}_{i,1200})$, where $R^{imp}_{i,1200}$ is the imputed poll tax revenue in dinar per capita for *kura* i in 1200, $\bar{\tau}_{i,641-1100}$ is the average τ_{hi} in *kura* i in 641–1100, $\bar{F}_{i,1200}$ is the proportion of villages in *kura* i that have no church or monastery circa 1200.[44] Panel (E) of Supplemental Appendix Figure B.1 shows the map of $R^{imp}_{i,1200}$.

We first compare $R^{imp}_{i,1200}$ between *Fayum*, which received Arab settlers in 700–969, and *Aphrodito*, which did not. Both *kuras* have the same HF-visit status (=0). We find that by 1200, *Fayum* had half of *Aphrodito*'s poll tax revenue per capita (0.13 dinar versus 0.27 dinar), suggesting that $\frac{\partial R_i}{\partial c_i} < 0$.[45] This is consistent with the identity-based model. Second, we compare $R^{imp}_{i,1200}$ of *Ihnas* and *Hermopolis*, where the HF-visit status =1, to that of *Fayum*, where the HF-visit status =0. All three *kuras* have the same Arab settlement status (=1). We find that by 1200, *Fayum* had a slightly lower per-capita poll tax revenue (0.13 versus 0.15). This suggests that $\frac{\partial R_i}{\partial r_{ji}} > 0$, which is consistent with both models. When we use the saint-martyr measure, we obtain even stronger results: *Fayum* and *Hermopolis*, where the saint-martyr status =1, had higher $R^{imp}_{i,1200}$ than *Ihnas*, where the saint-martyr status =0 (0.18 versus 0.08).

Post-Caliphate evidence. Next, we examine if c_i and r_{ji} under the Arab Caliphate had an impact on the total tax revenue, post the Caliphate period.

[44] If we weight villages by their population size in 1897, we obtain similar results for Arab settlement and the saint-martyr measure, but the (positive) difference between HF and non-HF *kuras* disappears.

[45] (1) Because $N = 4$, we cannot conduct a statistical test of the difference. (2) The result holds if we control for the saint-martyr measure instead, i.e. comparing *Ihnas* and *Aphrodito* where the saint-martyr =0. (3) Village-level data in 1245 from *Fayum* reveal that it had a small number of non-converts, and a low poll tax revenue (Rapoport 2018), which is consistent with our findings.

Measuring T_{ji}. For this purpose, we construct data on the total tax transfer (*'ibra*) per unit of land, $\tilde{T}_{ji}$, in village j within *kura* i, from the Mamluk-period cadastres of 1315 (land area) and 1375 (total tax transfer), the earliest extant cadastres with local-level data on total tax transfer.[46] We do not observe the discriminatory tax revenue (R_{ji}), though (recall that $T_{ji} = R_{ji} + \lambda_{ji}$). In 1375, the state estimated the village's average yearly tax revenue, when granting its tax collection right to a beneficiary (LA). The assignment of villages to LAs depended on total tax revenue: Mamluks were granted high-revenue villages, depending on their military rank, as a compensation for their military services to the Sultan (mobilizing and training soldiers, paying for military expenses).[47] Non-Mamluk LAs included Bedouin tribes who were granted villages in compensation for their military service, and other individuals who had to pay the village's tax worth in advance in a first-price auction. Panel (F) of Supplemental Appendix Figure B.1 shows the spatial distribution of $\tilde{T}_{ji}$.

From a theoretical perspective, Arab settlement in 700–969 may still (negatively) affect the tax revenue in 1375. Mamluks' CA and LAs were probably extractive ($c = c_i = 0$): they only cared about the village's tax worth, which they had to pay for upfront in the form of military expenses or in an auction.[48] The identity-based dynamic model (Section 4.1) thus predicts that, ceteris paribus, the LAs under the Mamluks would (reluctantly) keep the Caliphate's tax rate and revenue unaltered (ratcheting) in Arab-settled *kuras*, and would collect the same or more revenue in non-Arab-settled *kuras*.

Empirical specification. We thus examine the impact of Arab settlement, and of the HF-visit status, on the total tax transfer in 1375, using the same OLS and 2SLS specifications as in equations (1) and (2) with one additional control: population per unit of land. We also estimate three alternative specifications (Supplemental Appendix Table B.5): (1) Instead of controlling for region fixed effects, we include a full set of province fixed effects according to the administrative division in the 1315/1375 cadastre, finding similar results to our main specification.[49] (2) We include an interaction term of Arab settlement and the HF-visit status, finding that the interaction term is insignificant. (3) We control for the village-level presence of a Mamluk LA in 1375, finding that Mamluk

[46] Supplemental Appendix Section B.3 shows that $\tilde{T}_{ji}$ is equal to T_{ji} only if population per unit of land, and yield per unit of land, are both held constant for all j. We control for both variables.

[47] Regardless of revenue, Mamluks were also more likely to be granted villages closer to Cairo (where the Sultan was) and to the southern border (for military defense against Nubian attacks).

[48] Persecutions under the Mamluks took place during the early Mamluk period from 1250 to 1354, i.e. before the 1375 cadastre, but subsided afterwards. We explain these persecutions in Section 5 by arguing that Mamluk CAs had signaling concerns in 1250–1354.

[49] The 1315/1375 cadastre was administered at the province level: Egypt's administrative division changed circa 1100 from about 40 *kuras* to about 20 provinces (*a'mal*). Because *kuras* are not nested within provinces, we estimate the standard errors using two-way clustering by both *kura* and province.

villages generate higher revenue.[50]

Findings. We show the results with region fixed effects (FE) in Table 3, while we relegate the findings without region FE to Supplemental Appendix Table B.2b for conciseness. The findings reveal two things: (1) The coefficient on Arab settlement is systematically negative in both the OLS and 2SLS, whether we include region FE or not. The fact that the effect is imprecisely estimated and statistically insignificant when we do not control for region FE, and is larger and statistically significant in most FE specifications, suggests that regional unobserved heterogeneity in total tax revenue mitigates the (negative) effect of Arab settlement. This heterogeneity may arise due to inter-region differences in land quality, public investment in irrigation, and the composition of LAs. In terms of magnitude, based on the 2SLS estimates in column (9) of Table 3, villages in *kuras* where Arabs settled in 700–969 generated a 21% lower tax revenue per unit of land in 1375 than villages in non-Arab-settled *kuras* within the same region. The negative effect of Arab settlement on tax revenue comes in support of the identity-based model, and not the extraction model. (2) The coefficient on the HF-visit status is systematically positive, large in magnitude, and statistically significant with or without region FE. The positive effect of the HF-visit status on tax revenue is consistent with both models. We obtain similar results when we use the saint-martyr measure.[51]

3.2.5 Discussion of the cross-sectional evidence

Our findings indicate that Arab LAs imposed a higher poll tax in 641–1100 than Coptic LAs, induced more conversions to Islam between 641 and 1200, and thus faced lower (imputed) poll tax revenue in 1200. Furthermore, Arab settlement continued to have a negative effect on total tax revenue within regions in 1375, after the collapse of the Caliphate. While these findings are consistent with the identity-based model, there are alternative interpretations of Arab settlement, which we discuss below.

State capacity. A refinement of the extraction model takes into account the constraints faced by tax collectors, here the LAs, in enforcing tax payments.[52] "State capacity" reflects the ability to collect taxes. In our context, the identity of the LA, captured by Arab settlement, might be correlated with the LA's local state capacity. The sign of

[50] We do not prefer this specification, though, because Mamluk presence is endogenous, and may thus bias the coefficients of the other regressors.

[51] In addition to the robustness checks that we have already mentioned, we also measure T_{ji} in 1477 which yields similar results, and we allow for spatial autocorrelation which yields statistically significant results only in certain specifications in the 2SLS region-FE model (see Supplemental Appendix Table B.5 and Webpage Appendix D).

[52] A large literature has investigated optimal taxation from the point of view of Beckerian "economics of crime": see e.g. Allingham and Sandmo (1972), Cremer and Gahvari (1993) and Slemrod (2019).

Table 3: **Local determinants of the total tax transfer in 1375**

Dependent variable: Tax transfer (*'ibra*) in army dinars per unit of land in 1375									
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.51		-0.51	-0.56	-0.45	-0.73	-0.73	-0.85	-0.70
	(0.26)*		(0.26)*	(0.25)**	(0.18)**	(0.34)**	(0.34)**	(0.29)***	(0.28)**
=1 if HF visit (r_{ji})		0.96	0.97	0.98	0.97		0.97	0.99	0.98
		(0.45)**	(0.46)**	(0.46)**	(0.50)*		(0.46)**	(0.46)**	(0.49)**
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1511	1511	1511	1511	1485	1511	1511	1511	1485
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.03	0.03	0.03	0.03	0.06	0.03	0.03	0.03	0.06
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						17.32	17.30	14.40	19.94

Notes: Tax transfer (*'ibra*) is in army dinars (≈13.3/20 dinars) per *feddan* (= 6,368 square meters) of land. Standard errors clustered at the *kura* level are in parentheses. Regions, Byzantine controls, and geographic controls are defined as in Table 1. Population per unit of land is the population in 1897 ÷ land area in 1315. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. A constant is included in all regressions.
Sources: See Supplemental Appendix Section B.1.

the correlation is not obvious in general. A first argument suggests that the tax revenue should be higher in Arab-settled areas: Coptic LAs may have turned a blind eye to tax evasion by friends or been reluctant to impose high fines on detected evaders. A second argument suggests the opposite: Coptic LAs had finer information and therefore a higher ability to collect taxes; so the revenue should have been higher in areas with Coptic LAs.[53]

We argue that the latter information channel is less relevant for the poll tax: papyrological evidence reveals that conversion to Islam (hence, the poll tax base) was perfectly observed by the CA and LAs, whether Arab or Copt.[54] The exclusion of the information channel thus confers predictive power on the state capacity model in our context: In the case of an Arab LA, the poll tax rate and revenue should be the same as in the extraction model. For a Coptic LA, both should be weakly smaller than in

[53] The lack of clear theoretical prediction of the state capacity approach has been noted in the literature; empirically and in a different context, Balan et al. (2021) find in an RCT that local tax enforcers receive more bribes than state enforcers, but that they hold information that in the end enables them to raise more tax revenue for the state despite the leakage.

[54] A papyrological list of converts in 700–900 reveals that a convert had to declare his new Muslim faith in front of the authorities, adopt an Arabic name, and become a client of an Arab patron.

the extraction model.[55] Note, finally, that like the extraction model, the state capacity model does not allow for the possibility of Arab LAs' being on the downward-sloping side of the Laffer curve. Hence, whereas the findings that Arab LAs levied a higher tax rate, and induced more conversions, are consistent with both the state capacity and the identity-based models, the distinguishing prediction of the latter model lies in its implication of a negative, rather than a positive, effect of Arab settlement on tax revenue. Our findings that Arab-settled *kuras* raised lower (imputed) poll tax revenue circa 1200, and that Arab settlement had a negative effect on total tax revenue within regions in 1375, thus come in support of the identity-based model.

Migration. Arab settlement might have affected the proportion of Muslims via migration. This is an unlikely interpretation for the following reasons: (1) At Egypt level, Arab settlement was small relative to Egypt's population, and Copts rarely emigrated from Egypt. (2) Our measure of F_{ji}, Coptic churches and monasteries in 1200, actually depends on the absolute number of non-converts, rather than their population share. Hence, even if Arab settlement reflects a large Arab immigration wave, this will impact our measure only if Arabs turned churches into mosques, in spite of the existence of a large non-convert population (a scenario on which there is no historical evidence). (3) Copt local migration is unlikely, because the Caliphate restricted rural-rural migration (see Supplemental Appendix Section B.2). (4) The other two outcomes, the tax rate and revenue, cannot be explained by the migration mechanism.

Persecutions. Arabs may have coerced Copts to convert via persecutions. This is unlikely, though, because persecutions were relatively rare (see Section 5).

Persuasion. Arabs may have persuaded Copts of the attractiveness of Islam, reducing Copt religiosity and inducing more conversions. But then these areas should have faced a lower, not a higher, tax rate, which is contrary to what we find.

4 Time-series analysis

4.1 Theory

Conceptual framework Although the formal time-series analysis is performed in Supplemental Appendix A, it is useful to explain how the static framework of Section 3 translates into a dynamic one. Time is indexed by $t \in \{1, 2, \ldots, +\infty\}$. Because we

[55] In theory, it is equal if the CA employs a forcing contract and demands a revenue at the peak of the Laffer curve. This latter result is more fragile: if the LA has incomplete information about the district tax capability, then there will be in general a bit of slack and the tax collector can indulge in being a bit nicer to fellow Copts. In that case, the tax rate and revenue would be smaller than in a similar Arab LA.

now are interested in time-series rather than cross-section analysis, we presume for expositional simplicity that there is no heterogeneity among districts and that the LAs have the same identity preferences as the CA: $c_{it} = c_t \geq 0$ for all i (results carry over if there is an identity wedge between CA and LAs). The implementable discriminatory taxes are therefore described by the interval $[\tau^m, \tau^a(c_t)]$. We allow parameters to vary across periods. Let $\{\lambda_t, \tau_t\}$ denote the uniform and discretionary taxes. We initially take λ_t as exogenous, an assumption that is consistent with revenue-maximization that occurs in the absence of threat of rebellion (λ_t is then the extractive uniform tax at date t).

We assume that conversions are permanent. Reswitching may be costly either because of apostasy rules (as in the case of conversions to Islam) or because of the existence of human investments (Jewish intellectuals who left Germany for the US did not come back once politics in Germany returned to normal). Incentive compatibility and apostasy imply that each date t is characterized by a cutoff θ_t^* such that types $\theta \geq \theta_t^*$, and only them, have kept their identity up to date t (included). Because identity switches are permanent, for all t

$$\theta_t^* \geq \theta_{t-1}^*. \tag{3}$$

The date-t CA's instantaneous objective function is (normalizing $r_i = 0$ for all i) $w_t = \lambda_t + \tau_t[1 - F(\theta_t^*)] - c_t[1 - F(\theta_t^*)]$, where $c_t \geq 0$ is the date-t CA's identity strength (we allow this strength to vary over time). We assume for expositional simplicity that all parameters are deterministic (the results fully generalize if they are stochastic). The date-t CA (ruler)'s intertemporal welfare under discount factor $\beta \in [0, 1)$ is:

$$W_t = \sum_{k=0}^{+\infty} \beta^k [\lambda_{t+k} + (\tau_{t+k} - c_t)[1 - F(\theta_{t+k}^*)]]. \tag{4}$$

Because we allow the ruler to change over time, what will be chosen by ruler $t+k$ is evaluated from the point of view of ruler-t preferences. Agent θ's intertemporal welfare is

$$U = \sum_{t=1}^{+\infty} \beta^{t-1} [-\lambda_t - z_t(\tau_t - \theta)]$$

where $z_t = 1$ if the agent has maintained his identity up to date t (included) and $z_t = 0$ otherwise.

We also assume that the CA must meet a per-period budget constraint at level B_t (for example, budgetary needs may be high because of a war),

$$\lambda_t + R_t \geq B_t, \tag{5}$$

where $R_t \equiv \tau_t[1 - F(\theta_t^*)]$ is the discriminatory tax revenue. This constraint may or may not be binding, but we assume that, in equilibrium, it can always be met through some choice of taxes.

No internal or external challenge Assume, first, that there is no internal or external threat to ruler power. The equilibrium of the dynamic game is surprisingly simple: at each moment, (a) the past is summarized by a sufficient statistic, the highest discriminatory tax rate so far; (b) agents behave as if they were myopic; that is, at date t, an agent who has not converted yet converts if and only if the current discriminatory tax, τ_t, exceeds his per-period willingness to pay, θ, for keeping his identity; (c) rulers also behave myopically (maximize w_t subject to (3)), regardless of whether their identity is stronger or weaker than that of their predecessors or successors.

To understand why agents behave myopically, note first that the discriminatory tax is weakly increasing: because agents cannot return to their former identity, there is no point for a ruler lowering the tax below his predecessor's level; put differently, the elasticity of conversions is only one-sided. The agent's option value of waiting for a more empathic ruler therefore does not exist: If the willingness to pay for keeping one's identity lies below the current discriminatory tax, it will also lie below all future discriminatory taxes. To understand why the date-t ruler optimally behaves myopically, assume away budget constraints for simplicity and suppose that the date-$(t+1)$ ruler will have a stronger identity $(c_{t+1} \geq c_t)$. The only way for the date-t ruler to affect his successor's behavior is to induce even more conversions than the latter would want; but this strategy lowers the date-t ruler's payoff relative to playing myopically, both at date t and at date $(t+1)$ as well as the future dates. Conversely, suppose that the date-$(t+1)$ ruler is less eager to convert agents. Then picking the myopic optimum has a double benefit for the date-t ruler as this policy also forces the date-$(t+1)$ to select the date-t ruler's optimum.

This result's corollaries help guide the empirical analysis. They imply that the tax base (weakly) shrinks and the discriminatory tax (weakly) increases over time. But the "ever more religious tax base" result is not a foregone conclusion. Indeed, in a stationary economy in which successive rulers have identical tastes and budgetary needs do not change, the equilibrium tax base and the tax rate are time-invariant. More generally, the tax rate can increase exactly when the current ruler has a stronger identity than his predecessors (keeping the budgetary need the same) or when budgetary needs are relaxed relative to past ones (keeping the rulers' identity strength constant). The reason why a laxer budgetary requirement leads to tax hikes is that identity-based taxation induces taxes on the wrong side of the Laffer curve, where raising money requires lowering taxes.

