THÈSE de DOCTORAT

de l'UNIVERSITÉ TOULOUSE CAPITOLE

Présentée et soutenue par

Madame Lisa BOTBOL

Le 10 juillet 2024

Trois essais sur les marches d'appariement

École doctorale : **TSE Toulouse Sciences Economiques**

Spécialité : **Sciences Economiques - Toulouse**

Unité de recherche : **TSE-R - Toulouse School of Economics - Recherche**

Thèse dirigée par *Monsieur Thierry MAGNAC et Madame Ana GAZMURI*

Composition du jury

Rapporteur : M. Guillaume CHAPELLE
Examinateur : M. Laurent GOBILLON
Directeur de thèse : M. Thierry MAGNAC
Co-directrice de thèse : Mme Ana GAZMURI

Three essays on matching markets

Lisa Botbol

July 10, 2024

Acknowledgement

First and foremost, I am deeply grateful to my advisors Thierry Magnac and Ana Gazmuri for their invaluable guidance on research and career management. Thierry's unwavering support and the time and effort he invested in my work have been instrumental to my progress. Ana has been a fantastic mentor and co-author and her advice on non-work matters has also been incredibly helpful.

I would like to thank Martial Dupaigne for his enthusiasm and for keeping me excited about research. His insightful questions and his help in defining my research questions have been particularly beneficial. Christian Hellwig's support has been crucial, and I particularly thank him for teaching me how to better communicate my work to other researchers. I am also thankful to Stéphane Grégoir for his many useful suggestions on the second chapter of this thesis.

I am grateful to Arnaud Maurel for inviting me to Duke University and providing excellent comments and advice. My gratitude also extends to Guillaume Chapelle and Julien Grenet for agreeing to be referees of my PhD thesis, and to Laurent Gobillon for serving on my jury. I also thank Nour Meddahi and Fabrice Collard for their dedication as heads the PhD Program during my time at TSE.

I am immensely grateful to all members of the Econometrics and Industrial Organization departments at TSE for valuable comments and feedback during seminars, workshops, and beyond. The Economics department at Duke University also offered great comments that helped me define my job market paper. I additionally appreciate the useful interactions I have had with the members of ECHOPPE, which have enriched my research experience.

My journey would not have been as enjoyable without the friends I met through TSE: Alae, Alfonso, Anaïs, Diego, Esteban, Gilles, Gökçe, Guillem, Gyung-Mo, Hakan, Hippolyte, JF, Léa, Louis, Luca, Max, Moritz, Peter, Stephan, Tanja, Tim, Tuuli, and Young. I thank all of you for all the great times we had at work and outside, whether on the Toulouse night scene or during our travels. A special thanks to everyone who played basketball with me, making my time at TSE even more memorable.

I am deeply thankful for the support of my family: my mother Evelyne and stepfather Miquele, my father Michel and stepmother Mirjana, my aunt Dominique and her partner Bernard, and my grandmother Claudine. Your love made me feel like I was never alone, and you have always been my constant source of inspiration. This thesis would probably never have happened without my late uncle Philippo, who was a passionate researcher himself and is dearly missed.

I want to thank my friends Ingrid and Odile for always believing in me, making me laugh, and picking me up when I was down. I also feel very privileged to be friends with Eve-Cécile, Julie, Laetitia, Lola, Marie, Marianne, and Nesrine, who have supported me throughout the years, despite them being much cooler.

Finally, I want to express my deepest gratitude to my partner, Simon. Your unwavering care and love have been my rock throughout this journey. You continually challenge me intellectually and have shown me that I am capable of far more than I ever imagined. This thesis is dedicated to you.

Summary

This thesis explores the design of matching markets, where resources are allocated by public bodies without prices as a market-clearing mechanism, focusing on school admissions and social housing assignments. I analyze how these mechanisms influence socioeconomic inequalities and household welfare by studying the decision-making processes of both applicants and authorities. Using reduced form and structural econometric methods applied to data on social housing and school applications and assignments, I model the behavior of applicants (chapters 1 and 2) and authorities (chapters 2 and 3) in the allocation. The aim is to identify frictions that prevent reforms from effectively altering allocation outcomes to reduce segregation, inequalities, or improve welfare.

I find that applicants' socioeconomic backgrounds significantly influence their behavior and preferences. This insight is crucial for policymakers designing reforms targeting vulnerable households, as neglecting these differences may lead to poor estimations of the responses to changes in allocation mechanisms. Additionally, I show that French local social housing authorities apply heterogeneous criteria when assigning housing, varying significantly between cities. This is an important contribution, as the selection criteria used by social housing authorities are typically opaque in France and other developed countries.

Methodologically, I draw on empirical industrial organization literature to model and estimate applicants' behavior using demand estimation techniques in both static and dynamic frameworks. The first two chapters develop counterfactual algorithms to replicate allocations under modified mechanisms, enabling direct comparisons of different designs. Additionally, chapters 1 and 3 use program evaluation tools to identify treatment effects in quasi-experimental contexts.

The first essay, co-authored with my advisor Ana Gazmuri, analyzes the effects of a 2016 Chilean reform that centralized primary school applications and introduced priorities for low-socioeconomic status (SES) students to reduce education inequalities. Using a difference-in-differences framework, we quantify the reform's impact on access to high-quality schools across SES groups. The findings indicate that the reform did not achieve its goal. To understand this outcome, we estimate a demand model based on students' rank-ordered application lists. The analysis reveals significant heterogeneity in preferences across SES groups, explaining why counterfactual analysis indicates increased affirmative action levels do not significantly reduce segregation. The study also highlights that students often rank too few schools, limiting the reform's effectiveness by leaving many unassigned and ultimately enrolled in under-demanded schools. Counterfactual simulations sug-

gest that if students ranked more schools, the reform could better achieve its objectives.

The second essay examines the allocation of social housing in France, using a dynamic model and a novel comprehensive dataset of French social housing applications. This paper compares applicant welfare under different social housing allocation rules. To simulate applicants' behavior under modified rules, the empirical strategy models and estimates applicants' preferences, separating them from expectations of future offers. Estimates of applicants' utility surplus suggest that the current system favors households with French nationality while disadvantaging precarious households such as single mothers. Counterfactual analysis shows that better targeting low-income households would significantly improve welfare, while mechanisms based solely on applicant waiting time, like first-come-first-serve, reduce welfare.

The third essay also examines the French social housing context, focusing on whether social housing is allocated opportunistically around municipal elections. By analyzing data on social housing assignments and leveraging the timing of municipal elections, the study investigates whether the demographic composition of assignees changes in response to electoral cycles. The empirical strategy exploits the dual administration of social housing in France, where some units are managed nationally and remain unaffected by local elections. The results indicate little to no evidence of opportunistic behavior, suggesting that social housing allocations are influenced more by the political leaning of the municipality's mayor rather than by electoral cycles. This finding challenges the notion of widespread electoral manipulation when decentralizing the allocation of local public goods.

These three essays collectively highlight the critical role of allocation mechanisms in shaping outcomes in matching markets. The conclusions shed light on design features that reduce inequalities and improve welfare, providing valuable insights for policymakers aiming to reform matching markets.

Résumé

Cette thèse explore la conception des marchés d'appariement, où les ressources sont allouées par des organismes publics sans utiliser les prix comme mécanisme de régulation, en se concentrant sur les admissions scolaires et les attributions de logements sociaux. J'analyse comment la conception de ces mécanismes influence les inégalités socio-économiques et le bien-être des ménages. Pour ce faire, j'étudie la prise de décision des candidats et des autorités dans le processus d'allocation. J'utilise des méthodes économétriques en forme réduite et structurelle appliquées aux données des candidatures et des attributions de logements sociaux et d'écoles pour modéliser le comportement des candidats (chapitres 1 et 2) et des autorités (chapitres 2 et 3) dans l'allocation. L'objectif est de comprendre quels obstacles peuvent empêcher une réforme de modifier précisément le résultat de l'allocation pour réduire la ségrégation, les inégalités ou améliorer le bien-être des ménages.

Je constate que, dans les deux contextes étudiés, le milieu socio-économique des candidats influence en partie leur comportement et leurs préférences. Cette information est cruciale pour les décideurs lorsqu'ils conçoivent une réforme ciblant les ménages vulnérables, car ne pas tenir compte de cette différence pourrait conduire à mésestimer la réponse à un changement dans le mécanisme d'allocation. Je constate également que les autorités locales françaises en charge des logements sociaux pondèrent de manière hétérogène les caractéristiques des candidats, la méthode de sélection variant d'une ville à l'autre. C'est une contribution importante, car les critères de sélection des candidats par les autorités de logement social sont très opaques en France, comme dans de nombreux autres pays développés.