Again, the predictions differ sharply from those of the extraction model. Revenue maximization implies that neither the tax rate nor the tax revenue respond to ruler-specific parameters such as identity strength or budgetary needs. However, in the identity-based model (of which, recall, the extraction model is a special case), there is a sense in which the extraction model gains relevance over time. Supplemental Appendix Section A.2 puts a formal content on the following claims: (a) Later rulers have less influence on tax policy and outcomes than earlier ones, and (b) the extraction model becomes more and more relevant over time. In our model, the tax policy has two components: a uniform tax that is driven by extractive motives, and a discriminatory tax that reflects both extractive and identity considerations. The claims are based on the idea that over time the ruler (statistically) induces fewer conversions and a smaller decrease in discriminatory tax revenue than his predecessors. This might seem obvious as the group of agents having kept their identity shrinks over time. However, this is not a prediction of the extraction model; relatedly, the remaining non-converts have a very strong identity and are willing to pay much for keeping it.

External threats Suppose now that there is conditional probability $x_t \geq 0$ that the ruler is evicted at date t for some external reason. When the ruler is evicted, taxes – or at least the discriminatory tax– are no longer collected.[56] The uncertainty about the ruler's perennity makes agents more reluctant to convert as doing so eliminates the option value of having kept one's identity. This option value was shown to be equal to 0 in the absence of external threat, but is strictly positive here.

Supplemental Appendix Section A.3 characterizes the equilibrium discriminatory tax rate and revenue in an otherwise stationary economy in which only the conditional probability x_t of upheaval can vary over time. It shows that if the threat of external upheaval decreases over time, so do the discriminatory tax rate and revenue. Intuitively, the rulers charge agents for the option value of keeping their identity that results from a successful upheaval in the future. If the probability of upheaval decreases over time, this option value becomes smaller. Here the predictions are the same in the extraction model.

Internal threats and time-decreasing resistance The most interesting part of the time-series theoretical analysis is its investigation of internal legitimacy. Suppose that the ruler is toppled (and then will not return) when some given fraction of the population coalesce against him. The rebelling decision is rational: it involves a personal cost and occurs if and only if the personal stake of the agents' involved in a successful rebellion exceeds their cost. In each period, a rebelling coalition can form if taxes are "too high" and generate too much discontent. All players- agents, ruler- are far sighted. To

[56] We keep assuming that the ruler cares not only about taxes and current conversions, but also about his "legacy", that is the number of converts in the future even if he is kicked out (so, we keep the payoff function described in (4). In particular, the date-t ruler still internalizes $-c[1-F(\theta^*_{t+k})]$ even if an external challenge has annihilated taxes by date $t+k$).

make the result particularly striking, we assume that the environment is fully stationary: nothing changes over time.

We show that when the minimal coalition for a rebellion to be successful is not too small,[57] the discriminatory tax and tax revenue are not time-invariant. Namely, the threat of rebellion endogenously decreases over time: those who have converted in the past have a single reason to rebel (escaping the uniform tax) while prior to converting they had two (escaping the uniform and discriminatory tax). The ruler optimally back-loads the uniform tax, starting with a low level and increasing it once conversions have reduced the threat of rebellion. Of course, far-sighted agents realize that taxes will increase over time in the absence of rebellion; here the difference in objectives between the marginal rebel and inframarginal ones is important: The marginal rebel is then concerned solely with the discounted flow of uniform taxes; by contrast, agents who do not convert are affected by both the uniform and the discriminatory discounted taxes, as is the ruler. The ruler can soft-pedal uniform taxes and backload their flow so as to dissuade the marginal convert from rebelling. Put differently, he can divide and conquer the agent community. Once the resistance of the converts has been reduced, the ruler can then increase the tax burden. Endogenously time-decreasing resistance also obtains in the extraction model.

This analysis has an interesting corollary in our context of taxation in the early Islam: the land tax was initially capped on religious grounds at a relatively low level. It took over a century to lift the cap. To be certain, modifying a religious precept was costly, but it is unclear why this cost was incurred not at the start but later on. The model is extended to accommodate a cap on the uniform tax and a fixed cost of lifting this cap. The result of an endogenously time-decreasing resistance sheds light on the delay of the tax reform: the uniform tax should be kept low at the start when resistance is strong and raised once it has weakened. Hence, it is optimal for the rulers to wait to reform the tax system and enable higher uniform taxes.

In Supplemental Appendix Section A.4, we also consider the possibility that incomplete information about the discontent or above what it takes to topple the ruler generate equilibrium-path rebellions. At the start, rebellions necessarily imply almost exclusively non-converts. Later on, they involve both converts and non-converts, the latter being more prone to rebel as they lose more from the ruler's taxation.

[57] Otherwise the discriminatory tax and tax revenue are time-invariant.

4.2 Empirics

4.2.1 Time-series implications of the extraction and identity-based models

Both the extraction and the identity-based models imply that in a stationary environment, all conversions must occur at date 1, and the outcomes remain constant thereafter: (1) the poll tax rate (τ_t), (2) the proportion of converts (F_t), and (3) the poll tax revenue (R_t). The two models differ, though, in their explanation and prediction of later conversions, tax rises, and revenue busts, in non-stationary environments.[58]

The first key difference between the two models concerns the effect of the identity strength of the central authority (c_t). The identity-based model attributes tax hikes, conversions, and reductions in revenue, after date 1, to spikes in c_t beyond the maximum identity strength level of previous rulers. The extraction model, to the contrary, attributes no role to c_t increases. Two remarks are in order. (1) The identity-based model offers a second explanation of tax rises, conversions, and revenue reductions, that is not implied by the extraction model: busts in budgetary needs (B_t). (2) Both models imply that the poll tax rate and revenue both go up (down) as external threats (x_t) increase (decrease), while conversions remains unchanged. While our focus in the empirical analysis is on the impact of c_t, because it mirrors our cross-sectional analysis, and because it captures the core of the identity-based model (ruler's preferences), we control for the occurrence of foreign attacks and of adverse economic shocks that can partially account for budgetary busts and external threats.

The second key difference between the two models concerns their predictions of the long-term trajectories of the three outcomes. The extraction model does not predict any new conversions after date 1. It also predicts that the poll tax rate and revenue both remain constant in the absence of external threats. The identity-based model implies that (holding budget needs constant) the probability of poll tax hikes, conversions, and revenue busts decreases over time, as earlier rulers with high identity strength leave an ever-diminishing leeway for the posterity which faces an ever more religious tax base.[59]

4.2.2 Data

To test the two models apart, we exploit the time-series variation in the identity strength of the central authority (c_t) under the Arab Caliphate. We constructed two datasets, the first at the Caliph level ($N = 65$), and the second at the (lower) Egypt's

[58] We discuss the declining threat of rebellion as a potential explanation of the uniform tax rise circa 750, that is shared by both models, in Section 4.2.8.

[59] The tax base may have also grown richer over time due to the positive selection on socioeconomic status of non-convert Copts that was induced by the poll tax regressivity (Saleh 2018).

governor level ($N = 122$), spanning a period of 530 years from 641 to 1170. We use the Caliph-level dataset in the main analysis, and the governor-level dataset as a robustness check. Two remarks are in order. First, the two levels (Caliphs and governors) are identical in 868–905 and 935–1170, when Egypt was either autonomous or had its own Caliphate.[60] Second, both datasets use the same outcomes and controls, measured during the reign of the Caliph or the governor. They differ, though, on their measure of the main treatment, c_t. The Caliph-level dataset employs a Caliph-level measure based on Muslim narratives, whereas the governor-level dataset employs a governor-level measure based on Coptic narratives.

Measuring outcomes. Our econometric analysis focuses on the first two outcomes: the poll tax rate and conversions, which we are able to measure at the Caliph/governor level. The third outcome, the poll tax revenue, is only observed at a few scattered points in time based on Muslim narratives, that do not permit an econometric analysis. However, we include the tax revenue in the long-term trends discussion.

We employ two Coptic medieval chronicles to measure poll tax hikes and conversion waves: *The Chronicle of John of Nikiu* for the Rashidun period (641–661), and the *History of the Patriarchs of the Coptic Church of Alexandria*, the *Liber Pontificalis* (Book of the Popes) of the Coptic Church, for 661–1170.[61] We measure (1) poll tax hikes ($\Delta\tau_t$) by a dummy variable =1 if a poll tax rise is mentioned by Coptic narratives during the reign of Caliph or governor t, and =0 if no poll tax rise is mentioned, and (2) conversion waves (ΔF_t) by a dummy variable =1 if a conversion wave to Islam among Copts is mentioned by Coptic narratives during the reign of Caliph or governor t, and =0 if no conversion wave is mentioned. The zeros in both outcomes imply that τ_t and F_t remained constant; the chronicles do not mention tax decreases or reverse conversions.

Using Coptic narratives to measure tax hikes and conversions offers two advantages. (1) They provide details on the actual poll tax policies of Caliphs and (fiscal) governors (the enforcement of which was delegated to the LAs), instead of the *de jure* poll tax rate that shows little variation between 641 and 1170. (2) They describe conversion waves that are omitted by Muslim sources, and at a higher frequency than Courbage and Fargues (1997)'s and Saleh (2018)'s estimates of the proportion of converts (F_t).

[60] During the late Fatimid period (1074–1170), the Fatimid Caliph was a figurehead, and Egypt was de facto ruled by viziers. Thus, for this period, we treat the vizier as the effective Caliph/governor.

[61] John, the bishop of Nikiu in the 7th century, is a rare eyewitness of the Conquest. *History of the Patriarchs* was compiled by a single author for the period 661–1000: Severus ibn al-Muqaffa, the bishop of *Hermopolis* in the late 9th century, which mitigates the concern about comparability across multiple authors. Scholars agree that Severus "compiled" preexisting Coptic narratives, and did not "create" them, which mitigates the concern that the chronicles merely reflect Severus's own perceptions of events.

Identity strength of CA (c_t). We measure c_t by a binary variable ($\hat{c}_t$) that captures Caliphs' piety in the Muslim sense (i.e. independent of their attitudes toward non-converts), based on Muslim narratives. We focus on one aspect of piety that is arguably comparable across Sunni Caliphs in 641–969 and Ismaili Shiite (Fatimid) Caliphs in 969–1170: a dummy variable =1 if a Caliph is not known for drinking alcohol. We are only able to measure Muslim piety at the Caliph level. Unlike Caliphs (who were absentee rulers of Egypt), governors typically receive only a brief mention in Muslim narratives. We employ a secondary source based on Muslim medieval narratives for 641–868, and two medieval sources: al-Dhahabi's *The Lives of Noble Figures* for 868–969 and al-Maqrizi's *History of the Fatimid Caliphs* for 969–1170.[62]

As a robustness check, we construct a second governor-level measure of c_t that captures Egypt's (fiscal) governors' attitudes (and not policies) toward non-converts, according to the portrayals of governors in the two aforementioned Coptic chronicles. Specifically, we construct a dummy variable =1 if a governor t is portrayed as hostile to non-converts, and =0 if neutral, unmentioned, or friendly.[63] Although Coptic chronicles are often ideologically biased against Muslims, this bias is unlikely to vary systematically across governors.[64]

Although it is more aggregated, we prefer to employ the Caliph-level piety variable in the main analysis, because being based on Muslim narratives on Caliphs' drinking habits, it is plausibly exogenous to their tax policy outcomes: poll tax hikes and conversion waves. By contrast, the governor-level hostility variable may raise an (arguably, inevitable) concern: even if it is based on Coptic perceptions of governors' attitudes, these perceptions can themselves be shaped by governors' behavior, most importantly their tax policy. This can generate a spurious correlation (or reverse causality) between governors' hostility and each of poll tax hikes and conversion waves. However, we think that it is valuable to employ the governor-level measure as a robustness check for two reasons. First, Coptic chronicles offer an important advantage over Muslim narratives: they provide rare fine-grained information on the attitudes of Egypt's (fiscal)

[62] al-Dhahabi is considered one of the main trusted sources among Sunni Muslims on the personal biographies of Muslim politicians and clergy. al-Maqrizi (despite being Sunni) is considered one of the main objective medieval sources on Fatimid Caliphs.

[63] Out of 122 governors, there are 15 who are coded as hostile, 12 coded as friendly, 28 coded as neutral, and 67 who are not mentioned. Although unmentioned governors are the majority, they have a much shorter tenure, on average: Mentioned governors ruled for a total of 374 years, or 71% of the period of 530 years (hostile: 119 years; friendly: 117 years; neutral: 138 years). We pool neutral, unmentioned, and friendly governors, because they are theoretically unlikely to generate poll tax rises and conversions. And indeed they do not differ statistically with respect to the likelihood of poll tax rises and conversions.

[64] Examples of hostile governors include Amr ibn al-As (641–646; 659–663), who "had no mercy on the Egyptians," al-Asbagh (685–705), who "was a hater of the Christians." Examples of friendly governors include al-Layth ibn al-Fadl (798–803) who "was a good man and favored the Christians."

governors, even some of the least known ones, in 641–868 and 905–935, when Egypt was a province of the Caliphate. Second, plotting the two measures in Webpage Appendix Section D reveals that they co-move over time. According to both measures, Caliphs' and governors' identity strength was high during the Rashidun and Umayyad Caliphates (641–750), when most Caliphs are depicted by Sunni Muslims as pious, and most governors are depicted by Coptic chronicles as hostile. Both measures suggest too that ruler's piety/hostility declined during the First Abbasid Caliphate (750–868), when Caliphs are negatively portrayed by Muslim narratives, and governors are positively portrayed by Coptic narratives. The two measures deviate, though, after 868: While Caliph's piety increased during the Second Abbasid Caliphate (868–969) and the Fatimid Caliphate (969–1170), governor's hostility remained low in 868–1170.

Control variables. We include two control variables that are plausibly exogenous. First, we include a dummy variable =1 if there is at least one foreign attack during the reign of Caliph or governor t. A foreign attack is a war that is initiated by a foreign empire, and is hence unanticipated by the Caliphate.[65] Foreign attacks can affect taxation and conversions via increasing the Caliphate's budget need, lowering non-converts' expectations about the Caliphate's tenure, or both. Second, we include a dummy variable =1 if there is at least one adverse Nile shock under Caliph or governor t (Chaney 2013). This captures negative economic shocks that may affect taxation and conversions via reducing taxpayers' ability to pay, increasing the budget need, or both.

4.2.3 Testable hypotheses

Implications of binary measurement. While the theory takes ruler identity to be continuous, our measurement of this variable is binary ($\hat{c}_t = 0, 1$). Supplemental Appendix Section A.5 derives the testable properties that obtain when reality is continuous but the measurement is binary (truncated), implying a loss of information. The probability of a discriminatory-tax hike and conversions is higher when the current ruler has a strong rather than a weak identity. Because the first Caliph/governor in both of our measures is coded as 1 ($\hat{c}_1 = 1$), the maximum measure of level of identity strength of previous rulers for every ruler $t \geq 2$ is equal to 1. This implies the following testable hypothesis:
Hypothesis 1. The identity-based model implies that the probability of poll tax hikes and conversion waves is higher under $\hat{c}_t = 1$ than under $\hat{c}_t = 0$. The extraction model implies no such difference.

[65] Foreign attacks were relatively rare in 641–750, mainly initiated by the Byzantine and Nubian empires. They disappeared after 750, before they spiked in the late Fatimid period, due to the First and Second Crusades (1096–1099, 1147–1152), and the Crusaders' (failed) invasion of Egypt (1154–1169).

The second time-series implication of the identity-based model is that the probability of tax hikes and conversions should decrease over time, as earlier strong-identity rulers leave an ever more religious tax base to the posterity:

Hypothesis 2. The identity-based model implies that the probability of poll tax rises and conversions is decreasing in the number of previous strong identity rulers (for whom $\hat{c}_t = 1$). The extraction model implies no poll tax hikes.

4.2.4 Long-term trends

We first plot the long-term trends of poll tax rises and conversion waves in Supplemental Appendix Figure B.2, and the poll tax revenue in Supplemental Appendix Figure B.3. First, we observe poll tax hikes and conversion waves after "date 1," that can be roughly defined as the Rashidun period (641–661). While the trajectory of conversions is not consistent with the extraction model (which does not predict any new conversions after date 1), the identity-based model explains later tax hikes and conversions by spikes in c_t. Second, both time series reveal a declining trend: tax rises and conversions became less frequent over time, which is consistent with the identity-based model. Third, the poll tax revenue declined rapidly between 641 and 813. Based on these numbers, and assuming a constant *de jure* poll tax rate, Courbage and Fargues (1997) estimate that the proportion of non-Muslims declined from 100% in 641 to 75% in 661, 42% in 680, 33% in 786, and 23% in 813, suggesting that most conversions took place by 680. The poll tax revenue continued to decline albeit at a slower rate through 1090, which is again consistent with the identity-based model. To investigate if later poll tax hikes and conversions can be indeed explained by spikes in c_t, and if the probability of both events is decreasing in the number of previous strong identity rulers, as implied by the identity-based model, we now turn to the econometric evidence.

4.2.5 Empirical specification

To test Hypotheses 1 and 2, we estimate the following model separately for poll tax hikes and conversion waves in 641–1170 at the Caliph level:

$$outcome_t = \beta_0 + \beta_1 \hat{c}_t + \beta_2 n^c_{t-1} + \beta_3 Z_t + \beta_4 year_t + \varepsilon_t \qquad (6)$$

where $outcome_t$ is: (1) a dummy variable =1 if a poll tax rise is mentioned under Caliph t, and (2) a dummy variable =1 if a conversion wave is mentioned under Caliph t. There are two main regressors: (1) $\hat{c}_t$ =1 if Caliph t is not known for drinking alcohol according to Muslim narratives, (2) n^c_{t-1} is the number of previous pious (non-drinking) Caliphs. The vector Z_t includes two control variables: (1) a dummy variable =1 if there is at least one foreign attack during the reign of Caliph t, (2) a dummy variable =1 if there is at least one adverse Nile shock under Caliph t. We control for a linear trend in time by the variable $year_t$, the year when Caliph t ascended to power.

Three remarks are in order. First, while Supplemental Appendix Section A.5 shows that tax hikes and conversions are decreasing in the number of previous pious Caliphs, we also employ the number of previous tax hikes as an alternative measure, finding largely similar results (Supplemental Appendix Table B.6).[66] Second, we weight each regression by the length of Caliph t's tenure. This presumes that Caliphs who stayed longer in power had more influence, which is realistic from a historical viewpoint: it generally took time for Caliphs to consolidate power, align the interests of the governors and LAs with their own, and hence implement policies. Third, we estimate Newey-West standard errors assuming that the error structure is heteroskedastic and serially correlated up to 11 lags.[67]

According to Hypotheses 1 and 2, H_0 on each of β_1 and β_2 is the extraction model which implies that both coefficients are equal to 0, whereas H_1 is the identity-based model which implies that $\beta_1 > 0$ and $\beta_2 < 0$.

4.2.6 Findings

The results, shown in Table 4, come in support of the identity-based model, and are similar for both poll tax hikes (panel (a)) and conversion waves (panel (b)). First, columns (1)–(3) in both panels show that the probability of poll tax hikes and of conversion waves is each increasing in the current Caliph's piety, suggesting that ruler's preferences are an important driver of identity taxation and conversions (Hypothesis 1). In terms of magnitude, column (3) in the two panels shows that pious Caliphs are 29% more likely to trigger a tax hike, and 26% more likely to induce conversion waves, than non-pious Caliphs. Second, columns (4)–(6) show that the effect of the number of previous pious Caliphs on tax hikes and conversions is negative and statistically significant only when we do not include a linear time trend, presumably because of the collinearity between the two variables.[68] The time trend coefficient itself is systematically negative, and reaches statistical significance in the conversions regressions. Taken together, this suggests a role for the diminishing fiscal headroom left for the posterity over time (Hypothesis 2). Finally, when we include both Caliph's piety and the number of previ-

[66]We do not prefer this specification, though, because the number of previous tax hikes is endogenous from a theoretical perspective. We also estimate a specification that includes an interaction term of Caliph's piety and the number of previous pious Caliphs (Supplemental Appendix Table B.6). However, the coefficients become noisy and statistically insignificant.

[67]We determined the number of lags (m) using Lazarus et al. (2018)'s rule of thumb: $m = (1.3)T^{1/2}$, with rounding up, where T is the number of Caliphs.

[68]By the law of large numbers, and assuming that c_t is i.i.d, t is almost exactly proportional to the number of previous pious Caliphs: $n^c_{t-1} \approxeq [1-F(c^*)]t$, where c^* is the cutoff ($\hat{c}_t = 1$ iff $c_t \geq c^*$).

ous pious Caliphs in a horserace in columns (7)–(9), the coefficient on Caliph's piety retains its magnitude and statistical significance, whereas the coefficient on the number of previous pious Caliphs is negative and statistically significant only when we do not include a linear time trend (the latter retains its negative coefficient). We obtain similar effects when we use the governor-level hostility measure, while the latter's coefficient is systematically larger than that of Caliph's piety (Supplemental Appendix Table B.6).[69]

4.2.7 Discussion of the time-series evidence

Both the long-term trends and the econometric evidence are consistent with the identity-based model. There are alternative interpretations of both findings, however: state capacity, migration, persecutions, and persuasion. While the arguments that we put forward regarding these theories for the cross-sectional evidence (Section 3.2.5), apply to the time-series evidence, state capacity requires a separate discussion.

State capacity. The dynamic version of the state capacity model predicts that the Caliphate's extractive power will increase over time, due to the Arabization of the LAs that will reduce the LAs' collusion with Coptic taxpayers.[70] It follows that the proportion of converts, the poll tax rate and revenue will all increase over time. This is consistent with the long-term trends, with the exception of the trajectory of the poll tax revenue that shows a steady decline. Furthermore, if pious Caliphs were more likely to appoint Arab LAs, and if Arabization were difficult to reverse, the state capacity model could also explain why the probability of poll tax hikes and of conversion waves is increasing in Caliph's piety and decreasing in the number of previous pious Caliphs. We think that this theoretically vindicated interpretation is unlikely to hold here for two reasons: (1) Caliphs' Arabization efforts were constrained by both the spatial distribution of Arabs and the limited conversion of Coptic mid-low bureaucrats to Islam (Section 3.2.1). (2) Recall that Arab LAs collected lower, not higher tax revenue.

4.2.8 The uniform tax reform: analytic narrative

As we demonstrated in Section 2, the Caliphate imposed the *kharaj* land tax rate on converts circa 750, which implies, in the language of our model, a removal of the cap on the uniform tax. The long delay between the Arab Conquest and the tax reform that lifted the cap on the uniform tax raises a puzzle for early Islam. While we cannot study

[69]The larger magnitude of the coefficient on governor's hostility can be due to two reasons: (1) agency: Proposition 1, applied to the Caliph-governor agency relationship, implies that (fiscal) governors whose $c^{governor} < c^{Caliph}$ had more influence on tax policy than Caliphs, (2) the potential reverse causality between the outcomes and governor's hostility, which can magnify the effect of the latter.

[70]This resembles the role of fiscal centralization in building state capacity in Europe (Dincecco 2011).

Table 4: **Time-series determinants of poll tax hikes ($\Delta\tau_t$) and conversion waves (ΔF_t) in 641–1170**

(a) Dependent variable =1 if a poll tax hike mentioned during the reign of Caliph t

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if Caliph pious ($\hat{c}_t$)	0.25	0.26	0.29				0.26	0.26	0.28
	(0.12)**	(0.10)***	(0.14)**				(0.10)**	(0.10)***	(0.14)**
No. prev. pious Caliphs (n^c_{t-1})				-0.02	-0.01	-0.01	-0.02	-0.01	-0.00
				(0.01)***	(0.02)	(0.02)	(0.01)***	(0.02)	(0.02)
Caliph's start year		-0.46	-0.40		-0.18	-0.19		-0.29	-0.30
		(0.18)**	(0.30)		(0.64)	(0.71)		(0.53)	(0.60)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (Caliphs)	64	64	64	65	65	65	64	64	64
Years	526	526	526	530	530	530	526	526	526
R^2	0.07	0.16	0.18	0.09	0.09	0.10	0.16	0.16	0.18
p-value (Breusch–Godfrey test)	0.04	0.05	0.02	0.17	0.07	0.05	0.08	0.04	0.02
Mean dep. var.	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25

(b) Dependent variable =1 if a conversion wave mentioned during the reign of Caliph t

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if Caliph pious ($\hat{c}_t$)	0.23	0.25	0.26				0.24	0.25	0.27
	(0.15)	(0.13)*	(0.14)*				(0.13)*	(0.13)*	(0.15)*
No. prev. pious Caliphs (n^c_{t-1})				-0.02	0.02	0.02	-0.02	0.02	0.02
				(0.00)***	(0.01)	(0.01)	(0.01)***	(0.02)	(0.02)
Caliph's start year		-0.65	-0.68		-1.19	-1.25		-1.30	-1.35
		(0.17)***	(0.20)***		(0.47)**	(0.49)**		(0.43)***	(0.47)***
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (Caliphs)	64	64	64	65	65	65	64	64	64
Years	526	526	526	530	530	530	526	526	526
R^2	0.07	0.25	0.29	0.13	0.20	0.23	0.19	0.28	0.31
p-value (Breusch–Godfrey test)	0.01	0.02	0.01	0.01	0.00	0.00	0.02	0.01	0.00
Mean dep. var.	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18

Notes: There are 65 Caliphs. We omit Caliph 1, because for every Caliph $t \geq 2$, $\hat{c}_t = 1$ is interpreted as equal to the maximum $\hat{c}_t$ of previous Caliphs ($\hat{c}_1 = 1$). Newey-West standard errors, assuming that the error structure is both heteroskedastic and autocorrelated up to 11 lags, are in parentheses. Controls are (1) =1 if at least one foreign attack occurred under Caliph t, (2) =1 if at least one adverse Nile shock occurred under Caliph t. Caliph's start year is normalized $\in [0,1]$ by subtracting 641 and dividing it by the maximum value. Regressions are weighted by the length of Caliph's tenure. H_0 for the Breusch-Godfrey test is that there is no serial correlation up to 11 lags. * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Sources: See Supplemental Appendix Section B.1.

the determinants of the timing of the tax reform econometrically, because it was a one-time Caliphate-wide policy change, we use theory to shed light on its potential cause(s). To be certain, there was a religious cost of lifting this constraint, but doing so a century earlier would have given the CA more leeway in raising finances, which was particularly valuable at a time of high budget needs caused by the expansionary Arab conquests in 641–750. One classic "rationalization" for the delay is that the poll tax revenue dwindled due to conversions. This argument however is inconsistent with the extraction model, or for that matter any theory that would not put discriminatory taxation on the downward sloping side of the Laffer curve. The hypothesis we propose is that the CA took advantage of the decreasing-resistance property. Recall that, according to Courbage and Fargues (1997)'s estimates, the proportion of non-converts declined to 33% by 786, implying a strong decline in the internal threat of rebellion. The Abbasid Caliphate thus became more daring to raise the uniform tax. Although the composition of rebels in tax revolts now included both converts and non-converts, the Abbasids eventually managed to suppress the tax revolts by violence, and thus kept the new tax system intact.[71]

5 Extensions

This section discusses two extensions, persecutions and emigration. Webpage Appendix Section C develops other extensions: discrimination through discriminatory access to public goods and services, discriminatory empathy, social norms, and Malthusian ruler.

Persecutions. Can the CA benefit from replacing a discriminatory tax with an alternative proselytic strategy such as coerced conversions or persecutions[72]? Given their ignorance of individual preferences, their ability to reach their goals is constrained by incentive compatibility, the fact that agents with the strongest identity are necessarily less likely to convert. A straightforward generalization of the analysis in Stokey (1979) and Riley and Zeckhauser (1983) for our model shows that a CA obtains its highest welfare through a discriminatory tax, and so there is no restriction involved in assuming this particular approach to inducing conversions.[73] Because a discriminatory tax brings the

[71] We observe 15 tax revolts under the Arab Caliphate, all erupted between 726 and 866. The first 5 revolts (726–783) included only non-converts. Starting from 783, though, rebels included both converts and non-converts.

[72] This does not mean that forced conversions cannot result from our model. Consider the European-African slave trade (suggested to us by Itzchak Tzachi Raz); Europeans force-converted Africans to Christianity, arguing that they were saving their souls from eternal hell (Africans' actual utility obviously differed from Europeans' perception of it). Forced conversions can be understood in the following way in our model: due to their "benevolent" intent, Europeans had a very high utility of conversion (a high c), and so the solution may have been a corner solution with all converting to Christianity (an outcome equivalent to forced conversion). Of course for this to hold, either there must be an upper bound on the support of θ, or the Africans' wealth was limited so that they could not pay a large τ, or both.

[73] The easy observability of individual religious choices matters for the choice of instrument. The deterrence literature stresses that random monitoring calls for much higher penalties to affect behavior,

most revenue to the ruler for any desired level of conversion, a question arises as to why rulers may use (or tolerate) a priori inferior non-price instruments such as persecutions. We offer two theories for why persecutions may arise.

(a1) *Agency*. The first theory flows directly from our agency model. The CA may find the LA too soft toward non-converts and too preoccupied with revenue. Allowing mob persecution may then be a second-best way of inducing more conversions.[74]

Supplemental Appendix Section A.6 introduces the possibility of a dual instrument to induce conversions: a discriminatory tax and persecutions. The latter raises no revenue and would seem to be dominated by the former for that reason. However, persecutions may be optimal for the ruler in the presence of agency. We show that persecutions do not occur when the CA's identity is weaker, the same or slightly stronger than that of the LA; but when the CA's identity is much stronger than the LA's and so the LA is not trusted to enforce religious fervor, the CA (who relies on the LA for tax collection) can in part bypass the LA by ordering persecutions even though persecutions are an inefficient instrument to induce conversions.

(a2) *Signaling*. The ruler may use persecutions as a signal. The signaling hypothesis is a bit more subtle than appears, however. First, a concern for signaling a high c to establish legitimacy with respect to the Muslim community need not per se lead to persecutions: the ruler can be as tough with non-convert Copts by stopping persecutions and raising the poll tax sufficiently, thereby levying a larger revenue. Second, it does not shed light on whether we should expect persecutions, if any, to be a substitute or to complement the poll tax.

In Supplemental Appendix Section A.6, we develop a simple signaling model in which the ruler attempts to establish legitimacy. When the poll tax and tax enforcement are perfectly observable by the ruler, no persecution occurs (again, raising the poll tax is more efficient than allowing persecutions). We argue, though, that imperfect tax observability is a reasonable assumption: the ruler does obtain some information indirectly through occasional observations of conversions, but this does not give him a fully precise idea of the tax collection effort. We show that when the poll tax is imperfectly observable, persecutions (which are assumed to be observable) may arise and furthermore co-vary positively with the poll tax; the intuition is that the marginal cost of persecutions for highly religious rulers is smaller than for less religious ones.

Historians of early Islam are divided on whether conversions to Islam in the Middle

because they are enforced with small probability. This raises the issue of risk aversion or limited liability. Therefore, with infrequent monitoring, non-price instruments, such as jail, the pillory or the death penalty, are more frequent.

[74]The level of persecution could be district specific (as here) or else uniform across districts (in which case only part (i) of the following corollary is relevant).

East were tax-induced or persecution-induced. Historians who endorse the persecution-based narrative trace conversions to persecutions (and state-sponsored mob violence) under the Bahri Mamluks in 1250–1354 (El-Leithy 2005).[75] To shed light on persecutions, we trace the persecution waves that are mentioned in Coptic and Muslim narratives. We find that persecutions were rare under the Arab Caliphate, taking place only under al-Mutawakkil (847–861) and al-Hakim (996–1021), but intensified under the Bahri Mamluks. Furthermore, while we do not observe conversion waves when the poll tax did not rise, we do observe conversions when persecutions did not occur, suggesting that tax-induced conversions were more important.

We then use theory to shed light on the potential cause(s) of these persecution waves. We focus on the signaling model for two reasons. First, we do not have localized data on persecutions under the Caliphate which are necessary for the implications of the agency model. Second, there was no agency problem under the early Mamluks between 1250 and 1315; the poll tax was collected by the CA's agents sent from the capital, and not by the LAs. In line with the signaling model, we split rulers into two categories: most were legitimate (in the sense of having consolidated political power) and had no need to signal; others were less so. Consistent with the model implications, we document that persecutions took place under the less legitimate rulers, and that persecutions were complementary to tax hikes under these same rulers: Al-Mutawakkil was among the least religious Caliphs in history (well known for drinking and sexual scandals) but is widely credited by Sunni Muslims for the "Sunni Revival." Al-Hakim was yet another controversial figure in history considered by some to be an infidel tyrant and by others to be a divine incarnation. Bahri Mamluks were considered as outsiders who stole the Caliphate from the legitimate Caliphs, and hence not very good Muslims.

Emigration. The model allows for emigration in reaction to a strong ruler identity. Suppose that polities do not allow agents to change identity, or that identity is inalterable (race, ethnicity), so an agent's choice is between paying the taxes and emigrating. Think of θ as the agent's willingness to pay to stay in the country. The remaining population corresponds to $\theta \geq \theta^* = \lambda + \tau \equiv \hat{\tau}$. Assuming the uniform tax λ (related to the productivity of land, say) remains constant, $V = \lambda + (\hat{\tau} - c)[1 - F(\hat{\tau} - r)] - T$. Thus a simple relabeling shows that our model captures emigration as well, provided that return is impossible (or at least costly).

Although emigration is less relevant to early Islamic Egypt, it is prominent in many

[75] A third explanation that dates back to al-Maqrizi (died in 1442) attributes conversion to the violent suppression of tax revolts in the 9th century that crashed non-convert Copts' ability to rebel ever since. This explanation is actually a variant of the tax-based narrative, and can be accommodated by our model when there are rebellions in equilibrium.

other historical cases. The first example is the *emigration of Zoroastrians from Iran*. The first wave of emigration occurred under the Sunni Samanids (819–999), an independent state of the Abbasid Caliphate. A second wave occurred under the Qajars (1789–1925). Both emigration waves were in response to the imposition of the poll tax and persecutions.

A second example is the *emigration of Christian minorities in Europe* in response to persecutions. Many Anabaptists (Dutch Mennonites) fled from the south of the Netherlands to Germany, England, and the north of the Netherlands, in response to the discriminatory measures put in place by Phillip II of Spain in 1566. Huguenots (French Calvinists) fled from France to England and other parts of the world, in response to Louis XIV's Edict of Fontainebleau, and the consequent persecution of Protestants.

In both examples, rulers were presumably characterized by $c > 0$. In other situations, though, a ruler is not a unitary unit, but an elite with potentially divided preferences. An important example is the *expulsion of the Moriscos from Spain in 1609*. Following the Christian reconquest of Spain that was completed in 1492, Spanish Muslims were first forced to convert to Christianity via a series of edicts between 1500 and 1525. Forced converts then were perceived to be "crypto-Muslims," until they were eventually expelled from Spain in 1609. Spain's Christian elites were divided on the value they attached to (religious) identity (c), though. On one side, the nobles preferred to exploit their Muslim vassals through forced labor services and a share of their harvest; in the language of our model, they were extractive rulers ($c = 0$). On the other side, the Church and the King attached higher value to religion by achieving religious (Christian) demographic homogeneity, even at the cost of economic loss ($c > 0$). From 1238, date of the conquest of Valencia by King Jaume I of Aragon, through 1525, when Muslims were forced to convert to Christianity, the nobility's extractive motives were prevalent as they succeeded in exploiting Muslims. They kept being so after 1525, but lost the battle in 1609 when the Moriscos were expelled from Spain.[76]

6 Conclusion

The paper offered a theory-guided empirical and historical analysis of identity taxation under the Arab Caliphate. It first developed a simple model of optimal one-shot and repeated taxation/extraction by a polity that trades off its hostility towards a group's identity and its reluctance to let exile, conversions or quits erode the contribution base. It provided a set of testable predictions on how discriminatory taxation and the erosion of the contribution base are impacted by the ruler's and the governed's identity prefer-

[76]See Chaney and Hornbeck (2015) for a detailed study of the economic impact of this episode.

ences and by agency in tax collection. Changes in these explanatory variables as well as uncertainty about the ruler's tenure generate interesting fiscal and identity dynamics. Finally, it showed that the permanent loss of identity dampens one's incentive to rebel, and that the threat of rebellion against fiscal extraction peters out over time, even when those who have altered their identity stay in the constituency (as is the case for religious conversions).