Sur le plan méthodologique, je m'appuie sur la littérature en économie industrielle empirique pour modéliser et estimer le comportement des candidats, en utilisant spécifiquement des techniques d'estimation de la demande dans des cadres statiques et dynamiques. Les deux premiers chapitres de la thèse développent également des algorithmes contrefactuels pour reproduire les allocations sous un mécanisme modifié, permettant une comparaison directe de différents modèles. De plus, les chapitres 1 et 3 utilisent des outils d'évaluation de programme pour identifier un effet de traitement dans un contexte quasi-expérimental.

Le premier chapitre, coécrit avec ma codirectrice Ana Gazmuri, analyse les effets d'une réforme chilienne de 2016 centralisant les admissions à l'école maternelle et introduisant un avantage pour les élèves socio-économiquement défavorisés afin de réduire les inégalités scolaires. Nous quantifions l'impact de la réforme sur les inégalités d'accès à une école de qualité avec une méthode

de doubles différences et n'observons pas d'amélioration significative. Pour comprendre pourquoi, nous estimons un modèle de demande basé sur les choix d'école dans les candidatures, révélant des préférences hétérogènes entre les élèves de niveaux socio-économiques différents. Nos analyses contrefactuelles montrent que même avec un avantage supplémentaire à l'admission pour les élèves défavorisés, l'amélioration resterait limitée. Toutefois, nos simulations suggèrent que si les élèves postulaient à plus d'écoles, la réforme pourrait être plus efficace, car le nombre d'élèves assignés par défaut à une école qu'ils n'avaient pas choisie diminuerait.

Le deuxième chapitre traite de l'allocation des logements sociaux en France. En utilisant un modèle dynamique et une nouvelle base de données sur les demandes et les attributions, je compare différents mécanismes d'allocation en termes de surplus des ménages. Pour simuler la réaction des ménages à une modification du mécanisme, je modélise et estime les préférences des demandeurs pour les différents logements sociaux, en les identifiant séparément des attentes concernant les offres futures. Les résultats montrent que le surplus des ménages français est supérieur à celui des autres, tandis que les ménages défavorisés comme les familles monoparentales sont désavantagés. L'analyse contrefactuelle indique que favoriser l'attribution aux ménages les plus pauvres améliorerait significativement le surplus, tandis que des attributions basées uniquement sur le temps d'attente le diminueraient.

Pour améliorer la compréhension de l'allocation de logements sociaux en France, le troisième chapitre examine l'opportunisme électoral des maires dans les décisions d'attribution. Grâce à des données sur les attributions, et en les comparant en fonction de leur proximité avec les élections municipales, j'évalue si le type de ménage recevant un logement change selon le cycle électoral. Ma stratégie empirique utilise le fait que certains logements sont administrés par l'État et ne sont donc pas affectés par les élections municipales. Les résultats ne montrent pas de comportement opportuniste des maires, suggérant que les allocations sont influencées par leur couleur politique plutôt que par le cycle électoral. Ces conclusions remettent en question l'idée que la décentralisation augmenterait les manipulations électorales.

Ces trois chapitres soulignent l'influence cruciale du mécanisme sur l'allocation dans un marché d'appariement. Leurs conclusions mettent en lumière les caractéristiques qui réduisent les inégalités et améliorent le bien-être des ménages, informant les politiques de réforme.

Contents

List of Figures

List of Tables

Chapter 1

Affirmative action in centralized admissions: effects on education equality

with Ana Gazmuri

1.1 Introduction

In most countries, access to education is unequal for children of different socio-economic status. Even when allowing parents to choose their most preferred school, disadvantaged students attend on average schools of lower quality, which reinforces inequalities and generates school socio-economic segregation. A common way to mitigate this issue is to implement affirmative action policies giving an advantage to low SES students in the school allocation process.

The effectiveness of affirmative action policies on increasing quality for disadvantaged students depends on how these students choose a school and whether schools are allowed to select their students. There are concerns that schools may have incentives to select out low-SES students.[1] However, disentangling school selectivity from students' application behavior is challenging, as a segregated distribution of students could indeed be the result of preference heterogeneity.

In 2016, Chile implemented a reform that centralized school admissions and suppressed opportunities for schools to engage in student selection. Additionally, it introduced affirmative action measures to give an advantage to low-SES students when applying.[2] In this paper, we first study the effects of this policy change on two main outcomes, the level of school segregation and the quality of schools attended by different types of students. Furthermore, we use information on individual applications to estimate a demand model in order to understand the mechanisms that explain these effects. We use the model to simulate alternative policies that further increases the advantages for low-SES students in the school admission process.

In the first part of the empirical analysis, we take advantage of the sequential implementation of the reform to estimate its effects in a difference-in-differences framework. We look at measures of segregation and average academic quality across socioeconomic types. Our results suggest that contrary to expectations, the reform did not lead low-SES students to enroll in privately managed schools that might have previously discriminated against them. In order to assess the effects on the access to better quality schools, we construct a time-invariant measure of school quality using test scores over several years prior to the reform, controlling for student socioeconomic characteristics. We then use the same difference-in-difference approach to investigate whether school quality has increased for low SES students after the reform. Our results indicate that school quality increases significantly, but only for the subgroup of low-SES students that apply to schools and are assigned

[1] In the case of Chile, some studies find evidence of such behavior which in part motivates the reform studied in this paper (Gazmuri, 2018, Hsieh and Urquiola, 2006)

[2] Chile has a nationwide school choice system, where subsidies are the same for public and voucher schools. The new system affects admission to all public and voucher schools. Private schools which serve about 7% of the population keep having decentralized admissions.

in the first round of the application process. A large part of students remain unassigned after this first round and then participate in a second round with underdemanded schools. The high number of unassigned students is mostly explained by the fact that many of them do not rank enough schools in their applications.

In order to investigate the mechanisms behind these results, we estimate a demand model to recover students' preferences for schools. As admissions are now centralized, (almost) all the responses from the policy can be attributed to student demand.[3] The reform could affect student allocation in several ways. First, a direct effect of the preferential access for low-SES students[4]: if preferences for schools were homogeneous across students, low-SES should displace some medium- and high-SES students from good schools. Second, an indirect effect: if students have a taste for peers of the same socioeconomic background, a shift in the distribution of low-SES students will alter student preferences for schools in the following years and therefore future applications. A third effect is through the number of students assigned to a school they had not chosen. As the system is centralized, applicants need to submit a list of schools all at once. Some students that fail to rank enough schools are unassigned after the algorithm runs and will eventually get assigned a school they had not initially ranked in their application.

We use students rank-ordered lists, as well as students' home location and a measure of their SES to estimate an exploded logit model. We model students' choices taking into account the distance between a student's home and the school, the average percentage of low-SES in the school, and unobserved school quality. We allow for all the coefficients, including unobserved school quality, to vary with student SES category to understand preference heterogeneity across SES types.

Our first results indicate that there is indeed significant heterogeneity in preferences for schooling across SES. In particular, students of different SES levels, have vastly different perceptions of unobserved quality of schools after taking into account distance and type of peers. Furthermore, high- and medium- SES students have a clear distaste for higher proportions of low-SES students in the school, which is not the case for others. All of these results are consistent with the fact that low-SES students rank schools differently than the rest and partly explain, the lack of effect of the reform on student allocation.

[3] We say almost because private subsidized schools do have the possibility to respond by leaving the public system and becoming fully private, although this is relatively marginal.

[4] 15% of seats in each school are reserved for disadvantaged students. See section 1.3.2 for a description of the reform.

In order to investigate potential ways to reduce school segregation when accounting for preference heterogeneity, we study the counterfactual scenario of a change in policy design resulting in an increase in the level of affirmative action. To do so, we first replicate the allocation mechanism, and then use the parameters of our demand estimation to simulate a counterfactual allocation under a different level of affirmative action over several years. Results indicate that even if the advantage to low-SES students in the applications drastically increased, the reform would have little effect on student sorting. This is the result of heterogeneous preferences leading low-SES to under-apply to high-quality schools.

In Section 2 we review the existing literature on the several topics that our work relates to. Section 3 contextualize the analysis and detail the specifics of the reform at study. We present the data in Section 4 and show a few relevant statistics. Section 5 shows the estimation of the reform effects on student sorting across private- and publicly-funded schools, and across schools according to their quality. Section 6 is dedicated to describing our demand model and estimation, of which we present preliminary results in Section 7. In Section 8 we estimate the effect of the reform on within-school-type segregation. In Section 9 we discuss next steps of our research, before concluding in Section 10.