The second contribution is empirical/historical. The paper considered one particular historical event, the incentivized conversion of Egyptian Copts following the Arab conquest in the 7$^{\text{th}}$ century. Building on novel data sources, including tax papyri in 641-1100, list of churches and monasteries in 1200 (as a proxy for conversions), and proxies for local authorities' and Copts' identity strengths, it provided local-level support for the identity-based model, showing that a stronger enforcer identity increased conversions and the discriminatory tax, and reduced tax revenue, suggesting taxation on the downward-sloping side of the Laffer curve and thus providing evidence for the identity-based model. Then, exploiting Coptic and Muslim medieval narratives, it constructed Egypt-level time-series proxies for the central authority's identity strength, and traced discriminatory tax hikes and conversion waves in 641–1170. The Egypt-level time-series evidence also comes in support of the identity-based model. Finally, the combination of theory and history sheds light on how the decline in the threat of rebellion, due to conversions, may have triggered the Caliphate-wide circa 750 tax reform lifting the cap on the non-discriminatory tax. Understanding the determinants of this reform matters because it endogenizes a major "Islamic" institution: the canonical post-750 tax system, instead of treating it as "Islamic," exogenous, and ahistorical.

The theory can in principle be tested in a variety of historical environments where a discriminatory policy was used to induce an unwanted group to change its identity by adopting that of the ruling group, and where the optimal mix of discriminatory and uniform policies evolved in response to changes in taxpayers' identity composition. We mentioned some of these examples in the introduction.

Even though persecutions and emigration played a minor role in our historical context, we discussed how the identity-based model accommodates them. Persecutions are an interesting area of future study, with regard to both the agency problem and signaling, and to the substitutability/complementarity of price and non-price tools of discrimination. Emigration was prominent in many historical episodes, during which oppressed groups dwindled in size. Extending our exploratory theoretical treatment of persecutions and emigration, and performing empirical work along these lines, would be fascinating. For that, one will need to delve in greater depth into the foundations of the ruler's preference function. For example, does the ruler care primarily about population homogeneity? Or does he take a more religious stance of caring about conversions, and if so, how does he conceive his legacy (narrowly as the fraction of minority members in the polity, or broadly as his impact on worldwide conversions)? Particularly interesting would be the study of the strategic interaction, static and dynamic, among multiple rulers to offload or to the contrary attract the minority.

We view this paper as a first step toward further empirical and theoretical studies of optimal identity taxation with time-persistent status changes and their implications for the tax structure and the dynamics of ruler's legitimacy. We hope that it will stimulate empirical work in different historical contexts. We leave these promising alleys for research to future work.

References

Acemoglu, D. (2006). Modeling Inefficient Institutions. In Blundell, R., Newey, W., and Persson, T., editors, *Advances in Economic Theory, Proceeding of World Congress 2005*. Cambridge University Press, Cambridge. Chapter 9.

Allingham, M. G. and Sandmo, A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, 1(3):323 – 338.

Anderson, R. W., Johnson, N. D., and Koyama, M. (2017). Jewish Persecutions and Weather Shocks: 1100–1800. *Economic Journal*, 127(602):924–958.

Artunç, C. and Agir, S. (2017). The Wealth Tax of 1942 and the Disappearance of Non-Muslim Enterprises in Turkey. mimeo.

Augenblick, N., Cunha, J. M., Dal Bó, E., and Rao, J. M. (2016). The Economics of Faith: Using an Apocalyptic Prophecy to Elicit Religious Beliefs in the Field. *Journal of Public Economics*, 141:38–49.

Balan, P., Bergeron, A., Tourek, G., and Weigel, J. (2021). Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the D.R. Congo. *American Economic Review*, Revise and Resubmit.

Barro, R. J. and McCleary, R. M. (2003). Religion and Economic Growth across Countries. *American Sociological Review*, 68(5):760–781.

Becker, C. H. (1902). *Beiträge zur Geschichte Ägyptens unter dem Islam*. Verlag Von Karl J. Trübner, Strasbourg.

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago Press, Chicago.

Becker, S. O. and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, 124(2):531–596.

Bell, H. I. (1910). *Greek Papyri in the British Museum: Catalogue with Texts*, volume IV: The Aphrodito Papyri. London.

Belloc, M., Drago, F., and Galbiati, R. (2016). Earthquakes, Religion and Transition to Self-Government in Italian Cities. *Quarterly Journal of Economics*, 131:1875–1926.

Besley, T. and Persson, T. (2011). The Logic of Political Violence. *The Quarterly Journal of Economics*, 126(3):1411–1445.

Botticini, M. and Eckstein, Z. (2005). Jewish Occupational Selection: Education, Restrictions, or Minorities? *Journal of Economic History*, 65(4):922–948.

Bovenberg, A. and de Mooij, R. A. (1994). Environmental Levies and Distortionary Taxation. *American Economic Review*, 84:1085–1089.

Bovenberg, A. and Goulder, L. H. (1996). Optimal Environmental Taxation in the Presence of Other Taxes : General-Equilibrium Analyses. *American Economic Review*, 86:985–1000.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Cantoni, D., Dittmar, J., and Yuchtman, N. (2018). Religious Competition and Reallocation: The Political Economy of Secularization in the Protestant Reformation. *The Quarterly Journal of Economics*, 133(4):2037–2096.

Chaney, E. (2013). Revolt on the Nile: Economic Shocks, Religion, and Political Power. *Econometrica*, 81:2033–2053.

Chaney, E. and Hornbeck, R. (2015). Economic Dynamics in the Malthusian Era: Evidence from the 1609 Spanish Expulsion of the Moriscos. *Economic Journal*, 90:1404–1440.

Courbage, Y. and Fargues, P. (1997). *Christians and Jews under Islam*. I. B. Tauris Publishers, London-New York. J. Mabro (Tr.).

Cremer, H. and Gahvari, F. (1993). Tax Evasion and Optimal Commodity Taxation. *Journal of Public Economics*, 50(2):261 – 275.

De Long, J. B. and Shleifer, A. (1993). Princes and Merchants: European City Growth before the Industrial Revolution. *The Journal of Law and Economics*, 36(2):671–702.

Delavande, A. and Zafar, B. (2018). Information and anti-american attitudes. *Journal of Economic Behavior & Organization*, 149:1–31.

Diamond, P. (2006). Optimal Tax Treatment of Private Contributions for Public Goods with and without Warm Glow Preferences. *Journal of Public Economics*, 90:897–919.

Dincecco, M. (2011). *Political Transformations and Public Finances: Europe, 1650–1913*. Cambridge University Press.

Dittmar, J. E. and Meisenzahl, R. R. (2020). Public Goods Institutions, Human Capital, and Growth: Evidence from German History. *Review of Economic Studies*, 87(2):959–996.

El-Leithy, T. (2005). *Coptic Culture and Conversion in Medieval Cairo 1293-1524 AD*. PhD thesis, Princeton University, Princeton.

Farhi, E. and Werning, I. (2010). Progressive Estate Taxation. *Quarterly Journal of Economics*, 125(2):635–673.

Fleurbaey, M. and Maniquet, F. (2011). *A Theory of Fairness and Social Welfare*. Cambridge University Press, Cambridge.

Frantz-Murphy, G. (2004). Conversion in Early Islamic Egypt: The Economic Factor. In Conrad, L., editor, *Muslims and Others in Early Islamic Society*. Ashgate, Burlington, VA. pp. 323-330.

Greif, A. and Tadelis, S. (2010). A Theory of Moral Persistence: Crypto-Morality and Political Legitimacy. *Journal of Comparative Economics*, 38(3):229–244.

Grohmann, A. (1932). Aperçu de Papyrologie Arabe. *Etudes de papyrologie*, 1:23–95.

Kaplow, L. (1995). A Note on Subsidizing Gifts. *Journal of Public Economics*, 58:469–477.

Kuran, T. (2012). *The Long Divergence: How Islamic Law Held Back the Middle East*. Princeton University Press, Princeton.

Lazarus, E., Lewis, D. J., Stock, J. H., and Watson, M. W. (2018). HAR Inference: Recommendations for Practice. *Journal of Business & Economic Statistics*, 36(4):541–559.

Michalopoulos, S., Naghavi, A., and Prarolo, G. (2018). Trade and Geography in the Spread of Islam. *Economic Journal*, 128(616):3210–3241.

O'Donoghue, T. and Rabin, M. (2006). Optimal Sin Taxes. *Journal of Public Economics*, 90:1825–1849.

Rapoport, Y. (2018). *Rural Economy and Tribal Society in Islamic Egypt: A Study of al-Nabulusi's Villages of the Fayyum*. Brepols Publishers, Turnhout.

Riley, J. and Zeckhauser, R. (1983). Optimal Selling Strategies: When to Haggle, When to Hold Firm. *Quarterly Journal of Economics*, 98:267–287.

Roodman, D., MacKinnon, J. G., Nielsen, M. Ø., and Webb, M. D. (2018). Fast and Wild: Bootstrap Inference in Stata Using Boottest. mimeo.

Rubin, J. (2017). *Rulers, Religion, and Riches: Why the West Got Rich and the Middle East Did Not*. Cambridge University Press, Cambridge.

Saez, E. and Stantcheva, S. (2016). Generalized Social Welfare Weights for Optimal Tax Theory. *American Economic Review*, 106(1):24–45.

Saleh, M. (2018). On the Road to Heaven: Taxation, Conversions, and the Coptic-Muslim Socioeconomic Gap in Medieval Egypt. *Journal of Economic History*, 78(2):394–434.

Sandmo, A. (1975). Optimal Taxation in the Presence of Externalities. *The Swedish Journal of Economics*, 77:86–98.

Sijpesteijn, P. M. (2013). *Shaping a Muslim State: The World of a Mid-Eighth-Century Egyptian Official*. Oxford University Press, Oxford.

Slemrod, J. (2019). Tax Compliance and Enforcement. *Journal of Economic Literature*, 57(4):904–54.

Stokey, N. (1979). Rational Expectations and Durable Goods Pricing. *Bell Journal of Economics*, 12:112–128.

Tabellini, M. (2020). Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration. *Review of Economic Studies*, 87(1):454–486.

Tirole, J. (2016). From Bottom of the Barrel to Cream of the Crop: Sequential Screening with Positive Selection. *Econometrica*, 84(4):1291–1343.

Voigtländer, N. and Voth, H.-J. (2012). Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany. *Quarterly Journal of Economics*, 127(3):1339–1392.

Wellhausen, J. (1927 [1902]). *The Arab Kingdom and its Fall*. University of Calcutta, Calcutta. M. G. Weir (Tr.).

Supplemental Appendix

A Theory

A.1 Proof of Proposition 1

Let us first imagine that the LA's revenue-collection constraint is not binding. The monotone-hazard-rate condition implies that the LA's optimal discriminatory tax in district i, τ_i^a, is a weakly increasing function of r_i and c_i, with pass-through rates between 0 and 1.

Next we identify the discriminatory tax rates that are implementable by the CA through a transfer requirement. For $c_i < 0$, these are exactly described by the interval $[\tau_i^a(c_i), \tau_i^m]$: For any $\tau_i < \tau_i^a(c_i)$, the LA would raise the discriminatory tax to $\tau_i^a(c_i)$, keep the extra revenue for itself, raise more revenue and reach its optimal discriminatory taxation. For any $\tau_i > \tau_i^m$, the LA would lower the discriminatory tax to τ_i^m or below. Symmetrically, the implementable discriminatory tax rates for $c_i > 0$ are exactly described by the interval $[\tau_i^m, \tau_i^a(c_i)]$.

The upper bound on T_i for $c_i < 0$ is therefore equal to $\lambda_i + R_i^m$. Furthermore, setting $T_i = \lambda_i + R_i^m$ forces the LA to set discriminatory tax τ_i^m, which is as close to τ_i^c as the CA can get. Strict quasi-concavity of the latter's objective function then implies that this transfer requirement is optimal. For $c_i > 0$, the analysis is similar. For $c_i \geq c$, the CA can get its first best by setting $T_i = \lambda_i + R_i(\tau_i^c)$. The LA is then forced to moderate its discriminatory taxation so as to be able to raise enough revenue. Finally, for $c_i \in (0, c)$, the closest implementable tax rate (which is therefore optimal from strict quasi-concavity) is $\tau_i^a(c_i)$; the requested transfer is then $T_i = \lambda_i + R_i(\tau_i^a(c_i))$. ‖

A.2 Dynamics of conversion

Imagine first a world in which *both* rulers and agents are myopic ($\beta = 0$). Consider the tax that yields the CA's unconstrained static optimum under the budget constraint:

$$\tau^*(c_t, B_t - \lambda_t) \equiv \arg_{\{\tau_t[1-F(\tau_t)] \geq B_t - \lambda_t\}} (\tau_t - c_t)[1 - F(\tau_t)].$$

τ^* is increasing in c_t and decreasing in $(B_t - \lambda_t)$. Being myopic, agent θ converts whenever he has not yet converted yet and $\tau_t > \theta$. Ruler t chooses

$$\tau_t = \max\{\tau^*(c_t, B_t - \lambda_t), \theta_{t-1}^*\} \tag{A.1}$$

To understand (A.1), suppose first that $\tau^*(c_t, B_t - \lambda_t) \geq \theta_{t-1}^*$ (as is the case for instance if there have been few or no conversions yet). By definition, $\tau^*(c_t, B_t - \lambda_t)$ yields the static optimum and cutoff $\theta_t^* = \tau^*(c_t, B_t - \lambda_t)$. Next, suppose that $\tau^*(c_t, B_t -$

$\lambda_t) < \theta^*_{t-1}$. In the range $\tau_t \in [0, \theta^*_{t-1}]$, the demand for conversion is inelastic and so the objective function, $\lambda_t + [\tau_t - c_t][1 - F(\theta^*_{t-1})]$, is strictly increasing in τ_t. In either case, $\theta^*_t = \tau_t$. It turns out that these strategies are still optimal when the players value the future:

Proposition 2 ***(dynamics of conversion).*** *For any $\beta \in [0,1)$, there exists a Markov perfect equilibrium in which both the ruler and the agents behave as if they were myopic. The date-t tax and cutoff are $\tau_t = \theta^*_t = \max_{1\leq k\leq t} \tau^*(c_k, B_k - \lambda_k)$. This implies that the tax base shrinks and the discriminatory tax increases over time. In particular:*
(i) If only c_t varies, then $\tau_t = \tau^*(\max_{1\leq k\leq t}\{c_k\}, B - \lambda)$.
(ii) If only $B_t - \lambda_t$ varies, then $\tau_t = \tau^*(c, \min_{1\leq k\leq t}\{B_k - \lambda_k\})$.

The equilibrium can further be shown to be unique if the horizon is finite; and, under additional assumptions, under infinite horizon (the environment considered here) as well. The formal proof of Proposition 2 follows the lines in Tirole (2016).

It can further be checked that, even if the CA does not set taxes itself, it can still, through a transfer demand T_t, induce aligned LAs to implement the policy described in Proposition 2.[77]

The apostasy assumption and its ratcheting corollary validate this "ever more religious tax base" argument, but also show that it is not a foregone conclusion. Indeed, the discriminatory tax and tax revenue are constant in a stationary economy for the identity-based model. They are also constant for a non-stationary economy under the extraction model: In the extraction model, the ruler maximizes $\lambda_t + R(\tau_t, \theta^*_t)$ subject to $R(\tau_t, \theta^*_t) \geq B_t - \lambda_t$, where $R(\tau_t) = \tau_t[1 - F(\theta^*_t)]$. From our assumption that the budgetary need can always be met, then $\tau_t = \tau^m$, the monopoly level that maximizes $\tau[1 - F(\tau)]$, for all t.

Corollary 2 ***(time-series comparison with the extraction model).*** *In the extraction model, the tax base and the discriminatory tax are constant over time.*

Time-increasing relevance of extraction model. Our claim that the extraction model gains in predictive power over time is based on Proposition 2, which states that $\theta^*_t = \max_{1\leq k\leq t} \tau^*(c_k, B_k - \lambda_k)$. Suppose that the joint distribution of the ruler's type c_t and of the net budgetary needs $B_t - \lambda_t$ is the same over time. This generates a distribution $H(\tau_t)$ on some interval $[\underline{\tau}, \overline{\tau}]$ for the date-t ruler's *desired* discriminatory tax

[77] Either $\theta^*_{t-1} > \tau^a(c_t)$, and then LA i's objective function is $\lambda_t + \tau_t[1 - F(\theta^*_{t-1})] - T_t$ for $\tau_t \leq \theta^*_{t-1}$, or the smaller $\lambda_t + \tau_t[1 - F(\tau_t)] - T_t$ for $\tau_t > \theta^*_{t-1}$. Strict quasi-concavity then implies that the LA's optimum is at $\tau_t = \theta^*_{t-1}$. For $\tau^a(c_t) \geq \theta^*_{t-1}$, the equilibrium policy can be decentralized by similarly setting a transfer demand $T_t = \max\{B_t, \lambda_t + \tau^a(c_t)[1 - F(\tau^a(c_t))]\}$.