1.2 Literature review

This paper contributes to the broad literature on school choice. First, our work relates to papers about cream-skimming (McEwan, Urquiola, Vegas, Fernandes, and Gallego (2008)). Second, our work also relates to papers evaluating school choice programs, highlighting the importance of competitive forces through schools' responses (Epple and Romano (1998), Rothstein (2006)), and pointing out the sorting patterns that appear as a result (Hsieh and Urquiola (2006), Hanushek and W ößmann (2006)).

We also contribute to the literature on the estimation of students' preferences for schooling. Abdulkadiroglu, Pathak, Schellenberg, and Walters (2017) studies how preferences relate to value-added measures. Similarly to Hastings, Kane, and Staiger (2009), Gazmuri (2018), and Böhlmark, Holmlund, and Lindahl (2016), we are interested in preference heterogeneity for observed school characteristics. Gazmuri (2018) uses an earlier reform to estimate the differences in preferences for schooling across socioeconomic backgrounds. The 2016 reform allows for a better identification of preferences without concerns for school selection. Our focus is on evaluating the impact of centralizing admissions, and specifically on the effects of how priority rules affect student allocation.

Agarwal and Somaini (2018)) estimates demand when students have incentives to behave strategically.

Moreover, this paper adds to the growing literature on the evaluation of assignment mechanisms, such as Abdulkadiroğlu, Agarwal, and Pathak (2017), who compares assigment mechanisms in NYC public high schools. Specifically, our work also relates to recent papers studying affirmative action measures in centralized admission mechanisms, such as Barahona, Dobbin, and Otero (2023), Concha-Arriagada (2023), Tincani, Kosse, and Miglino (2023), Idoux (2022) and Escobar and Huerta (2021). Our paper differs in that it studies a reform where school preferences over students are ignored and replaced by a lottery[5]. As this was widely advertised, it could have impacted application behavior and calls for different methods to model and estimate demand for schooling. Additionally, few other contexts than offer the possibility to study admission mechanisms for a student's first entry into the school system, as we are not aware of another country where kindergarten slots allocated through a national centralized mechanism.

1.3 Institutional background

1.3.1 The Chilean education system

Chile has nationwide school choice system where public and private subsidized schools coexist and students are free to choose where to enroll. Both publicly or privately-managed schools receive a per-student voucher from the government. There is also a third type of fully private non-subsidized schools that enroll around 7% of students.

In summary, Chilean schools are divided in three categories:

- Public schools, which are managed by the local municipality and receive public funding in the form of a per-student voucher. They are tuition-free.
- Voucher schools, which are privately-managed but funded through the same public voucher as public schools. They can ask for a top-up tuition as a complement with restrictions for low-SES students. Low tuition (see below).
- Private schools, which are privately-managed and fund themselves entirely through tuition, which is high (see below).

Table 1.1 displays students' enrollment shares in the 3 types of schools in the country, although there is significant variation across regions.

[5]See section 1.3.2 for more detail on the specifics of the lottery.

The vouchers the government pays to the schools depend on the socioeconomic level of the student. Students are classified into three categories depending on parental income, and participation of some government social programs.

- Low-SES or first-label: household income in the lowest 40% of the total distribution of income or if the household is part of any of the program dedicated to socioeconomically disadvantaged households[6]. Participation on these programs is based on household characteristics flagging vulnerability.
- Medium-SES or second-label: household income is above the 40% but below the 80% threshold in the income distribution.
- High-SES or no-label: household income above the 80% threshold.

Appendix Table A.1 shows the distribution of tuition for voucher schools, which is small for the large majority of schools (median monthly income in Chile was 1,150 USD in 2020). Additionally, even schools that have some top-up tuition are not allowed to charge this to students in the low-SES categories.

School is compulsory from the first grade of primary school in Chile. However, most children enter the school system two years before, in pre-K. Many schools in Chile offer grades from pre-K to the end of secondary school. We focus on entry to Pre-K, when most students first enter the school system. Additionally, the choice of a school in Pre-K will be determinant for future educational outcomes.

1.3.2 The 2016 reform of the admission process

The School Inclusion Law program (*Ley de Inclusión Escolar*) was implemented in 2016 with three main objectives: centralize the admission system to prevent schools from selecting students, suppress tuition for all private-subsidized schools; switch the legal status of subsidized schools to not-for-profit. The reform is only aimed at public and voucher schools — private schools are not directly under its scope.

Previous to this reform, the admission system was fully decentralized: students had to apply to each school separately, and schools did not have to disclose the basis for their selection.

[6]Protección Social Chile Solidario, Programa de Ingreso Ético Familiar, Subsistema Seguridades y Oportunidades

The new process centralizes applications, and allocate students to schools according to an algorithm that does not allow schools to express preferences. The matching process follows a deferred-acceptance algorithm with multiple tie-breaking rules. Students submit a rank-ordered list of public and voucher schools. Schools rank students according to a ruled determined centrally: students are divided in categories, and schools' preference over those categories are imposed by the policy. The student categories are the following, in order of priority:

1. Students who are already enrolled in the school. Students are always able to continue their studies in the school where they are enrolled if their grade is offered. (henceforth referred to as *Enrollment priority*)
2. Student who have a sibling enrolled in the school (henceforth *Sibling priority*).
3. Children of employees working at the school (*Worker priority*)
4. All applicants who do not fall in any of the aforementioned categories.

In each school, a quota of 15% of the slots follow a slightly modified ordering of students:

1. Enrollment priority students
2. Sibling priority students
3. Low-SES students (as defined by the first-label in section 3.2)
4. Worker priority students
5. All other students

This quota introduces some affirmative action in the mechanism. There are very few students who have enrollment priority in PreK (the grade we focus on), as this is the grade were most Chilean kids enter the school system. Therefore, if we abstract from Sibling priorities, should there be less low-SES applicants than 15% of the slots in school j, all low-SES applicants will be accepted.

In secondary school, schools are allowed to set up an entrance test, and applicants with test results in the top 20% of the distribution get some priority at admissions. Since we focus on Pre-K, this does not affect our analysis.

The application process has two rounds and the timeline goes as follows:

1. First round:

(a) Applicants submit rank-ordered lists of schools. The deferred-acceptance algorithm runs.

(b) Applicants receive their offers and have to answer. They can decide to either accept/reject the assignment, or they can save it until after the wait list runs.

(c) Rejected offers are freed and the wait list runs. Applicants receive their final first-round assignments.

2. Second round:

   (a) After first-round assignments, some seats remain unfilled. Those seats are offered during a second round. Participating students can be those who were left unassigned or dissatisfied after the first round, as well as new applicants that did not participate in the previous round. They submit a new ranked-ordered list of remaining schools and the deferred-acceptance algorithm runs.

   (b) Applicants receive their assignment.

   (c) Unassigned applicants are assigned a school with remaining empty seats through an algorithm that minimizes the distance between the student and the school.

Participation in the application process is not compulsory, but students who enter school for the first time, or those for whom their current school does not offer the next grade are likely to participate. If they do not participate, they have the option of applying in the second round but only to schools that have available seats or they can choose to enroll in a private school. Students that want to change the school in which they are currently enrolled also need to apply to new schools through the centralized process.
In order to avoid the selection of students who participate in the process because they want to change their school, our analysis focuses on the most common entry-level grade, pre-K.

The reform was promulgated in 2015 and then implemented sequentially on the territory. The timeline of the implementation is as follows:

1. For the academic year 2017[7], the new application process was launched in a pilot region[8]. We refer to this region as *the first wave* region.

2. For the 2018 academic year, the reform was extended to 4 more regions.[9] We refer to these 4 regions as *the second wave* ones.

[7] The school year starts in March and ends in December.
[8] Magallanes, the country's most scarcely populated region.
[9] Coquimbo, O'Higgins, Los Lagos, Tarapaca

3. For the 2019 academic year, the reform was implemented in the entire country except for the region of the capital city, Santiago. We refer to all the regions where the reform was first implemented in 2019 as the *third wave* regions.

4. For the 2020 academic year the reform was implemented in the entire country. The region of Santiago is referred to as the *fourth wave* region.

1.4 Data and descriptives

Our data comes from the Chilean Ministry of Education. We use data from the application process, as well as census information on student enrollment, and schools' teachers and tuition.