$\tau^*(c_t, B_t - \lambda_t)$ (the actual one, as we showed, may be constrained by previous choices), and a cumulative distribution function $H^{t-1}(\max_{1\le k\le t-1} \tau^*(c_k, B_k - \lambda_k))$ for the highest discriminatory tax prior to date t. The expected number of conversions at date t is equal to $\int_{\underline{\tau}}^{\overline{\tau}}[\int_{\chi}^{\overline{\tau}}[F(\tau) - F(\chi)]dH(\tau)]dH^{t-1}(\chi)$, which after an integration by parts can be shown to be decreasing in t. Similarly, using the fact that the discriminatory tax is on the downward-sloping side of the Laffer curve, the expected reduction in discriminatory tax revenue from date $t-1$ to date t is $\int_{\underline{\tau}}^{\overline{\tau}}[\int_{\chi}^{\overline{\tau}}[R^c(\chi) - R^c(\tau)]dH(\tau)]dH^{t-1}(\chi)$ and is decreasing in t.

Delegation in time-series model. We assumed for expositional simplicity that LAs were congruent with the CA. However, even in the presence of (possibly district-specific) agency problems, the optimality of myopic behaviors and the ratchet property still hold. Consider for example the case in which, at date t, the CA has identity c_t and the LAs identity c_t^{LA} (again, this identity could be district specific, at the cost of heavier notation). The date-t CA must then account for the implementability constraint $\tau_t \in [\tau^a(c_t^{LA}), \tau^m]$ if $c_t^{LA} \le 0$ and $\tau_t \in [\tau^m, \tau^a(c_t^{LA})]$ if $c_t^{LA} \ge 0$. Let (I) denote this implementability condition. Let $\tau^*(c_t, c_t^{LA}, B_t - \lambda_t) \equiv \arg\max_{\{(I), \tau_t[1-F(\tau_t)] \ge B_t - \lambda_t\}} (\tau_t - c_t)[1 - F(\tau_t)]$ denote the CA's desired static discriminatory tax, characterized in Proposition 1 and Figure 1. The equilibrium discriminatory tax in the time-series model is then $\tau_t = \max\{\tau^*(c_t, c_t^{LA}, B_t - \lambda_t), \theta_{t-1}^*\}$. This implies for example that even ignoring time-varying budget constraints, a strong-identity earlier ruler may not have had the opportunity to convert as many Copts as he desired because of the strong presence of Copts among the LAs, leaving scope for further conversions by subsequent rulers who were not necessarily more religious.

A.3 External threats

In Supplemental Appendix Sections A.3 and A.4 (on external and internal threats, respectively), we make

Assumption 1 *In the rest of the section, $c_t = c$, $\lambda_t = \lambda$, and $B_t = B$ for all t.*

To capture external threats, we assume that there is probability $x_t \ge 0$ that the ruler is evicted at date t conditional on not having been evicted earlier. The sequence $\{x_t\}$ is (for simplicity) known and exogenous. As discussed in the text, the ruler cares about what happens when and when not in power; that is, he cares about his legacy. Except for the sequence $\{x_t\}_{t\ge1}$, all parameters are invariant as stated in Assumption 1, and we suppose that the budget constraint is never binding (the analysis can be generalized if that is not the case).

Proposition 3 ***(option value under external threats)***. *Let* $\tau^c \equiv \arg\max_{\{\theta\}}(\theta - c)[1 - F(\theta)]$ *and* $K_t \equiv \left(1 + \frac{\beta}{1-\beta}x_{t+1}\right)$. *In equilibrium,*[78] *the date-t discriminatory tax is* $\tau_t = K_t\tau^c$ *and the discriminatory tax revenue is* $R_t = K_t\tau^c[1 - F(\tau^c)]$. *In particular, if* x_t *is weakly decreasing (increasing) over time, so are* τ_t *and* R_t. *All conversions occur at date 1.*

Corollary 3 ***(external threats: comparison with the extractive model).*** *The external-threats dynamics for the extractive model are identical with those of the identity model, except that the stable fraction of converts is* $F(\theta^m)$ *where* θ^m *solves* $\max\{\theta[1 - F(\theta)]\}$.

Proof of Proposition 3

The agents' equilibrium strategy can be described by the following cutoff rule at date t:

$$\theta_t^* = \max\left\{\theta_{t-1}^*, \frac{\tau_t}{K_t}\right\}$$

and the discriminatory tax obeys:

$$\tau_t = K_t \max\{\theta^*, \theta_{t-1}^*\}.$$

To see that this is an equilibrium, note that the date-t cutoff, if interior ($\theta_t^* > \theta_{t-1}^*$) satisfies $\left(1 + \frac{\beta}{1-\beta}x_{t+1}\right)\theta_t^* = \tau_t$: Either the ruler is removed at date $(t+1)$ and then the cutoff type enjoys θ_t^* forever; or the ruler remains in place and then type $\theta_t^* \leq \theta_{t+1}^*$ prefers (weakly or strongly) to pay the tax τ_{t+1}.

As for the ruler, note that the equilibrium behaviors deliver the upper bound on his intertemporal payoff that would correspond to the no-external-challenge environment ($x_t \equiv 0$ for all t):

$$W_t^{\max}(\theta_{t-1}^*) = \begin{cases} \dfrac{1}{1-\beta}[\lambda + (\theta^* - c)[1 - F(\theta^*)]] & \text{if } \theta_{t-1}^* \leq \theta^* \\ \dfrac{1}{1-\beta}[\lambda + (\theta_{t-1}^* - c)[1 - F(\theta_{t-1}^*)]] & \text{if } \theta_{t-1}^* > \theta^* \end{cases}$$

To see this, assume that $\theta_{t-1}^* \leq \theta^*$, say (the proof is the same in the opposite case, due to strict quasi-concavity of the adjusted tax revenue). Let the ruler charge $K_t\theta^*$. Then

$$W_t = [\lambda + (K_t\theta^* - c)[1 - F(\theta^*)]] + \beta x_{x+1}\left[-\frac{c[1 - F(\theta^*)]}{1-\beta}\right] + \beta(1 - x_{t+1})W_{t+1}.$$

So W_{t+1} is equal to $\frac{1}{1-\beta}[\lambda + (\theta^* - c)[1 - F(\theta^*)]$, then W_t takes this value as well. The upper bound on the ruler's continuation payoff can be reached though a stationary policy $\theta_{t+k}^* = \max\{\theta^*, \theta_{t-1}^*\}$. So no deviation for any history can yield more that the equilibrium strategy. ∥

[78] As in Proposition 2, equilibrium uniqueness requires further assumptions in the case of an infinite horizon.

A.4 Internal threats and time-decreasing resistance

To facilitate the understanding of endogenously evolving internal challenges, we first gain intuition about the threat of rebellion by analyzing the static case (Proposition 4) and then state our main proposition (Proposition 5).

Static analysis of the rebellion threat. Let us first consider the (noiseless version of the) static case. Assume that it takes $[1-F(\hat{\theta})]$ rebels to topple the CA, and the individual cost of doing so is ρ. In the following, we will say that the threat of rebellion is low (resp. high) if $\hat{\theta}$ is low (high), that is if the number of required rebels is high (low); we could alternatively index the threat of rebellion by (minus) the cost ρ of rebelling. To avoid unnecessary notation, assume $\hat{\theta} \geq 0$. The no-rebellion constraint for taxes $\{\hat{\lambda} \leq \lambda, \hat{\tau}\}$ is that the rebellion cost exceeds the marginal rebel's gain $G(\hat{\theta})$ from a successful rebellion:[79]

$$\rho \geq \hat{\lambda} + \min\{\hat{\tau}, \hat{\theta}\} \equiv G(\hat{\theta}).$$

Assumption 2 ***(relevant rebellion threat).*** $\lambda + \min\{\hat{\theta}, \tau^c\} > \rho$.

Recall that in the absence of rebellion threat, the CA's first best is $\hat{\lambda} = \lambda$ and $\hat{\tau} = \tau^c$. Were Assumption 2 violated, the threat of rebellion would be irrelevant and the first-best level of taxes $\{\lambda, \tau^c\}$ would prevail. We look at the optimal pair $\{\hat{\lambda} \leq \lambda, \hat{\tau}\}$ of taxes that the CA would like to implement. Let $\tilde{\tau} < \tau^c$ be uniquely defined as $\arg\max\{R^c(\hat{\tau}) - \hat{\tau}\}$ or $\tilde{\tau} + \frac{F(\tilde{\tau})}{f(\tilde{\tau})} = c$ (this is the optimal discriminatory tax when an increase in that tax must be offset 1-for-1 by a decrease in the uniform tax). The CA picks the discriminatory tax rate that maximizes $\hat{\lambda} + (\hat{\tau} - c)[1 - F(\hat{\tau})]$ subject to $\hat{\lambda} \leq \lambda$ (feasibility), $\hat{\lambda} + \min\{\hat{\tau}, \hat{\theta}\} \leq \rho$ (no-rebellion constraint) and $\hat{\tau} \in [\tau^m, \tau^c]$ (implementability). For the sake of simplicity, we do not put any lower bound at 0 for $\hat{\lambda}$ (uniform subsidies are feasible).

Finally, let $\theta^* \in [\tau^m, \tau^c]$ be defined by $\theta^* \equiv R^c(\tau^c) - R^c(\tau^*) + \tau^*$, where $\tau^* \equiv \max\{\tau^m, \tilde{\tau}, \rho - \lambda\} \in [\tau^m, \tau^c]$.

Proposition 4 ***(capping the uniform tax to thwart rebellion: the static case).*** *Under Assumptions 1 and 2,*

(i) For a low threat of rebellion ($\hat{\theta} < \theta^*$), the marginal rebel is a convert; the optimal policy for the CA is to reduce the uniform tax to $\hat{\lambda} = \rho - \hat{\theta} < \lambda$, and to keep the discriminatory tax at $\hat{\tau} = \tau^c$.

(ii) For a high threat of rebellion ($\hat{\theta} > \theta^*$), the optimal policy for the CA is to reduce both the uniform tax from λ to $\rho - \tau^*$ and the discriminatory tax from τ^c to τ^*. The marginal rebel is a non-convert.

[79] Allowing for negative values of $\hat{\theta}$, this condition would be $\rho \geq \lambda + \min\{\max\{\hat{\theta}, 0\}, \tau\}$.

Proof of Proposition 4

Assume that $\lambda + \min\{\hat{\theta}, \tau^c\} > \rho$, so there is a real threat of rebellion. The CA's optimization program is

$$\max_{\{\hat{\tau},\hat{\lambda}\}} \hat{\lambda} + (\hat{\tau} - c)[1 - F(\hat{\tau})]$$

s.t.

$$\begin{cases} \hat{\lambda} + \min\{\hat{\theta}, \hat{\tau}\} \le \rho & \text{(no rebellion)} \\ \hat{\lambda} \le \lambda & \text{(uniform tax cannot exceed its extractive level)} \\ \hat{\tau} \in [\tau^m, \tau^c] \end{cases}$$

Suppose that the CA chooses $\{\hat{\lambda}, \hat{\tau}\}$ such that $\hat{\lambda} + \hat{\tau} \le \rho$ (that is, $\hat{\tau} \le \hat{\theta}$ and so the marginal rebel is a non-convert). Then the CA has welfare $\hat{\lambda} + (\hat{\tau} - c)[1 - F(\hat{\tau})] = \rho - \hat{\tau} + R^c(\hat{\tau})$, which is decreasing in $\hat{\tau}$ for $\hat{\tau} \ge \tilde{\tau}$, where

$$\tilde{\tau} + \frac{F(\tilde{\tau})}{f(\tilde{\tau})} = c.$$

Let us restrict the consideration set for the discriminatory tax. First, $\hat{\tau} < \tau^m$ is not implementable. Next, $\hat{\tau} < \tilde{\tau}$ is always weakly dominated: Consider a small change $\delta\hat{\tau} = +\varepsilon$ and $\delta\hat{\lambda} = -\varepsilon$; then the no-rebellion constraint, $\hat{\lambda} + \min\{\hat{\tau}, \hat{\theta}\} \le \rho$, remains satisfied and $\delta(\hat{\lambda} + R^c(\hat{\tau})) = \varepsilon((R^c)' - 1) > 0$ for $\hat{\tau} < \tilde{\tau}$. Finally, $\tau < \rho - \lambda$ is not feasible unless $\hat{\theta}$ is a convert, i.e. $\lambda + \hat{\theta} = \rho$ and $\hat{\theta} < \tau$.

Let $\tau^* \equiv \max\{\tau^m, \tilde{\tau}, \rho - \lambda\}$. Because we are interested only in the case of a rebellion threat $(\lambda + \tau^c > \rho)$, $\tau^* < \tau^c$.

We distinguish three regions:

Region 1: $\hat{\theta} < \tau^m$. Then $\hat{\theta}$ is a convert, $\lambda + \hat{\theta} = \rho$ and $\hat{\tau} = \tau^c$. Welfare is

$$W^1 \equiv \rho - \hat{\theta} + R^c(\tau^c).$$

Region 2: $\hat{\theta} > \tau^c$. Type $\hat{\theta}$ is then necessarily a non-convert, and

$$W^2 \equiv \rho - \tau^* + R^c(\tau^*).$$

Region 3: $\tau^m \le \hat{\theta} \le \tau^c$. Either $\hat{\tau} > \hat{\theta}$ (the marginal rebel is a convert) and then at the optimum $\hat{\tau} = \tau^c$. Welfare is then $W^3 = W^1$. Furthermore, welfare W^1 can be obtained for any $\hat{\theta} \in [\tau^m, \tau^c]$.

Or $\hat{\tau} \le \hat{\theta}$ (the marginal rebel is a non-convert). Then $\hat{\tau} = \tau^*$, yielding welfare $W^3 = W^2$. But, unlike for W^1, W^2 is not feasible for any $\hat{\theta} \in [\tau^m, \tau^c]$: It requires that $\tau^* \le \hat{\theta}$.

Optimal welfare is therefore W^1 for $\hat{\theta} \in [\tau^m, \tau^*]$. On $[\tau^*, \tau^c]$, note that $dW^1/d\hat{\theta} = -1$ while $dW^2/d\hat{\theta} = 0$. Furthermore

$$W^1(\tau^*) - W^2(\tau^*) = R^c(\tau^c) - R^c(\tau^*) > 0 > W^1(\tau^c) - W^2(\tau^c) = -(\tau^c - \tau^*).$$

Therefore in this interval $W^3 = W^2$ if and only if $\hat{\theta} \ge \theta^*$ where

$$\theta^* \equiv R^c(\tau^c) - R^c(\tau^*) + \tau^*.$$

Putting all three regions together

(1) For $\hat{\theta} < \theta^*$, $W = W^1$, $\hat{\lambda} = \rho - \hat{\theta}$, and $\hat{\tau} = \tau^c$.

(2) For $\hat{\theta} > \theta^*$, $W = W^2$, $\hat{\lambda} = \rho - \tau^*$, and $\hat{\tau} = \tau^*$. ‖

Dynamic analysis of the rebellion threat. Suppose next that $t = 1, 2, ..., +\infty$ and that agents and the CA apply the same discount factor β to future utilities. The assumption that $T = +\infty$ is important here; for, with a finite horizon, the gain from a successful rebellion would decrease over time, generating an artificial increase over time in the relative cost of rebellion (expressed relative to future benefits). We assume that the cost of rebellion is $\rho/(1-\beta)$: while rebellion is a one-shot activity, we normalize its per-period cost to be ρ to facilitate the comparison with the static legitimacy model. The willingness to pay to keep one's identity is still θ per period.We focus on Markov perfect equilibria (MPE).

Proposition 5 ***(far-sighted players and decreasing resistance).*** *Let* $\tau^* \equiv \max(\tau^m, \tilde{\tau}, \rho - \lambda)$ *and* $\theta^* \equiv R^c(\tau^c) - R^c(\tau^*) + \tau^* \in (\tau^*, \tau^c)$*. Under Assumptions 1 and 2:*
(i) If $\hat{\theta} < \theta^*$*, the marginal rebel* $\hat{\theta}$ *converts at date 1. In the CA's optimal MPE, the CA backloads the uniform tax, charging a low uniform tax at date 1 and raising the uniform tax to* $\min\{\lambda, \rho\}$ *once the threat of rebellion has subsided. The discriminatory tax is equal to* τ^c *in all periods.*

If $\hat{\theta} > \theta^*$ *and* $\rho - \tau^* \leq \lambda$*, the marginal rebel* $\hat{\theta}$ *never converts. The discriminatory tax and the uniform tax are equal to* $\tau_t = \tau^*$ *and* $\lambda_t = \rho - \tau^*$ *for all* t*.*

Despite the lack of commitment, the CA's per-period welfare is in both cases the same as in the static model, namely $\rho - \tau^* + R^c(\tau^*)$ *for* $\hat{\theta} \leq \theta^*$ *and* $\rho - \hat{\theta} + R^c(\tau^c)$ *for* $\hat{\theta} \geq \theta^*$*.*
(ii) The MPE maximizing the CA's payoff (characterized in part (i)) is furthermore coalition-proof à la Bernheim-Peleg-Whinston (1987) if $\tau^c \geq \beta \frac{\hat{\theta}+\lambda-\rho}{1-\beta}$ *when* $\hat{\theta} < \theta^*$*, or if* $\hat{\theta} > \theta^*$*.*

Intuition. Assume in a first step that all parties are myopic ($\beta = 0$); in particular, each generation cares about its own welfare, but apostasy creates a linkage between periods as conversions apply to future generations. A key insight is that, when the marginal rebel is a convert, the marginal rebel's incentive to rebel decreases over time, as depicted in Figure A.1(a) in the two-period case. Earlier converts' gain from a successful rebellion is limited to the uniform tax and no longer includes the preservation of their foregone identity. Thus, suppose that the threat of rebellion is not too high: $\widehat{\theta} < \theta^*$. A myopic CA then selects $\{\widehat{\lambda} = \rho - \widehat{\theta}, \widehat{\tau} = \tau^c\}$ at date 1. All types $\theta \leq \tau^c$ including the marginal rebel convert at date 1. Because the marginal rebel cares only about the uniform tax from date 2 on, the no-rebellion constraint at date 2 and at any subsequent date t yields:

$$\lambda_t = \min\{\rho, \lambda\}$$

and the date $t \geq 2$ welfare becomes $\min\{\rho, \lambda\} + R^c(\tau^c)$. This is also the maximal welfare that can be obtained in any given period: The uniform tax cannot exceed $\min\{\rho, \lambda\}$ without triggering a rebellion, and $R^c(\tau^c)$ is the maximum adjusted revenue from the discriminatory tax. All conversions occur at date 1, as the discriminatory tax is constant at τ^c from date 1 on. But the uniform tax increases from $\lambda_1 = \rho - \widehat{\theta} < \min\{\rho, \lambda\}$ to $\lambda_2 = \lambda_3 = \cdots = \min\{\rho, \lambda\}$ once the threat of rebellion has decreased. In particular, it increases to equal the extractive tax if $\lambda \leq \rho$.