The applications dataset gives us access to information on schools, applicants and the matching process itself. The latter contains students' rank-ordered lists, school-specific priority and lottery numbers. Other information includes geographic coordinates of both students and schools, school capacities, seats available, school status as well as students' gender and family relations with other applicants. We have access to this information for both rounds of the process, from the first set of applications to 2020.

The census dataset displays all enrolled students in the country, from 2014 to 2020. It provides information on their gender, age, priority label[10], as well as the school they are enrolled in and the course attended. On the school side, we observe each school's type (public, voucher or private), its location, as well as the pool of teachers - on which we observe age, gender, grades taught, experience - and tuition.

The information on students' priority labels allows us to divide the student population in 3 categories: low, middle and high SES. Low SES students roughly belong to the poorest 40% of the income distribution and high SES to the richest 20%. We use this categorization to study socioeconomic segregation throughout the paper.

Table 1.2 shows the distribution of students of each SES category across school types for the year prior to the implementation of the reform. As expected, low and medium SES students are very rarely enrolled in private schools given the high levels of tuition in those schools. We also see that low-SES students are more likely to be enrolled in public schools than others, a difference that is

[10] as defined in section 3.2

unlikely to be explained by tuition alone (see section 3.2).

Table 1.3 shows the same distribution for each wave of the reform for the year before and the year after the implementation. Comparing the proportion of low SES students enrolled in a public school across years, we see that the reform does not seem to bring much of a change. This is at odds with the fact that the reform removes barriers to entry for low-SES students and increases their comparative chances to get into their favorite school.

Regarding the application behavior, there is some evidence that students rank schools differently across SES. Table 1.4 shows the distribution of students of each SES according to the type of school they ranked as their first choice. Higher-SES students tend to rank a voucher school first with higher probability.
These differences in ranking patterns are especially important to consider given that most students end up attending their first choice. Considering that the system encourages students to rank programs truthfully, these differences could be interpreted as reflecting differences in student preferences. Figure 1.1 shows the distribution of the rank in the list of the school a student is assigned after the first round.

1.5 Evaluation of the reform

We investigate the impact of the reform taking advantage of the implementation in waves. We focus on two different outcomes. First, socioeconomic segregation across school types, and second, the quality of schools attended by low-SES students.

1.5.1 Socioeconomic segregation across school types

As a measure of school-type socioeconomic segregation, we use the proportion of low-SES students enrolled in voucher schools compared to the proportion of low-SES students in the municipality. The measure is constructed at the municipality level, as follows:

$$S_{m,t} \equiv \frac{\%\text{ low SES students enrolled in voucher schools}_{mt}}{\%\text{ low SES students}_{mt}}$$

It measures the propensity of low-SES students to enroll in voucher schools. The higher the ratio, the more prone low-SES students are to enroll in a voucher school.

As mentioned above, low SES students have been traditionally more likely to enroll in public schools, which has been partly attributed to being selected out of voucher schools. Thus, we would expect a shift of low SES students towards voucher schools after the reform and therefore for this ratio to increase.

Table 1.5 shows some statistics about the distribution of S_{mt} where regions are aggregated by wave of implementation of the reform. As expected, most municipalities have $S_{mt} < 1$, meaning that low SES students tend to enroll in public schools more often than in voucher schools.

We run the following specification to study the impact of the reform on S_{mt}. We make use of the fact that the reform was implemented sequentially in different regions to set up a difference-in-differences framework:

$$S_{m,t} = \alpha + \beta \, \text{Reform}_t + \delta_t + \xi_r + \varepsilon_{m,t} \tag{1.1}$$

$$\Delta S_{m,t} = \alpha + \beta \, \Delta\text{Reform}_t + \omega_t + \eta_{m,t} \tag{1.2}$$

with $\omega_t \equiv \Delta\delta_t$ and $\eta_{m,t} \equiv \Delta\varepsilon_{m,t}$. Results are shown in Table 1.6. Column (1) corresponds to equation (1) without region fixed-effects. The negative coefficient on the reform suggests that regions treated first tend to have lower average values for the ratio (fewer proportion of low-SES students in voucher schools). Column (2) includes region fixed-effects and Column (3) corresponds to the difference-in-differences specification. We do not find any significant effect on this measure of segregation.

1.5.2 School quality

Even if the reform did not increase the probability of low-SES students enrolling in voucher schools, there could still be an effect on the quality of schools low-SES students are accessing.
In order to test whether low-SES students are ranking better schools given the new priorities established by the program, we need to construct a measure of school quality. We want to use a time-invariant measure, constructed before the policy was implemented, to abstract from posible quality differences associated to the policy.
We construct a measure of school quality following Abdulkadiroğlu, Pathak, Schellenberg, and Walters (2020) estimating value-added measures using students' standardize test scores controlling for a rich set of individual characteristics. We include parental education and income, student

gender, whether the child has repeated a grade, and whether the student is categorized as low-SES. Using this dataset, we run the following regression on observations between 2010 and 2015 (before the implementation of the reform):

$$\text{Score}_{i,j} = X_i + \delta_t + \xi_j + \varepsilon_{ij} \tag{1.3}$$

where $\text{Score}_{i,j}$ is the score of student i in school j. X_i is a set of observed student characteristics, and δ_t and ξ_j are respectively time and school fixed-effects.

We use the estimates of ξ_j from equation 1.3 as a measure of time-invariant school quality. Quality estimates are robust to alternative specifications and time periods.
Figure 1.2 shows the distribution of school quality for each type of school. It roughly coincides with general perception of quality differences, with private school significantly above. Voucher schools on average better than public schools, but both types showing significant heterogeneity and overlap in the distribution.

We use these measures to investigate the reform's impact on the average school quality attended by each of the 3 student SES categories. Again, our identification strategy takes advantage of the sequential implementation of the reform, and run the following regression:

$$\text{School quality}_{i,t} = \alpha + \beta \text{Reform}_{r(i),t} + \delta_t + \xi_{m(i)} + \varepsilon_{i,t}$$

where $\text{School quality}_{i,t}$ is the ξ_j for the school attended by student i in year t, in region r and municipality m. We run three separate regressions, one for each student SES category.
Table 1.7 shows these results. It seems like the reform had a small but significant positive effect on the school quality for low and high SES students and a non-significant effect for medium SES.

Using the student final assignment confounds several possible effects that could be affecting these results. First, a student has a choice of participation, by submitting a rank-ordered list in the platform at the correct time. If the student participates, but ranks an insufficient number of alternatives, she could remain unassigned after the first round. Students not participating in the first round or failing to submit a sufficiently long list, would not be able to take advantage of the priorities from the new program.

In order to better understand the heterogeneity of the impact across students, we run a similar regression adding information on whether the student participated to the applications, and on which specific round she participated in. Results can be found in Table 1.8. The coefficient on the variable *treat* displays the impact of the reform for students in treated regions that decided not to participate in the applications. The impact of the reform on participants assigned after one round, is the sum of the coefficient of *treat* and that of *round 1 only.* Students participating in both rounds, are students that submitted a rank-order list, but not sufficiently long so they were not assigned in the first round and need to apply again in the second round. The effect for them is the sum of the coefficients of *treat* and *both rounds.* Same for the ones that only submitted applications in the second round, the effect would be the sum of *treat* and *round 2 only.*

Results suggest that the reform indeed had a large positive effect for low-SES students that participate in the first round of applications. These are the students that are effectively taking advantage from the priority system created in the reform. The overall small impact for low SES students is mostly explained by those who did not apply, did not rank a sufficiently large number of schools, or apply only during the second round.

The effect for high SES students is negative no matter their application decisions, except for those who did not participate to the applications. This is explained by a big share of these students choosing to enroll in private schools, which are usually of higher quality. An effect of the reform for this group of students is to increase the proportion enrolling in private schools. The next three columns in Table 1.8. show the same regression withoutstudents enrolled in private schools. High-SES students that submit applications and get enrolled in the first round, get a positive overall effect.

This analysis shows that the reform has heterogeneous impacts on school quality for students of different SES, and who made different application decisions. Overall, the results for students assigned within the first round of applications seem to be in line with the objectives of the new policy.

One mechanism highlighted by these results is that students tend to rank too few schools. Figure 1.3 shows the proportion of students by SES in each possible enrollment path, showing that only around 40% of students are effectively enrolling to a school assigned in the first round of applications. Next, we show some descriptives to shed some light on this issue.

1.5.3 Unassigned students

Panel A of Table 1.9 shows the percentage of unassigned students by SES among the ones that participated in the first round of applications. In all waves, this increases over time. Panel B, shows the average number of schools ranked by SES among assigned and unassigned students. As expected, unassigned students tend to rank fewer programs than assigned students. Furthermore, low-SES students rank fewer schools. In this table we focus on students without sibling priority, as these tend to rank very few, and are likely to be assigned to the sibling's school.

Table 1.10 shows a regression of an indicator for being unassigned on student SES and number of ranked programs. As expected, the probability of being unassigned decreases as the number of ranked schools increase. Low-SES students are less likely to be unassigned, and the effect of increasing the number of ranked schools is less severe for them. This is expected given the priorities from the program.

1.6 Demand estimation

In order to understand the mechanisms that explain the previous results and to be able to predict allocations under alternative policies we use the information on individual rank order list to estimate a student demand model.

1.6.1 Model

We assume a linear utility function with observable and unobservable school characteristics. Students' tastes for school characteristics are allowed to vary according to their SES category (First, Second or no label). The utility of student i from school j in year t is assumed to take the following form:

$$\begin{aligned}\mathcal{U}_{ijt} = \alpha + \gamma_i \text{D}_{ij} + \phi_i \text{Perc. low SES}_{jt} + \delta_j^0 + \delta_j^p \mathbb{1}\{SES_i = p\} + \beta_i^S \text{Sibling}_{ij} \\ + \beta_i^W \text{Worker}_{ij} + \tilde{\mu}_{ij} + \epsilon_{ijt}\end{aligned} \tag{1.4}$$

where D_{ij} is the distance between student i and school j, and Perc. low SES_j is the percentage of low-SES students in school j, Sibling_{ij} takes the value one if student i has sibling priority in school j, similarly for Worker_{ij} if the student has a parent working in the school. p is the student SES. All coefficients are given the flexibility to vary by student SES, such that:

$$\gamma_i = \gamma^0 + \gamma^p \mathbb{1}\{SES_i = p\}$$

$$\phi_i = \phi^0 + \phi^p \mathbb{1}\{SES_i = p\}$$

$$\beta_i = \beta^0 + \beta^p \mathbb{1}\{SES_i = p\}$$

The unobserved idiosyncratic component is identically and independently distributed, $\epsilon_{ij} \sim T1EV$. $\mathcal{I}$ is the set of all students and $\mathcal{S}$ that of all available schools on the market. A market is defined as a year and a region.

In their applications, students are assumed to rank schools in order of their preference. We do not explicitly model the decision to stop ranking, and therefore the size of a rank order list. While we are aware this is a limitation to our counterfactual analysis, it avoids assuming that students stop ranking when they prefer the outside option to the next school, which does not seem plausible in a context where students who are unassigned after the first round of the allocations do participate in the second round with schools they had not listed initially. Counterfactual simulations will take list size as observed - which means our results are unbiased under the condition that the counterfactual change does not affect the amount of schools listed. Given the relatively low variation in list size observed in the data - 76% of students rank between 2 and 4 schools -, one could think that students may have a set list size in mind and seldom deviate from it.

When ranking schools in their application, applicants are assumed to be truth-telling, meaning that they rank schools in the order of their genuine preference. This is a standard assumption when the assignment algorithm is deferred-acceptance, since this is a strategy-proof mechanism. In practice, if an applicant's probability of being assigned to a given school is negligible, they could fail to rank it even if it is preferred to some other ranked schools - a behavior called "skipping the impossible" (Fack, Grenet, and He (2019),Haeringer and Klijn (2009)). In order to assess whether some schools are so oversubscribed that the probability of admission is negligible, we study the distribution of the following metric:

$$P_j^r = \frac{\text{\# applicants who ranked school } j \text{ in first } r \text{ choices}}{\text{\# seats available in } j}$$

Table 1.11 looks at the distribution of this variable for different values of r. It shows that few schools receive significantly more applications than seats available, and almost none have a probability of entry close to zero for P_j^2. This suggests that very few schools should be skipped based on a negligible probability of entry. Furthermore, we had already shown in Figure 1.1 the distribution of the position in the rank-ordered list of the assigned school after the first round of the algorithm, showing that the large majority of applicants get assigned their first or second choice. In light of this information, P_j^2 seems to be the relevant measure to assess the distribution of probabilities of being assigned.

There is a few of concerns related to this measure of the risk of skipping-the-impossible behavior. First, P_j^r is only informative about the actual probability of entry, not the belief parents have when submitting their rank. If a school seems really inaccessible, parents may not list it, even if in reality it is indeed accessible. Second, the probability of entering a school is different for low-SES students compared to other students because of the relative advantage they have in the lottery. Nevertheless, this metric in a useful approximation to give an idea of how oversubscribed schools are in reality.

1.6.2 Estimation

We use the information from students' rank-ordered lists to estimate the model. We estimate our model by maximum likelihood to recover

$$\theta \equiv \{\delta_j^0, \delta_j^p, \gamma^0, \gamma^p, \beta^0, \beta^p, \phi^0, \phi^p, \sigma_\gamma, \sigma_\phi\}$$

We make use of the distributional assumption on ε_{ij} to write the probability for student i to rank school j first as follows:

$$P_{ij} = \frac{exp(V_{ij})}{\sum_{k \in \mathcal{S}} exp(V_{ik})}$$

where $\mathcal{S}$ is the set of schools in student i choice set.[11] Then, the probability for student i to rank school k second, knowing j is first is given by:

$$P_{ik}^2 = \frac{exp(V_{ik})}{\sum_{l \in \mathcal{S} \setminus \{j\}} exp(V_{il})}$$

More generally, we model the probability for i to rank school m in the r^{th} position as:

$$P_{im}^r = \frac{exp(V_{im})}{\sum_{l \in \mathcal{S}^r} exp(V_{il})}$$

The exploded-logit formulation of the likelihood consists in R parts, with R the number of ranked schools considered per student. Although students are allowed to rank up to 10 schools, they rarely rank more than 3 and thus we use only the first 3 ranks for estimation (R=3).

[11] Given that some schools are girls-only or boys-only, these schools would not belong to the choice set for a student of the opposite gender.

The contribution to the likelihood for student i if l_i and is given by:

$$l_i = \sum_{j \in \mathcal{S}^2} \sum_{k \in \mathcal{S}^2} \cdots \sum_{l \in \mathcal{S}^R} \left[(P^1_{ij} \times d^1_{ij}) \times (P^2_{ik} \times d^2_{ik}) \times \cdots \times (P^R_{il} \times d^R_{il}) \right]$$

where d^r_{ij} is a dummy taking the value one if i indeed ranked j in the r^{th} position and S^r is the set of all schools except the ones ranked 1^{st} to $(r-1)^{th}$.

1.6.3 Results

Current results use only 2 SES categories (Low SES, others), and do not include the random coefficient component $\tilde{\mu}_{ij}$.

Figure 1.4 shows the distribution of mean values for vouchers schools across SES. The blue bars display the distribution of δ^0_j for voucher schools, and the red bars the distribution of estimates of $\delta^0_j + \delta^p_j$, where p= Low SES here. Low SES students seem to discriminate less among schools than other students.

Table 1.12 below shows the estimates for $\beta^0, \beta^p, \gamma^0$, γ^p, ϕ^0, ϕ^p. Interestingly, contrary to what is mostly found in the literature, low-SES students appear to have less disutility from distance than others. This could come from their internalizing that they have a lot to benefit from this new mechanism, making them more ready to travel.

High-SES students seem to have a stronger disutility for schools with a high proportion of low-SES, again consistently with the literature.

Interestingly, low-SES students seem to avoid schools where they have a sibling. This could be explained by the new priorities if the older siblings were enrolled in a less preferred school because they did not have access to better schools.

1.7 Model Validation

We validate our model by simulating rank-order lists and running the algorithm using observed lottery numbers. To do so, we first develop a replicating algorithm which we test using observed rank-order lists and lottery numbers. We manage to replicate 99.5% of the real allocation.

The way in which the quota level enters the algorithm occurs as follows:

- Each course is split into two sub-courses, one of which is stamped "low-SES".

- In the low-SES sub-course, a low-SES applicant is always preferred to any other, except those who have a sibling in the course.[12]

- The low-SES sub-course has a capacity corresponding to 15% of the total capacity of the course [13].

Results of the matching from simulated rank-order lists evidences that the main features or the current sorting are well captures by our model. Appendix Figure A.1 shows that the distribution of the proportion of low-SES across courses is reasonably well replicated, although tails are thicker in the replication than in reality and the center of the distribution in the simulations is left-biased. In this figure, as in the following, where we simulate rank-order lists, we draw show several simulations that result from various draws of the shocks. In all figures, simulations are reassuringly very similar. Appendix Figure A.2 shows that the simulations replicates very well the number of students in voucher schools, and the distribution of those students across SES. As for the distribution of students across school qualities, Appendix figure A.4 shows our simulations are reasonably accurate, albeit imperfect. However, Appendix Figure A.3 shows the simulations leave too many students unassigned compared to reality.