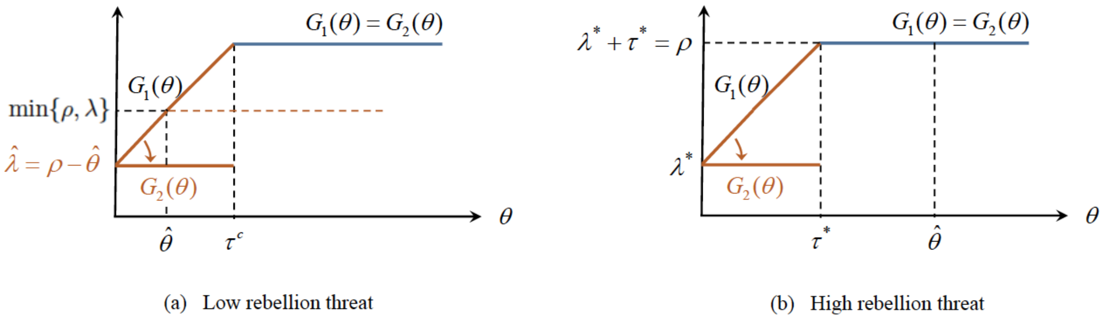


Figure A.1: **Time-decreasing resistance**

Note: $G_t(\theta)$ = type θ's gain from a successful rebellion at date t.

By contrast, when the marginal rebel is a non-convert in the static analysis ($\widehat{\theta} > \theta^*$, see Figure A.1(b)), the threat of rebellion remains the same over time. The CA in each period must still satisfy $\widehat{\lambda}_t + \widehat{\tau}_t \leq \rho$ for each t. And so, $\widehat{\tau}_t = \tau^*$ and $\widehat{\lambda}_t = \rho - \tau^*$ for all $t \geq 1$. The equilibrium is stationary and replicates the static analysis in each period.

When agents are far-sighted ($\beta > 0$), one might guess that the agents' resistance in this case would no longer subside over time, as they internalize the fact that not rebelling will lead to an increase in future taxes. Interestingly, this is not the case. The reason has to do with the difference in objectives between marginal and inframarginal agents when the marginal rebel is a convert; the marginal rebel is then concerned solely with the discounted flow of uniform taxes; by contrast, agents who do not convert are affected by both the uniform and the discriminatory discounted taxes, as is the ruler. The ruler can soft-pedal uniform taxes and backload their flow so as to dissuade the marginal convert from rebelling. Put differently, he can divide and conquer the agent community. Once the resistance of the converts has been reduced, the ruler can then increase the tax burden. The proof of Proposition 5 can be found in Webpage Appendix C.

Timing of the tax reform. We obtain two corollaries in the simple context of myopic agents and rulers. These corollaries also hold when $\beta > 0$. Consider Proposition 5: The uniform tax is initially low to avoid a rebellion, and so a tax reform is not necessary or at least yields low benefits. Once the threat of rebellion has decreased, though, the uniform tax is optimally raised, which may require a tax reform if the initial cap was low.

Corollary 4 ***(timing of tax reform).*** *Suppose that the uniform tax is initially capped by some level $\overline{\lambda}$, and that removing this cap, allowing any level of uniform tax up to the extraction level $\lambda > \overline{\lambda}$, generates some instantaneous cost $C > 0$ for the CA. Under Assumptions 1 and 2,*
(i) If the threat of rebellion is low ($\widehat{\theta} < \widetilde{\theta}$, for some $\widetilde{\theta}$)[80] *and the cap on the uniform tax is binding ($\lambda < \rho - \widehat{\theta}$), then the tax reform occurs at date 1 if $\overline{\lambda} < \rho - \tau^c - C$ and at date 2 if $\overline{\lambda} \in (\rho - \tau^c - C, \min\{\rho, \lambda\} - C)$ (it never occurs if c if higher).*
(ii) If the threat of rebellion is high ($\widehat{\theta} > \widetilde{\theta}$), then the tax reform, if it ever occurs, always occurs at date 1.

Second, we have assumed for simplicity that the CA is well-informed about the threat of rebellion. As a consequence, rebellions constrain the tax system but do not occur on the equilibrium path. With imperfect information about the threat of rebellion, rebellions in general will occasionally occur in equilibrium. When there is little uncertainty, rebellions will be rare. To obtain results about the composition of the rebel group after date 1 (at date 1 all start non-converts, so only non-converts can rebel), we consider the limit of distributions of the rebellion parameters ρ and $\widehat{\theta}$ converging to the certainty case[81]. The intuition behind the following proposition can be grasped from Figure A.1 (a) and (b). Suppose for instance that $min\{\rho, \lambda\} = \rho$ and that the marginal rebel is a convert; a small overestimation of the cost of rebellion will lead converts with types in roughly $[\widehat{\theta}, \tau^c]$ and non-converts with types $\theta \geq \tau^c$ to join the rebellion. Compare this with the case in which the marginal rebel is a non-convert. Then a small overestimation of the level of ρ will lead (almost) only non-converts to rebel.

Corollary 5 ***(composition of rebel group).*** *When the uncertainty about the cost of rebellion is small, at date 1, only non-converts rebel when a rebellion occurs. Later on:*
(i) If the threat of rebellion is low ($\widehat{\theta} < \theta^*$), actual rebellions involve both converts and

[80] $\widetilde{\theta}$ in general differs from θ^*, as the cap affects the welfare in the two regions.

[81] Two comments are in order here. First, we keep the analysis informal. The notion of vanishing uncertainty is the same as in Nash's celebrated non-cooperative Nash demand game when the uncertainty about the size of the endowment vanishes. Second, the uncertainty could affect other parameters than ρ and $\widehat{\theta}$ without changing the analysis.

non-converts. (ii) If the threat of rebellion is high ($\widehat{\theta} > \theta^*$), actual rebellions involve almost only non-converts (the fraction of rebels who are converts tends to 0 as the uncertainty vanishes).

A.5 Implication of theory under binary coding

We here investigate the effect of the binary measurement of c_t on the probability of poll tax hikes and conversion waves under ruler t. We think of our binary measure, $\hat{c}_t$, as truncation at some level c^*: $\hat{c}_t = 1$ if $c_t \geq c^*$ and $\hat{c}_t = 0$ otherwise. Let $c^{t-1} \equiv max_{1\leq k\leq t-1} c_k$, and $\hat{c}^{t-1} \equiv max_{1\leq k\leq t-1} \hat{c}_k$, the associated binary variable. Suppose that c_t is an independent random draw from a distribution $G(c_t)$. Let $n^c_{t-1} \equiv \sum_{1\leq k\leq t-1} \hat{c}_k$ denote the number of realizations $\hat{c}_k = 1$ up to $t-1$. We have for $n^c_{t-1} \geq 1$, $E[F_t - F_{t-1}] = \hat{c}_t \int_{c^*}^{+\infty} [\int_{c^{t-1}}^{+\infty} [F(\tau^a(c_t)) - F(\tau^a(c^{t-1}))] dG(c_t)] \frac{d}{dc^{t-1}} [(\frac{G(c^{t-1}) - G(c^*)}{1-G(c^*)})^{n^c_{t-1}} (c^{t-1})]$. So, in reduced form $E[F_t - F_{t-1}] = \hat{c}_t W(n^c_{t-1})$, where W is a decreasing function converging to 0 as n^c_{t-1} goes to infinity. Similarly, the probability of a tax hike is $P_t = \hat{c}_t \int_{c^*}^{+\infty} [1 - G(c^{t-1})] \frac{d}{dc^{t-1}} [(\frac{G(c^{t-1}) - G(c^*)}{1-G(c^*)})^{n^c_{t-1}} (c^{t-1})]$, and satisfies the same properties as $E(F_t - F_{t-1})$. To sum up, the probability of poll tax rises and conversion waves is increasing in $\hat{c}_t$, and is decreasing in n^c_{t-1}.

A.6 Persecutions

Agency model of persecutions. Consider a CA with identity c and an LA with identity c_i. Express the cost of persecution borne by a non-convert in district i, $p_i \geq 0$, in terms of money, so that the agents' total cost of keeping their identity is $\tau_i + p_i$. Persecution does not bring any cash; it only serves to deter the agents from keeping their identity. The CA chooses the level of acceptable persecutions and the LA then collects taxes.[82]

Corollary 6 ***(agency and persecutions).*** *Consider an economy with parameter sequence $\{c_t, c_{it}, B_t, \lambda_t\}_{t\geq 1}$. Then the tax base shrinks and the discriminatory tax increases over time:*
(i) Persecutions do not occur as long as the CA's identity is not much stronger than the LAs' identity: There exists a function c^* satisfying $c^*(\tilde{c}) > \tilde{c}$ for all $\tilde{c}$ such that there are no persecutions ($p_{it} = 0$) if and only if $c_t \leq c^*(c_{it})$.
(ii) The ruler is more likely to allow persecutions in districts with the weakest identity.

Proof of Corollary 6. For the sake of the argument, suppose that the LA is soft ($c_i > 0$) rather than counterattitudinal (the same reasoning works in the latter case). For a

[82]Persecutions under the Arab Caliphate were ordered by the CA (Caliph or governor).

given p_i, the implementable set is $[\tau_i^m(p_i), \tau_i^a(c_i, p_i)]$, where $\tau_i^a(c_i, p_i)$ is the LA's preferred discriminatory tax, which solves: $max\ (\tau_i - c_i)[1 - F(\tau_i + p_i - r_i)]$ (and $\tau_i^m(p_i)$ solves the same program for $c_i = 0$). So $\tau_i^a(c_i, 0) = \tau_i^a(c_i)$, the discriminatory tax in the persecution-free environment. It is easily shown that persecutions reduce the discriminatory tax as the LA absorbs a fraction of its effect: $\frac{\partial \tau_i^a(c_i,p_i)}{\partial p_i} \in (-1, 0)$. The CA's payoff when the LA sets $\tau_i^a(c_i, p_i)$ (which is the tax in the implementable set that the CA prefers) is $W_i = \lambda_i + [\tau_i^a(c_i, p_i) - c][1 - F(\tau_i^a(c_i, p_i) + p_i - r_i)]$. Simple computations show that $\frac{\partial W_i}{\partial p_i} |_{p_i=0} \propto \frac{\partial \tau_i^a(c_i,p_i)}{\partial p_i} |_{p_i=0} [c - c_i] + c - \tau_i^a(c_i, 0)$. For $c = c_i$, the first term on the RHS is equal to 0 while the second term is strictly negative. The RHS is strictly increasing in c; and for c sufficiently large the CA can guarantee itself λ_i by choosing an infinite level of persecutions and gets strictly less than λ_i when choosing $p_i = 0$. Finally, the cutoff level c_i^* is defined by $\frac{\partial \tau_i^a(c_i,p_i)}{\partial p_i} |_{p_i=0} [c_i^* - c_i] + c_i^* - \tau_i^a(c_i) = 0$. ‖

Signaling model of persecutions. Consider a ruler with unknown religiosity $c \in \{c_L, c_H\}$ with $c_L < c_H$. Make the extreme assumption that the poll tax rate or enforcement, τ, is unobserved by the Muslims, while the level of persecution, p, is perfectly observed.

The ruler is image-concerned and has payoff

$$(\tau - c)[1 - F(\tau + p)] + \mu \hat{c},$$

where $\hat{c} \equiv E[c|p]$ is the rulers' estimated religiosity conditional on the Muslims' information (p) and μ is the intensity of image concerns. Without loss of generality, we have ignore the additive, non-discriminatory tax.

We will look for a separating equilibrium $\{\tau_k, p_k\}_{k \in \{L,H\}}$. As usual, the low type behaves as under full information (i.e. as if there were no image concerns: $p_L = 0$). And so, let

$$W_L \equiv \max_{\{\tau_L\}} (\tau_L - c_L)[1 - F(\tau_L)] + \mu c_L$$

denote the low type's separating equilibrium payoff. As for the high type, $\{\tau_H, p_H\}$ is given by the least-cost-separating policy:

$$\max_{\{\tau_H, p_H\}} (\tau_H - c_H)[1 - F(\tau_H + p_H)] + \mu c_H$$

s.t.

$$\max_{\{\tau\}} (\tau - c_L)[1 - F(\tau + p_H)] + \mu c_H \leq W_L. \tag{A.2}$$

Let us define, letting $\hat{\tau} \equiv \tau + p$,

$$w(c, p) \equiv \max_{\{\hat{\tau}\}} (\hat{\tau} - p - c)[1 - F(\hat{\tau})].$$

For the separating equilibrium to exist, the sorting condition must be satisfied. This is indeed the case, using the envelope theorem:

$$\frac{\partial^2 w}{\partial c \partial p} = f(\hat{\tau}) \frac{\partial \hat{\tau}}{\partial c} > 0,$$

where $\partial\hat{\tau}/\partial c > 0$ results from revealed preference.

The "least-cost-separating equilibrium" corresponds to the value $p_H = p_H^*$, where p_H^* satisfies (A.2) with equality.

Finally, let us see whether the separating equilibrium is consistent with $\tau_H > \tau_L$ (co-variation of τ and p). Suppose a uniform distribution on $[0,1]$: $F(\theta) = \theta$. Then $\tau_H - \tau_L = [(c_H - c_L) - p_H^*]/2$ and p_H^* us given by $\frac{(1-c_L)^2-(1-c_L-p_H^*)^2}{2} = \mu(c_H - c_L) = \mu\Delta c$. Fix c_L and increase Δc.

Note that for (A.2) to have a solution, it must be the case then $p_H^* + c_L \leq 1$, which requires μ not too large. Note also that $(1 - c_L - p_H^*)dp_H = \mu d(\Delta c)$. And so

$$\frac{d(\tau_H - \tau_L)}{d(\Delta c)} = \left[1 - \frac{\mu}{1 - c_L - p_H^*}\right]/2$$

is equal to $1/2$ for $\mu = 0$ (for which $p_H^* = 0$ as well) and remains positive as long as μ is not too large. Because at $\mu = 0$, $\tau_H = \tau_L$, then $\tau_H > \tau_L$ in this range.

B Empirics

B.1 Data sources

B.1.1 Cross-sectional analysis

- Identity strength of local authorities (c_i): Locations of Arab tribes that settled in Egypt in 700–969 are constructed from al-Barri (1992), a secondary source that draws on *al-bayan wal-i'rab 'amman fi ard misr min al-a'rab* (Arab Tribes in Egypt) by al-Maqrizi (died in 1442).

- Identity strength of Copts (r_{ji}): The list of the Holy Family visit villages is from Anba Bishoy (1999) and Gabra (2001); both are based on the apocryphal book *Vision of Theophilus* in Mingana (1931). The list of pre-641 Coptic saints and martyrs is from the Coptic Synaxarium, *Le Synaxaire arabe-jacobite* translated by R. Basset (Basset 1907).

- Proportion of converts (F_{ji}): The list of Coptic churches and monasteries circa 1200 is from *History of Churches and Monasteries* (Abul-Makarim 1984).

- Discriminatory tax rate (τ_{hi}): The individual-level poll tax payments are from Morimoto (1981, pp. 67-79, 85-87) for Greek papyri, and the Arabic Papyrology Database for Arabic papyri.

- Total tax transfer (T_{ji}): The village-level data on total tax transfer (*'ibra*) per unit of land are from Ibn al-Ji'an (1898).

- Byzantine-period *kura*-level controls: The natural logarithm of urban population circa 300 is based on Wilson (2011, pp. 185-187). Byzantine military garrisons circa 600 are constructed from Maspero (1912). *Autopract* estates circa 600 are constructed from Hardy (1931).

- Geographic village-level controls are from the Food and Agriculture Organization's Global Agro-Ecological Zones (FAO-GAEZ) Data Portal 3.0.1. Crop suitability indices are under irrigation and intermediate input level. Population is from the 1897 population census (Ministère des finances 1898).

B.1.2 Time-series analysis

- Poll tax hikes and conversion waves are based on *The Chronicle of John, the Bishop of Nikiu* for 641–661 (John of Nikiu 1916), and *History of the Patriarchs of the Coptic Church of Alexandria* for 661–1170 (Ibn al-Muqaffa 1910, 1943).

- Identity strength of central authority (c_t): Caliph-level piety (not drinking alcohol) is based on Sirhan (1978) for 641–868, al-Dhahabi's *The Lives of Noble Figures* (al-Dhahabi 1982) for 868–969, and al-Maqrizi's *History of the Fatimid Caliphs* (al-Maqrizi 1996) for 969–1170. Governor-level hostility toward non-convert Copts is based on *The Chronicle of John, the Bishop of Nikiu* for the Rashidun period (641–661), and *History of the Patriarchs of the Coptic Church of Alexandria* for 661–1170.

- Control variables: The yearly number of foreign attacks is constructed from Mikaberidze (2011). The yearly occurrence of a Nile adverse shock is constructed from Chaney (2013).