1.8 Counterfactuals

The objective of this exercise is to understand the potential mechanisms through which socioeconomic segregation could be significantly mitigated through a centralized mechanism. In the following counterfactual analysis, we simulate first-round matches[14] under modified mechanism rules or applicant behavior. In our first set of counterfactuals, we increase the percentage of seats for which low-SES students have priority from 15 to 40% of the total available slots. We implement the change in quota for the region of Coquimbo, which is a relatively big region where the reform was implemented in 2018. We set the new quota to be 40% and run counterfactual allocations for 2018 using a matching algorithm that reproduces closely the one used in reality. Again, we need to take into account changes in the peer group to compute the new school peer compositions.

In a second set of counterfactuals, we compare the match in the cases where students rank alternatively 3 or 10 schools. This allows us to see the advantages from students listing more schools, mainly that less of them are unassigned and that the distribution of low-SES students across

[12]Applicants who used to be enrolled in the school more than a year ago or those who have a parent working in the school are also preferred to low-SES applicants, but those are very rare.

[13]The number of low-SES students who were enrolled in the preceding grade and do not participate in the applications is subtracted from the 15% of slots dediated for low-SES in the applications. However, since we are looking at the first grade of entrance to school, very few children are enrolled prior to the applications.

[14]See section 1.3.2 for a description of first and second round of the allocation algorithm.

schools is more homogeneous.

1.8.1 Change quota level

Altering the quota level directly changes the capacities for each sub-course. When increasing the quota level from 15 to 40%, we effectively increase the number of vacancies in the sub-course where a low-SES applicant is preferred.

In our analysis, we focus on the region of Coquimbo, which is part of the second wave of implementation of the reform, leaving us some perspective, and is relatively densely populated. We choose not to focus on the pilot region since it is in many ways unrepresentative of the rest of the country. The counterfactuals are simulated for two consecutive years. First, we change quota levels from 15 to 40% in 2018, the first year of implementation of the reform in the region. We simulate the match using observed rank-order lists and modified sub-course capacity as some seats from the regular sub-course are transferred to the low-SES sub-course. This leads students to allocate differently across schools, which can affect the proportion of low-SES students in each school. This will in turn affect student preferences and rankings in 2019, as they will rank schools based on the counterfactual percentage of low-SES students in the previous year. We therefore simulate students' rankings in 2019 and run the matching algorithm using those simulated applications.

Direct effects through changes in admission chances

Figure 1.5 shows the percentage of courses with less that 15% low-SES (top figure) and less than 40% low-SES (bottom figure). Unexpectedly, the proportion of courses with less than 15% low-SES students increases under the higher quota. Figure 1.6 compares the distribution of the proportion of low-SES students per course under the true 15% quota and the counterfactual one. Because low-SES students prefer different schools than high-SES, when increasing the quota the non-preferred schools end up with even fewer low-SES students. As a result, the proportion of courses with less than 40% of low-SES is lower under the counterfactual scenario. This is consistent with the idea that under the 15% quota a significant proportion of unlucky low-SES students are getting turned down in schools with too many low-SES applicants.

Figure **??** shows that under the counterfactual, slightly more low-SES students are getting assigned to a voucher school. Figure **??** also show that increasing the quota slightly increases the quality of the assigned school for low-SES students, and decreases it for high-SES students.[15] However, this

[15] Here we use the same measure of school quality as in section 1.5.2. That is the schools value-added in a regression of school grades on student characteristics using results of a national exam.

effect is small compared to the level of the increase, and suggests there is some inertia in students choices. This could be explained by the preference heterogeneity for peers across SES-types. High-SES students are slightly penalized, they are less likely to attend a voucher school and have a small decrease in school quality on the aggregate. However, these graphs investigate assigned schools from the allocation mechanisms, and therefore do not take into account the fact that a portion of high-SES students will reject their assignment and enroll in a private school (this is outside of the scope of the model).

Additionally, increasing the quota to 40% slightly increases the quantity of unassigned students - from 12.9% to 13.5% of applicants. Figure 1.7 compares the composition of the pool of unassigned students under the two scenarios: under the counterfactual, low-SES are less likely to end up unassigned. This is expected since the number of slots for which they have priority increases, making high-SES students more likely to be rejected from oversubscribed schools.

Indirect Effect through Changes in Distribution of Students

As mentioned above, changing the low-SES quota will affect the distribution of students across schools, which in turn affects how students rank schools through the peer-preference parameter identified in the demand estimation. Therefore, changing the quota levels in 2018 will have effects in 2019, even if the 2019 quota were back to 15%. We thus expect the impact of a quota change for a single year to be able to affect allocations over several years. To analyze these long-term effects, we simulate student rank-ordered lists[16] for 2019 using the counterfactual distribution of low-SES students from 2018 and the parameter estimates from our demand estimation.
We feed the simulated preferences together with modified sub-course capacities reflecting the 40% quota into our matching algorithm. We take the list of applicants and courses as observed in 2019. Capacities for each course are assumed to be the same as those stated in the real 2019 applications - although we change the allocation of those slots across sub-courses[17]. We assume all students rank the same number of schools in the simulations as they do in their observed ranks in order to abstract from the effect of changes in the number of choices ranked.

For each simulation of rank-ordered lists we draw one shock ε_{ij} per student-course pair. In order to abstract from possible effects because of the specific shock, we run several versions of the allocation algorithm for each counterfactual. Each version uses the resulting simulated rank-ordered lists from a given draw of shocks. When graphing the results, we display several simulations of the allocation.

[16] See section 1.8.1 for a description of rank-ordered lists.
[17] See section 1.7 for a definition of sub-courses.

Results are shown in figures 1.8 to 1.11. Under the 40% quota, the proportion of schools with very small and with very large percentage of SES students increases (the curve flattened). Additionally, it decreases the proportion of low-SES students in very low quality schools and increases the proportion of students in schools of medium quality.

1.8.2 Varying list sizes

To conduct this analysis, we simulate rank-order lists for students in 2019, imposing a list size for all students. We simulate first-round matches using observed lottery numbers and school capacities. In our first scenario, students rank 3 schools, and in the second they rank 10 schools. We run multiple simulations for both scenarios, each with a different draw of the iid shocks to student utility.

Results are displayed in figures 1.12 to 1.14. Figure 1.12 shows that when students rank 10 schools, the number of them who remain unassigned after the first round drops drastically compared to the case where they rank only 3. Figure 1.13 compares the distribution of the proportion of low-SES students in the school across schools. It appears that when students rank more schools, the distribution has slimmer tails, indicating more homogeneity across schools. This is because many students get assigned a lower-ranked choice, which alleviates the sorting coming from heterogeneous preferences. Figure 1.14 shows that, among assigned students after the first round, a smaller proportion is assigned to top-quality schools when students rank more schools. This is because a lot more students get assigned, and since top-quality schools are most likely to be filled in the 3-ranks scenario, proportionally less assigned students are in those schools in the 10-ranks case.

1.9 Conclusion

This paper assesses the impact of affirmative action in a newly centralized school admissions system on socioeconomic segregation and educational quality in Pre-K. Our evaluation of the reform showed it had little effect on student sorting. Estimating a demand model revealed significant preference heterogeneity for school characteristics, with high- and medium-SES students preferring schools with fewer low-SES peers. Counterfactual simulations show that increasing the level of affirmative action would not significantly decrease student sorting because of the inertia coming from preference heterogeneity. Additionally, we show that students who rank less schools are a lot worse off from the reform. Counterfactual analysis comparing outcomes resulting from various imposed school list sizes show that longer lists significantly reduce the number of unassigned students in the first round. This means less students going to the second round, where only underdemanded schools of relative low quality are available. We also show that longer rankings lead to a more homogeneous distribution of students across schools, which sugggest that incentivizing students to rank more schools could be a tool to reduce socioeconomic segregation. Future research should investigate the mechanisms behind the choice of a list size in order to efficiently design policy to make them longer.

Tables and figures

Table 1.1: Distribution of students across schools

	2015	**2016**	**2017**	**2018**
Public	35.59%	35.14%	35.2%	35.29%
Voucher	55.54%	55.4%	54.94%	53.81%
Private	8.86%	9.47%	9.86%	10.89%

Table 1.2: Enrollment in 2016

	school type			
SES	private	public	voucher	Total
Low	1%	44%	55%	100%
Medium	2%	35%	64%	100%
High	28%	20%	52%	100%
All	11%	34%	55%	100%
N	17,066	54,683	88,046	159,795

Table 1.3: Enrollment across School Types

	Before reform			After reform		
	Low SES (%)	Medium SES (%)	High SES (%)	Low SES (%)	Medium SES (%)	High SES (%)
Panel A: wave 1						
private	2.183	3.957	30.51	0.484	4.724	21.85
public	63.10	55.40	37.80	73.12	55.91	44.94
voucher	34.72	40.65	31.69	26.39	39.37	33.21
Total	100	100	100	100	100	100
Panel B: wave 2						
private	0.385	1.472	21.28	0.608	1.734	24.84
public	52.73	32.74	18.93	52.13	36.96	22.33
voucher	46.89	65.79	59.79	47.26	61.31	52.83
Total	100	100	100	100	100	100
Panel C: wave 3						
private	0.478	2.137	23.80	0.594	2.483	24.74
public	53.60	38.61	24.39	53.59	38.89	25.75
voucher	45.92	59.25	51.81	45.82	58.63	49.52
Total	100	100	100	100	100	100
N	504	278	590	413	254	801

Table comparing enrollment in the regions belonging to a given wave of the reform for the year before and the first year of implementation.

Numbers are percentage of the total population of this SES category.

Wave 1: Magallanes, wave 2: Coquimbo, Los Lagos, Tarapaca, O'Higgins, wave 3: all other regions but Santiago

Table 1.4: Type of the first-ranked school, 2019, all regions except Santiago

SES	school type		
	public	voucher	Total
	row %	row %	row %
Low	48.8	51.2	100.0
Medium	36.9	63.1	100.0
High	24.5	75.5	100.0
Total	41.0	59.0	100.0
N	112,877	162,113	274,990

Table 1.5: Distribution of S_{mt} by wave-aggregated regions, country

	Mean				p1				p50				p99			
	wave 1	wave 2	wave 3	wave 4	wave 1	wave 2	wave 3	wave 4	wave 1	wave 2	wave 3	wave 4	wave 1	wave 2	wave 3	wave 4
2014	0.751	0.942	0.915	1.237	0.441	0.661	0	0.693	0.821	0.945	0.941	0.975	0.991	1.288	1.223	4.769
2015	0.862	0.914	0.940	1.213	0.721	0.512	0.528	0.698	0.810	0.922	0.938	0.978	1.054	1.319	1.333	5.300
2016	0.833	0.915	0.924	1.213	0.592	0.350	0.389	0.655	0.865	0.939	0.936	0.980	1.042	1.374	1.198	3.783
2017	0.769	0.902	0.922	1.156	0.722	0.574	0.400	0.705	0.750	0.905	0.940	0.966	0.836	1.429	1.344	4.079
2018	0.853	0.890	0.904	1.234	0.710	0.357	0.371	0.674	0.800	0.903	0.918	0.975	1.048	1.323	1.334	4.043
2019	0.872	0.878	0.917	1.238	0.782	0.158	0.470	0.615	0.881	0.904	0.921	0.985	0.952	1.205	1.466	3.923
# municip.	18	338	907	286												

Statistics of the distribution of $\frac{\text{\% prio in vouchers}_m}{\text{\% prio}_m}$ for each municipalities in an area of implementation of the reform (wave 1, 2, 3 and 4)

Table 1.6: DiD on $\frac{\text{\% prio in vouchers}_m}{\text{\% prio}_m}$

	(1)	(2)	(3)
treat	-0.270***	-0.0218	
	(0.0441)	(0.0482)	
Δ treat			0.00382
			(0.0130)
2015	-0.00719	-0.00703	
	(0.0399)	(0.0380)	
2016	-0.00621	-0.00736	0.0105
	(0.0399)	(0.0380)	(0.0257)
2017	-0.0381	-0.0424	-0.0252
	(0.0399)	(0.0380)	(0.0237)
2018	0.0484	0.00134	0.0529**
	(0.0407)	(0.0390)	(0.0195)
2019	0.173***	0.0177	0.0218
	(0.0477)	(0.0478)	(0.0235)
_cons	1.068***	1.070***	-0.0135
	(0.0283)	(0.0270)	(0.0151)
Region FE	NO	YES	NO
Municipality FE	NO	NO	YES
Autocorrelation-robust SE	NO	NO	YES
N	1549	1549	1274

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

DiD estimation of the reform on $\frac{\text{\% prio in vouchers}_m}{\text{\% prio}_m}$ using municipalities from all of the country

All municipalities from a given region are treated at the same time,

and all regions from a given wave are treated at the same time by definition

Table 1.7: DiD on school quality, by student SES

	Low SES	Medium SES	High SES
treat	0.521***	-0.278	0.438**
	(0.124)	(0.205)	(0.148)
2015	-0.166		-0.368***
	(0.0906)		(0.102)
2016	0.0109		1.667***
	(0.0904)		(0.111)
2017	-0.805***	1.282***	2.515***
	(0.0905)	(0.175)	(0.113)
2018	-0.874***	0.259	1.407***
	(0.0970)	(0.169)	(0.113)
2019	-1.199***	0.243	1.107***
	(0.125)	(0.205)	(0.133)
_cons	-1.240***	2.711***	10.88***
	(0.0633)	(0.128)	(0.0744)
N	408391	97384	312045

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The label Medium SES (parental income between 40 and

80% of the distribution) appeared in 2016 only

DiD estimation of the reform on the added-value of enrolled school

One regression for all students of a given SES in the country

All students going to school in a given region are treated at the same time

Table 1.8: DiD on school quality by application behavior, by student SES

	All students			Without students in priv. schools		
	Low SES	Medium SES	High SES	Low SES	Medium SES	High SES
treat	-3.268***	0.197	6.072***	-3.865***	-2.905***	-5.430***
	(0.207)	(0.379)	(0.199)	(0.207)	(0.400)	(0.294)
round 1 only	5.053***	-0.00621	-8.082***	5.731***	3.475***	7.367***
	(0.200)	(0.359)	(0.213)	(0.200)	(0.382)	(0.296)
both rounds	2.183***	-3.579***	-10.72***	2.743***	-0.445	3.913***
	(0.394)	(0.525)	(0.478)	(0.394)	(0.542)	(0.528)
round 2 only	-0.297	-6.539***	-19.31***	0.342	-3.285***	-3.882***
	(0.345)	(0.667)	(0.593)	(0.344)	(0.676)	(0.616)
2015	-0.166		-0.373***	-0.166		-0.201
	(0.0905)		(0.102)	(0.0903)		(0.111)
2016	0.0119		1.678***	0.0170		1.089***
	(0.0903)		(0.111)	(0.0900)		(0.125)
2017	-0.804***	1.280***	2.548***	-0.709***	1.221***	1.854***
	(0.0904)	(0.175)	(0.112)	(0.0901)	(0.174)	(0.128)
2018	-0.873***	0.251	1.425***	-0.784***	0.128	0.462***
	(0.0969)	(0.169)	(0.113)	(0.0965)	(0.168)	(0.129)
2019	-1.275***	0.258	1.147***	-1.205***	0.0352	0.189
	(0.125)	(0.205)	(0.133)	(0.124)	(0.204)	(0.156)
_cons	-1.241***	2.710***	10.87***	-1.429***	2.446***	4.334***
	(0.0632)	(0.128)	(0.0741)	(0.0630)	(0.127)	(0.0813)
N	408391	97384	312045	406159	95686	225606

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

DiD estimation of the reform on the added-value of enrolled school

One regression for all students of a given SES in the country

All students going to school in a given region are treated at the same time

The label Medium SES (parental income between 40 and 80% of the distribution) appeared in 2016 for the 1st time

Variable *round 1 only* takes value 1 if the student only participated in round 1 of the applications

Similarly for *round 2 only* and *both rounds*

Variable *round 1 only* takes value 1 if the student only participated in round 1 of the applications

Columns 4-6 abstract from students who are enrolled in private schools

Table 1.9: Unassigned Students

Panel A: % of unassigned students
PreK, 1st round participants

	year		
wave	2018	2019	2020
1	13.1	12.2	16.9
2	7.9	10.7	11.1
3		10.5	11
4			12.5

Panel B: Average number of schools ranked
PreK, 1st round participants without sibling priority
By assignment status after round 1