B.1.3 Descriptive figures

Figure B.1 shows the spatial distribution of the cross-sectional outcomes and main regressors. Figures B.2 and B.3 show the evolution of poll tax hikes, conversion waves, and the total poll tax revenue, in 641–1170.

B.2 Measuring the proportion of converts (F_{ji})

We measure the proportion of converts (F_{ji}) by a village-level dummy variable =1 if there is no Coptic church or monastery in village j located within *kura* i circa 1200. Our measure is valid under the following assumptions: (1) the universe of villages is observed in 641 (no post-641 villages), (2) every village had at least one Coptic church

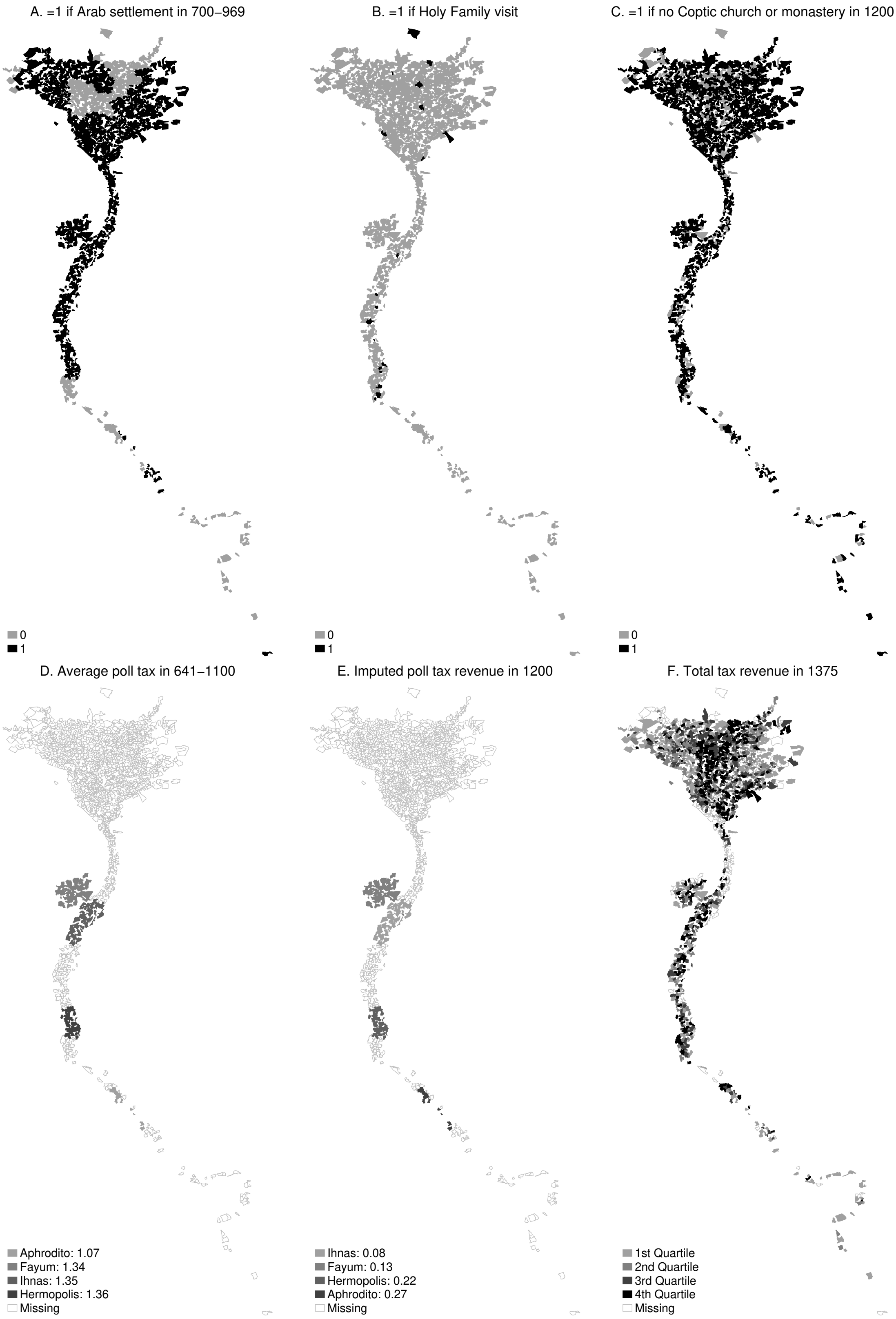


Figure B.1: **Cross-sectional spatial heterogeneity in determinants and outcomes**

Notes: (1) The map shows 1,782 villages in the 1315 cadastre, which defines our universe of villages, using the boundaries of villages in the 2006 population census. (2) Nile Delta refers to the northern triangle on the map. Nile Valley covers the whole region to the south of the Delta.
Sources: See Supplemental Appendix B.1.

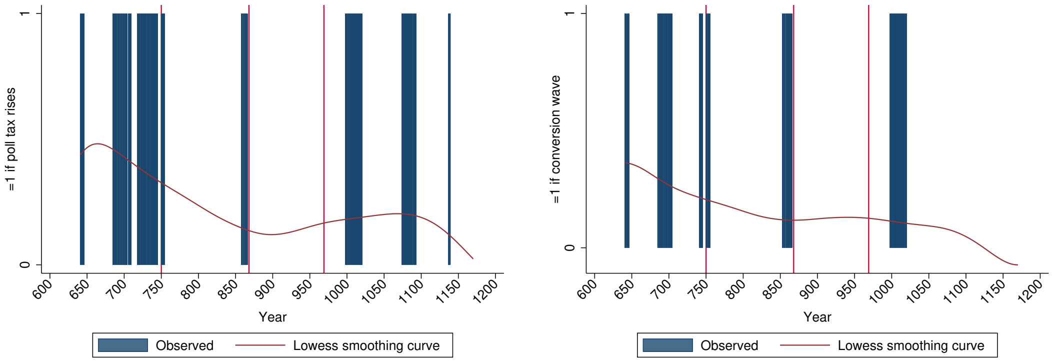


Figure B.2: **Poll tax hikes and conversion waves in 641–1170**

The vertical red lines at years 750, 868, and 969, indicate major dynastic changes. 641–750: Rashidun and Umayyads; 750–868: First Abbasid Period; 868–969: Tulunids, Second Abbasid Period, Ikhshidids; 969–1170: Fatimids. Source: See Supplemental Appendix Section B.1.

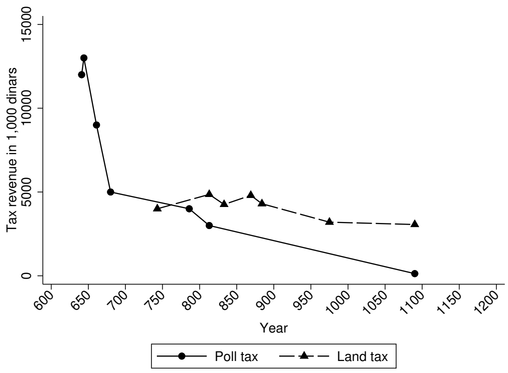


Figure B.3: **Total poll and land tax revenues in 641–1170**

Sources: Courbage and Fargues (1997); poll tax revenue in 1090: Mahmoud (2009).

or monastery in 641, (3) the non-existence of Coptic churches and monasteries in a village in 1200 is the result of the conversion of the vast majority of its population in 641–1200, which led to the desertion of the churches and monasteries or their transformation into mosques, rather than any non-conversion cause (e.g., abandoning a church for financial reasons), (4) the list of churches and monasteries in 1200 is complete, and (5) there is no differential movement of converts and non-converts across villages.

These assumptions are supported by a number of observations. In support of (1), we define the universe of villages based on the 1315 cadastre (Ibn al-Ji'an 1898).[83] Most of these villages existed before 641 (Ramzi 1954). As a robustness check, we further restrict our analysis to a subset of villages mentioned in Byzantine-period sources,

[83] While the earliest extant comprehensive list of Egyptian villages dates to the 1298 cadastre, we chose to digitize the 1315 cadastre instead, because it has information on land area and total tax revenue.

that was compiled by the French archaeologist Amélineau (1850–1915) (Amélineau 1893).[84] The results are qualitatively similar. In support of (2), Amélineau's villages are quite large (mean population in 1897 is 5,900, compared to 2,700 in non-Amélineau villages). Hence, they are most likely to have had at least one church or monastery in 641. In support of (3), our measure is negatively correlated ($\rho = -0.29$) with the actual number of non-convert Coptic households in 1245 among villages in *Fayum kura*, based on al-Nabulsi's *Fayum* cadastre.[85] We also use the individual-level religious affiliation in Egypt's first population censuses in 1848 and 1868, as a robustness check, finding similar results. In support of (4), Abul-Makarim's list is the most complete enumeration of churches and monasteries in medieval Egypt. It has more entries, and geographic coverage, than any other list. We obtain similar results if we use al-Maqrizi's list of churches and monasteries circa 1500, as a robustness check.[86] In support of (5), (a) rural-rural migration was outlawed: papyrological administrative records reveal that "fugitives" who fled their villages were forced to go back, (b) (tax-induced) rural-urban migration is unlikely because cities were controlled by Arab LAs.[87]

B.3 Measuring total tax revenue (T_{ji})

Tax transfer per unit of land ($\tilde{T}_{ji}$) is equal to T_{ji} only if population per unit of land, and yield per unit of land, are both held constant for all j. Let q_{ji} denote the amount of land, z_{ji} the average yield per unit of land, n_{ji} the number of inhabitants, total tax transfer is thus $T_{ji}^{Tot} = q_{ji}z_{ji}\lambda_{ji} + n_{ji}\tau_{ji}(1-F_{ji})$. In the theory, we normalized $q_{ji} = z_{ji} = n_{ji} = 1$. We observe $\tilde{T}_{ji} = \frac{T_{ji}^{Tot}}{q_{ji}} = \lambda_{ji}z_{ji} + \frac{n_{ji}}{q_{ji}}R_{ji}$. Hence, $\tilde{T}_{ji} = T_{ji}$ only if z_{ji} and $\frac{n_{ji}}{q_{ji}}$ are the same for all j. Empirically, we control for z_{ji} by the FAO-GAEZ cereals suitability index, and for $\frac{n_{ji}}{q_{ji}}$ by the population size in 1897 divided by land area in 1315.

B.4 Supplemental Appendix tables

This section presents the Supplemental Appendix Tables that are referenced in the paper.

[84] This is not an exhaustive list of pre-641 villages, though; it only includes villages that were large enough to be mentioned in the Byzantine sources.

[85] The number of Coptic households in the *kura* of *Fayum* is constructed from Rapoport (2018) based on the 1245 cadastre of *Fayum* in *Tarikh al-Fayum (History of Fayum)* by al-Nabulsi (died circa 1250).

[86] This is constructed from al-Maqrizi's *al-Mawa'iz wal-I'tibar fi Zhikr al-Khitat wal-'Athar* (Sermons and Considerations in Examining Plans and Monuments) (al-Maqrizi 2002).

[87] In 1848, when mobility restrictions and the poll tax were both still enforced, the proportion of rural-rural cross-*kura* immigrants is not statistically different between Muslims and Copts (5.7% versus 6.1%).

Table B.1: **Relevance and exogeneity of the instrumental variables**

	(1) =1 if Arab settlement in 700–969	(2) =1 if Holy Family visit	(3) Log (urban population) in 300	(4) =1 if Byzantine garrison in 600	(5) FAO-GAEZ cereals suitability	(6) Mean temperature	(7) Mean Temperature range	(8) Mean slope	(9) Mean rainfall
Kura's distance to *Arish* (km)	0.011	0.005	-0.005	0.008	-0.001	0.000	0.008	0.002	-0.198
	(0.004)**	(0.006)	(0.001)***	(0.005)	(0.000)	(0.003)	(0.004)*	(0.003)	(0.287)
=1 if *Kura* borders desert	3.715	1.168	-1.004	2.223	-0.130	-1.982	0.529	0.093	-36.071
	(1.020)***	(1.419)	(0.492)**	(1.211)*	(0.112)	(0.668)***	(1.058)	(0.677)	(67.489)
Bordering desert × Dist. *Arish*	-0.012	-0.006	0.006	-0.008	0.000	0.009	0.001	0.000	0.053
	(0.004)***	(0.006)	(0.002)***	(0.005)	(0.000)	(0.003)***	(0.004)	(0.003)	(0.288)
Observations	42	42	42	42	42	42	42	42	42
R^2	0.376	0.134	0.109	0.061	0.377	0.828	0.776	0.383	0.379

Notes: White-Huber robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. A constant term is included in all regressions.
Sources: See Supplemental Appendix Section B.1.

Table B.2: **Local determinants of conversions to Islam in 641–1200 and total tax transfer in 1375: No region fixed effects**

(a) Dependent variable: =1 if no Coptic church or monastery in village j circa 1200

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.08		0.08	0.08	0.08	0.13	0.12	0.13	0.12
	(0.03)**		(0.03)**	(0.03)**	(0.03)**	(0.06)**	(0.06)**	(0.06)**	(0.05)**
=1 if HF visit (r_{ji})		-0.59	-0.58	-0.59	-0.62		-0.58	-0.59	-0.62
		(0.08)***	(0.08)***	(0.08)***	(0.09)***		(0.08)***	(0.08)***	(0.09)***
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1782	1782	1782	1782	1751	1782	1782	1782	1751
Clusters (*kuras*)	42	42	42	42	42	42	42	42	42
R^2	0.01	0.03	0.04	0.04	0.05	0.01	0.04	0.04	0.05
Mean dep. var. in control	0.78	0.85	0.78	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat						17.23	17.33	16.40	16.65

(b) Dependent variable: Tax transfer (*'ibra*) in army dinars per unit of land in 1375

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	-0.13		-0.13	-0.30	-0.24	-0.46	-0.45	-0.57	-0.36
	(0.30)		(0.30)	(0.28)	(0.21)	(0.35)	(0.34)	(0.33)*	(0.30)
=1 if HF visit (r_{ji})		0.96	0.96	0.86	0.73		0.97	0.86	0.73
		(0.41)**	(0.42)**	(0.46)*	(0.53)		(0.43)**	(0.46)*	(0.53)
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Population per unit of land?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	1511	1511	1511	1511	1485	1511	1511	1511	1485
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40
R^2	0.00	0.00	0.00	0.01	0.05	-0.00	-0.00	0.01	0.04
Mean dep. var. in control	3.40	3.29	3.40	3.40	3.40	3.40	3.40	3.40	3.40
KP Wald F-stat						16.32	16.42	16.17	14.65

Notes: Tax transfer (*'ibra*) is in army dinars (≈13.3/20 dinars) per *feddan* (= 6,368 square meters). Standard errors clustered at the *kura* level are in parentheses. Byzantine-period *kura*-level controls are: (1) the logarithm of urban population in *kura i* circa 300, and (2) a dummy variable =1 if there was a Byzantine garrison in *kura i* circa 600. Geographic village-level controls are: (3) FAO-GAEZ suitability index to the cultivation of barley, wheat, beans, and maize, under irrigation and intermediate input level, (4) mean temperature, (5) temperature range, (6) slope, and (7) rainfall. Population per unit of land is (8) the population in 1897 ÷ land area in 1315. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. A constant is included in all regressions.
Sources: See Supplemental Appendix Section B.1.

Table B.3: **Local determinants of conversions to Islam in 641–1200: Robustness checks**

Dependent variable: =1 if no Coptic church or monastery in village j circa 1200 (except in columns (3)–(6))

	Amélineau's Byzantine-era villages		=1 if no church or mon. c. 1500		=1 if Copt in 1848–1868		Control interaction ($c_i \times r_{ji}$)		Control *autopract* c. 600		Alternative measure of r_{ji} (Saint-martyr)		SARAR model	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS	(9) OLS	(10) 2SLS	(11) OLS	(12) 2SLS	(13) OLS	(14) 2SLS
=1 if Arab settlement (c_i)	0.16	0.08	0.04	0.03	-0.03	-0.05	0.07	0.09	0.10	0.07	0.07	0.08	0.06	0.10
	(0.08)**	(0.15)	(0.01)***	(0.01)***	(0.01)***	(0.02)**	(0.04)*	(0.05)*	(0.07)	(0.19)	(0.03)*	(0.05)*	(0.02)***	(0.03)***
=1 if HF visit (r_{ji})	-0.60	-0.59	-0.27	-0.27	0.12	0.13	-0.64	-0.50	-0.64	-0.64			-0.62	-0.62
	(0.08)***	(0.07)***	(0.08)***	(0.07)***	(0.04)***	(0.04)***	(0.18)***	(0.28)*	(0.11)***	(0.11)***			(0.08)***	(0.08)***
$c_i \times r_{ji}$							0.03	-0.17						
							(0.21)	(0.35)						
=1 if *autopract* c. 600									-0.01	-0.01				
									(0.07)	(0.07)				
=1 if saint-martyr (r_{ji})											-0.51	-0.51		
											(0.09)***	(0.09)***		
Region FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Byzantine controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs (villages)	157	157	1751	1751	16195	16195	1751	1751	575	575	1748	1748	1730	1730
Clusters (*kuras*)	37	37	42	42	42	42	42	42	21	21	42	42		
R^2	0.15	0.14	0.12	0.12	0.09	0.09	0.06	0.06	0.09	0.09	0.05	0.05		
Mean dep. var. in control	0.54	0.54	0.95	0.95	0.12	0.12	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78
KP Wald F-stat		16.33		19.32		5.33		10.99		6.90		19.38		

Robust standard errors clustered at the *kura* level are in parentheses. Regions, Byzantine-period controls, and geographic controls are defined as in Table 1. In columns (7)–(8), the excluded instruments in the first-stage regressions for Arab settlement (c_i) and its interaction with the HF visit status ($c_i \times r_{ji}$) are: (1) *kura*'s distance to *Arish* ($DistancetoArish_i$) (2) =1 if *kura* borders desert ($BorderDesert_i$), (3) $DistancetoArish_i \times BorderDesert_i$, (4) $r_{ji} \times DistancetoArish_i$, (5) $r_{ji} \times BorderDesert_i$, (6) $r_{ji} \times DistancetoArish_i \times BorderDesert_i$. In columns (9)–(10), the excluded instrument in the first-stage regression for Arab settlement (c_i) is $DistancetoArish_i$. Column (13) reports the results of estimating a spatial autoregressive model with spatial autoregressive standard errors (SARAR) with inverse distance weighting matrix estimated using generalized spatial two-stage least squares (GS2SLS) (STATA command spreg). Columns (14) reports the results of estimating a SARAR model with endogenous variables (STATA command spivreg). * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions. The full results are shown in Webpage Appendix D.

Sources: See Supplemental Appendix Section B.1. Columns (1)–(2): Byzantine-era villages listed in Amélineau (1893). Columns (3)–(4): list of Coptic churches and monasteries circa 1500 in al-Maqrizi (2002). Columns (5)–(6): The 1848 and 1868 individual-level population census samples restricted to Egyptian local free Coptic and Muslim employed men of a rural district of origin who are at least 15 years of age and with non-missing information on age, religion, occupation, and district of origin.

Table B.4: **Evaluating the national representativeness of the poll tax sub-sample**

(a) Village-level variables

	Villages out of poll-tax sample			Villages in poll-tax sample			
	N	Mean	SD	N	Mean	SD	Diff
=1 if no church or monastery in 1200	1589	0.83	0.38	195	0.89	0.31	0.061*
=1 if no church or monastery in 1500	1589	0.98	0.15	195	0.94	0.24	-0.038
ibra per *feddan* in 1375	1336	3.23	2.33	176	4.01	6.40	0.793
ibra per *feddan* in 1477	1336	2.78	2.03	176	3.51	6.45	0.749
=1 if on H. Family route	1589	0.01	0.11	195	0.03	0.16	0.014
=1 if pre-641 Coptic saint or martyr	1585	0.02	0.12	195	0.03	0.17	0.016
FAO-GAEZ cereals suitability index	1560	0.68	0.10	191	0.66	0.10	-0.024
Mean temperature	1560	20.98	0.82	191	21.88	0.30	0.899***
Mean temperature range	1560	14.17	1.04	191	16.34	0.23	2.167***
Mean slope	1560	3.43	0.61	191	3.90	0.63	0.467***
Mean rainfall	1560	50.26	33.27	191	6.43	3.31	-43.832***

(b) *Kura*-level variables

	Kuras out of poll-tax sample			Kuras in poll-tax sample			
	N	Mean	SD	N	Mean	SD	Diff
=1 if Arab settlement in 700–969	38	0.63	0.49	4	0.75	0.50	0.118
Log (urban population) in 300	38	10.00	0.73	4	10.57	0.72	0.570
=1 if Byzantine garrison in 600	38	0.42	0.50	4	1.00	0.00	0.579***
Kura's distance to *Arish* (km)	38	354.07	148.34	4	425.86	83.63	71.792
=1 if *Kura* borders desert	39	0.74	0.44	4	1.00	0.00	0.256***

(c) Dep. var. =1 if no Coptic church or monastery in village j c. 1200: Poll tax sub-sample

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) 2SLS	(7) 2SLS	(8) 2SLS	(9) 2SLS
=1 if Arab settlement (c_i)	0.15		0.17	0.24	0.64	0.33	0.24	0.26	0.85
	(0.56)		(0.59)	(0.16)	(0.35)	(0.22)	(0.14)	(0.21)	(0.37)
=1 if HF visit (r_{ji})		-0.51	-0.52	-0.50	-0.66		-0.52	-0.50	-0.66
		(0.78)	(0.79)	(0.70)	(0.91)		(0.75)	(0.78)	(0.83)
Byzantine controls?	No	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls?	No	No	No	No	Yes	No	No	No	Yes
Obs (villages)	195	195	195	195	191	195	195	195	191
Clusters (*kuras*)	4	4	4	4	4	4	4	4	4
R^2	0.01	0.07	0.08	0.08	0.12	-0.01	0.08	0.08	0.12
Mean dep. var. in control	0.75	0.91	0.75	0.75	0.75	0.75	0.75	0.75	0.75
KP Wald F-stat						1.11	1.17	395.65	910.85

Notes: "Diff" reports the slope in the regression: $y = \alpha_1 + \alpha_2 polltaxsample_i + \varepsilon$, where y is the outcome of village j located in *kura* i in panel (a), or the outcome of *kura* i in panel (b), and $polltaxsample_i$ =1 if *kura* i is in the poll tax sub-sample. Standard errors are clustered at the *kura* level in panel (a), and are White-Huber SEs in panel (b). Panel (c): The IV in columns (6)-(9) is the *kura*'s distance to *Arish*. P-values are in parentheses, estimated by clustering standard errors at the *kura* level, using Wild Cluster Restricted bootstrap for OLS, and Wild Restricted Efficient clustered bootstrap for IV, with Webb weights and 999,999 replications. A constant is included in all regressions. Controls are defined as in Table 1.
Sources: See Supplemental Appendix Section B.1.

Table B.5: **Local determinants of the total tax transfer in 1375: Robustness checks**

Dependent variable: Tax transfer (*'ibra*) in army dinars per unit of land in 1375 (except in columns (1)–(2))

	Tax transfer in 1477		Province FE two-way clustering		Control interaction ($c_i \times r_{ji}$)		Control Mamluk LA in 1375		Alternative measure of r_{ji} (Saint-martyr)		SARAR model	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS	(5) OLS	(6) 2SLS	(7) OLS	(8) 2SLS	(9) OLS	(10) 2SLS	(11) OLS	(12) 2SLS
=1 if Arab settlement (c_i)	-0.39	-0.65	-0.46	-0.33	-0.50	-0.81	-0.45	-0.68	-0.45	-0.69	-0.20	-0.12
	(0.16)**	(0.26)**	(0.10)***	(0.12)***	(0.20)**	(0.33)**	(0.18)**	(0.27)**	(0.18)**	(0.28)**	(0.21)	(0.26)
=1 if HF visit (r_{ji})	0.68	0.69	0.84	0.83	0.90	1.04	0.82	0.82			0.86	0.86
	(0.53)	(0.53)	(0.45)*	(0.38)**	(1.08)	(1.30)	(0.45)*	(0.45)*			(0.70)	(0.70)
$c_i \times r_{ji}$					0.19	0.01						
					(1.19)	(1.42)						
=1 if LA in 1375 Mamluk							0.70	0.70				
							(0.18)***	(0.18)***				
=1 if saint-martyr (r_{ji})									0.54	0.54		
									(0.65)	(0.64)		
Region FE?	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE?	No	No	Yes	Yes	No	No	No	No	No	No	No	No
Byzantine controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population per unit of land?	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Obs (villages)	1485	1485	1467	1467	1486	1486	1460	1460	1482	1482	1456	1456
Clusters (*kuras*)	40	40	40	40	40	40	40	40	40	40		
Clusters (provinces)			19	19								
R^2	0.07	0.07	0.10	0.10	0.04	0.04	0.08	0.08	0.06	0.06		
Mean dep. var. in control	2.89	2.89	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.40	3.43	3.43
KP Wald F-stat		19.94		58.59		10.11		20.06		20.07		

Notes: Army dinar ≈13.3/20 dinars. *Feddan* = 6,368 square meters. Robust standard errors clustered at the *kura* level are in parentheses. In columns (3)–(4), standard errors are clustered at both the *kura* and province level (STATA commands reghdfe and ivreg2), where provinces are defined according to the administrative division in the 1315/1375 cadastre. Regions, Byzantine-period controls, and geographic controls are defined as in Table 1. Population per unit of land is the population in 1897 ÷ land area in 1315. In columns (5)–(6), the excluded instruments are: (1) *kura*'s distance to *Arish* ($DistancetoArish_i$) (2) =1 if *kura* borders desert ($BorderDesert_i$), (3) $DistancetoArish_i \times BorderDesert_i$, (4) $r_{ji} \times DistancetoArish_i$, (5) $r_{ji} \times BorderDesert_i$, (6) $r_{ji} \times DistancetoArish_i \times BorderDesert_i$. Column (11) is a spatial autoregressive model with spatial autoregressive standard errors (SARAR) with inverse distance weighting matrix estimated using generalized spatial two-stage least squares (GS2SLS) (STATA command spreg). Column (12) is a SARAR model with endogenous variables (STATA command spivreg). * $p<0.10$, ** $p<0.05$, *** $p<0.01$. A constant is included in all regressions. The full results are shown in Webpage Appendix D.

Sources: See Supplemental Appendix Section B.1.

Table B.6: **Time-series determinants of poll tax hikes ($\Delta\tau_t$) and conversion waves (ΔF_t) in 641–1170: Robustness checks**

(a) Dependent variable =1 if a poll tax hike mentioned during the reign of ruler t

	Governor-level dataset			Control n. prev. tax hikes			Control $\hat{c}_t \times n^c_{t-1}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if ruler's identity strong ($\hat{c}_t$)	0.53	0.53	0.50	0.21	0.16	0.17	0.15	0.20	-0.19
	(0.15)***	(0.14)***	(0.13)***	(0.10)**	(0.12)	(0.14)	(0.22)	(0.27)	(0.44)
N. prev. strong identity rulers (n^c_{t-1})	-0.00	-0.02	-0.05				-0.02	-0.01	-0.03
	(0.01)	(0.02)	(0.04)				(0.01)***	(0.03)	(0.03)
N. previous poll tax hikes				-0.04	-0.07	-0.09			
				(0.02)**	(0.04)*	(0.07)			
$\hat{c}_t \times n^c_{t-1}$							0.01	0.00	0.03
							(0.01)	(0.01)	(0.02)
Ruler's start year		0.38	0.78		0.50	0.96		-0.27	-0.11
		(0.35)	(0.76)		(0.45)	(0.96)		(0.60)	(0.66)
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (governors/Caliphs)	121	121	121	64	64	64	64	64	64
Years	526	526	526	526	526	526	526	526	526
R^2	0.29	0.30	0.33	0.21	0.23	0.27	0.16	0.16	0.20
p-value (Breusch–Godfrey test)	0.01	0.00	0.00	0.04	0.02	0.01	0.06	0.04	0.01
Mean dep. var.	0.13	0.13	0.13	0.25	0.25	0.25	0.25	0.25	0.25

(b) Dependent variable =1 if a conversion wave mentioned during the reign of ruler t

	Governor-level dataset			Control n. prev. tax hikes			Control $\hat{c}_t \times n^c_{t-1}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 if ruler's identity strong ($\hat{c}_t$)	0.47	0.47	0.46	0.19	0.26	0.31	0.32	0.64	0.29
	(0.19)**	(0.19)**	(0.19)**	(0.13)	(0.19)	(0.20)	(0.33)	(0.41)	(0.67)
N. prev. strong identity rulers (n^c_{t-1})	0.00	-0.00	0.01				-0.01	0.05	0.03
	(0.01)	(0.01)	(0.02)				(0.01)	(0.03)*	(0.04)
N. prev. poll tax hikes				-0.04	0.01	0.04			
				(0.01)***	(0.05)	(0.07)			
$\hat{c}_t \times n^c_{t-1}$							-0.00	-0.02	-0.00
							(0.01)	(0.02)	(0.03)
Ruler's start year		0.05	-0.12		-0.81	-1.21		-1.47	-1.36
		(0.20)	(0.44)		(0.72)	(1.07)		(0.48)***	(0.57)**
Controls?	No	No	Yes	No	No	Yes	No	No	Yes
Obs (governors/Caliphs)	121	121	121	64	64	64	64	64	64
Years	526	526	526	526	526	526	526	526	526
R^2	0.33	0.33	0.38	0.21	0.26	0.30	0.19	0.29	0.31
p-value (Breusch–Godfrey test)	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.00
Mean dep. var.	0.07	0.07	0.07	0.18	0.18	0.18	0.18	0.18	0.18

Notes: Newey-West standard errors are in parentheses, assuming that the error is both heteroskedastic and autocorrelated up to 15 lags (governors) and 11 lags (Caliphs). Controls are (1) =1 if foreign attack occurred, (2) =1 if an adverse Nile shock occurred. Ruler's start year is normalized $\in [0,1]$. Regressions are weighted by the length of ruler's tenure. H_0 for the Breusch-Godfrey test is that there is no serial correlation up to 15 lags (governors) and 11 lags (Caliphs). * p <0.10, ** p <0.05, *** p <0.01. A constant is included in all regressions.
Sources: See Supplemental Appendix Section B.1.

Appendix References

Abul-Makarim, S. (1984). *Tarikh al-Kana'is wal-Adyera (History of Churches and Monasteries)*. Dar al-Na'am, Cairo. Anba Samuel (Ed.).

al-Barri, A. K. (1992). *al-Qaba'il al-'Arabiya fi Misr fil Qurun al-Thalatha al-Ula lil Hijra (Arab Tribes in Egypt in the First Three Hijri Centuries)*. Egyptian General Book Authority, Cairo.

al-Dhahabi, S. (1982). *Siyar A'lam al-Nubala' (The Lives of Noble Figures)*. al-Risala, Beirut.

al-Maqrizi, T. (1996). *Itti'az al-Ḥunafa' bi-Akhbar al-A'immah al-Fatimiyin al-Khulafa' (History of the Fatimid Caliphs)*. Islamic Affairs Supreme Council, Cairo. Jamal-el-Din al-Shayyal (Ed.).

al-Maqrizi, T. (2002). *al-Mawa'iz wal-I'tibar fi Zhikr al-Khitat wal-'Athar (Sermons and Considerations in Examining Plans and Monuments)*, volume 4.1. Al-Furqan Islamic Heritage Foundation, London. A. F. Sayyid (Ed.).

Amélineau, E. (1893). *La Géographie de l'Egypte à l'Epoque Copte*. Imprimerie nationale, Paris.

Anba Bishoy, P. (1999). *The Flight of the Holy Family to Egypt*. Nubar Printing House, Cairo.

Basset, R. (1907). Le Synaxaire Arabe-Jacobite (Redaction Copte). In Blundell, R., Newey, W., and Persson, T., editors, *Patrologia Orientalis*. Firmin-Didot, Paris. 1907-1929.

Chaney, E. (2013). Revolt on the Nile: Economic Shocks, Religion, and Political Power. *Econometrica*, 81:2033–2053.

Courbage, Y. and Fargues, P. (1997). *Christians and Jews under Islam*. I. B. Tauris Publishers, London-New York. J. Mabro (Tr.).

Gabra, G. (2001). *Be Thou There: The Holy Family's Journey in Egypt*. American University in Cairo Press, Cairo.

Hardy, E. R. (1931). *The Large Estates of Byzantine Egypt*. Columbia University Press, New York.

Ibn al-Ji'an (1898). *al-Tuhfa al-Saniyya bi Asma' al-Bilad al-Misriya (Book of Names of Egyptian Localities)*. al-Matba'a al-Ahliya, Cairo.

Ibn al-Muqaffa (1910). History of the Patriarchs of the Coptic Church of Alexandria, Volume I. In *Patrologia Orientalis*, volume 1,5,10. Firmin-Didot Et Co., Paris. B. Evetts (Tr.).

Ibn al-Muqaffa (1943). *History of the Patriarchs of the Coptic Church of Alexandria, Volume II: Part I-III, Volume III: Part I*. Publications de la Société d'Archéologie Copte, Cairo. Yassa Abd al-Masih, Aziz Suryal Atiya, O.H.E. Burmester, Antoine Khater (Trs. and Eds.), 1943, 1948, 1959, 1968.

John of Nikiu (1916). *The Chronicle of John, Bishop of Nikiu*. Williams and Norgate, London-Oxford. R. H. Charles (Tr.).

Mahmoud, A. A. (2009). *al-jizya fi misr 1713-1856 (The Poll Tax in Egypt in 1713-1856)*. Supreme Council of Culture, Cairo.

Maspero, J. (1912). *Organisation Militaire de l'Égypte Byzantine*. Bibliothèque de l'École des Hautes-Études, Paris.

Mikaberidze, A. (2011). *Conflict and Conquest in the Islamic World: A Historical Encyclopedia [2 volumes]: A Historical Encyclopedia*. ABC-CLIO, Santa Barbara, CA.

Mingana, A. (1931). Vision of Theophilus. In *Woodbrooke Studies*, volume 3, pages 1–92. W. Heffer and Sons Limited, Cambridge.

Ministère des finances (1898). *Recensement Général de l'Egypte: 1 juin 1897*. Imprimerie nationale, Le Caire.

Morimoto, K. (1981). *The Fiscal Administration of Egypt in the Early Islamic Period*. Dohosha, Kyoto.

Ramzi, M. (1994 [1954]). *al-Qamus al-Gughrafi lil Bilad al-Misriya min 'Ahd Qudama' al-Misriyin ila Sanat 1945 (Geographic Dictionary of Egyptian Localities from the Time of Ancient Egyptians to 1945)*. Egyptian General Book Authority, Cairo.

Rapoport, Y. (2018). *Rural Economy and Tribal Society in Islamic Egypt: A Study of al-Nabulusi's Villages of the Fayyum*. Brepols Publishers, Turnhout.

Sirhan, J. (1978). *al-Musamara wal-Munadama 'ind al-'Arab hatta al-Qarn al-Rabi' al-Hijri (The Boon-Companionship in the Arab World until the 4th Century A.H.)*. PhD thesis, American University of Beirut, Beirut.

Tirole, J. (2016). From Bottom of the Barrel to Cream of the Crop: Sequential Screening with Positive Selection. *Econometrica*, 84(4):1291–1343.

Wilson, A. (2011). City Sizes and Urbanization in the Roman Empire. In Bowman, A. and Wilson, A., editors, *Settlement, Urbanization, and Population*, pages 161–195. Oxford University Press, Oxford.