		SES		
		low	medium	high
unassigned				
	sibling prio	1.94	2.36	2.28
	no sibling prio	2.87	3.15	3.27
not unassigned				
	sibling prio	2.47	2.66	2.68
	no sibling prio	3.08	3.55	3.87

Table 1.10: Probability of being unassigned, students without a sibling priority (OLS)

	(1)	(2)
prioritario	-0.130***	-0.187***
	(0.00234)	(0.00470)
preferente_alu	-0.0925***	-0.121***
	(0.00253)	(0.00511)
2018.agno	0.206***	0.203***
	(0.0180)	(0.0179)
2019.agno	0.202***	0.204***
	(0.0177)	(0.0176)
2020.agno	0.275***	0.278***
	(0.0176)	(0.0175)
2.wave	-0.0595***	-0.0627***
	(0.0125)	(0.0125)
3.wave	-0.0623***	-0.0652***
	(0.0123)	(0.0122)
4.wave	-0.0646***	-0.0578***
	(0.0123)	(0.0123)
2018.agno#2.wave	0.00903	0.0191
	(0.0181)	(0.0181)
2019.agno#2.wave	0.0657***	0.0697***
	(0.0179)	(0.0178)
2019.agno#3.wave	0.0667***	0.0685***
	(0.0175)	(0.0174)
number_ranked		-0.0228***
		(0.000801)
c.number_ranked#c.prioritario		0.0140***
		(0.00122)
c.number_ranked#c.preferente_alu		0.00683***
		(0.00124)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.11: Distribution of P_j^r

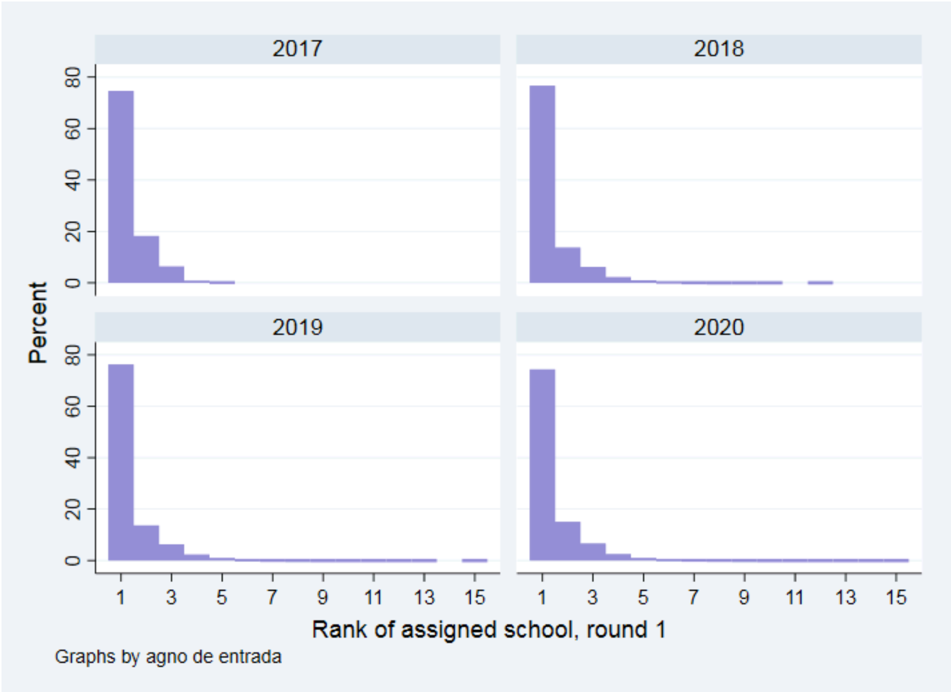
	Mean	P1	P25	P50	P75	P99
Panel A: P_j^2						
2017	74.35%	0.00%	5.13%	38.02%	75.00%	400%
2018	166.96%	0.00%	4.29%	66.67%	188.89%	1500%
2019	107.96%	0.00%	1.20%	38.00%	114.29%	1100%
2020	101.32%	0.00%	0.00%	34.48%	111.11%	1000%
Panel B: P_j^5						
2017	61.49%	0.00%	2.08%	21.67%	66.67%	400%
2018	116.85%	0.00%	3.28%	46.15%	128.57%	1000%
2019	87.24%	0.00%	0.73%	28.57%	89.56%	1000%
2020	77.69%	0.00%	0.00%	25.58%	82.14%	900%
Panel C: P_j^{all}						
2017	60.56%	0.00%	1.94%	20.47%	65.00%	400%
2018	109.12%	0.00%	3.03%	44.22%	121.74%	1000%
2019	83.88%	0.00%	0.71%	27.27%	86.21%	900%
2020	75.61%	0.00%	0.00%	24.74%	78.26%	900%

Statistics from the distribution of $\frac{\text{\# applicants who ranked school } j \text{ in first } r \text{ choices}}{\text{\# seats available in } j}$ across the country, $r \in \{2, 5, 10\}$

Table 1.12: Preliminary estimates of $\beta^0, \beta^p, \gamma^0$ and γ^p

	Main
Distance (km)	-0.4544
Distance*Low SES	0.0297
% low SES	-1.7117
% low SES *Low SES	0.8038
Sibling priority	0.0150
Sibling priority*Low SES	-0.3167
Worker priority	0.3429
Worker priority*Low SES	-1.2572
N	32435

Figure 1.1: Rank of school assigned after round 1



Only assigned students are included

Assigned students who rejected their offer are still included

Figure 1.2: Distribution of school quality measures across school type

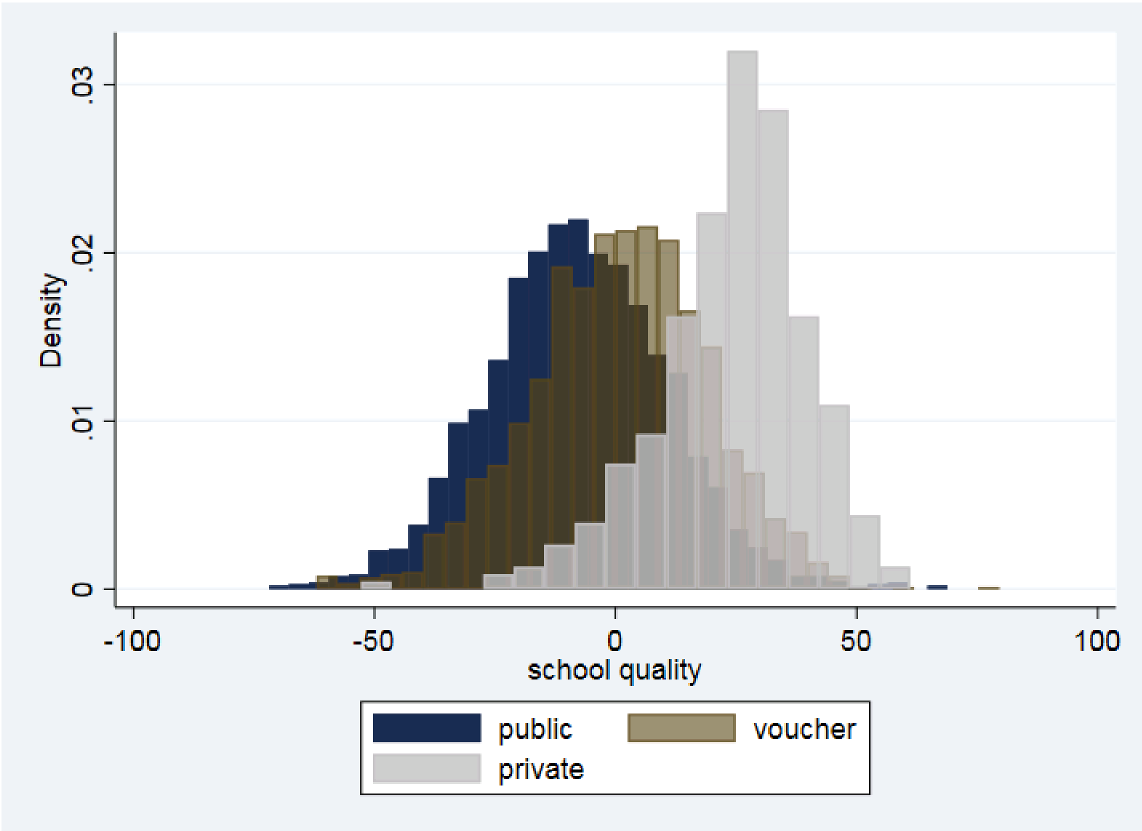
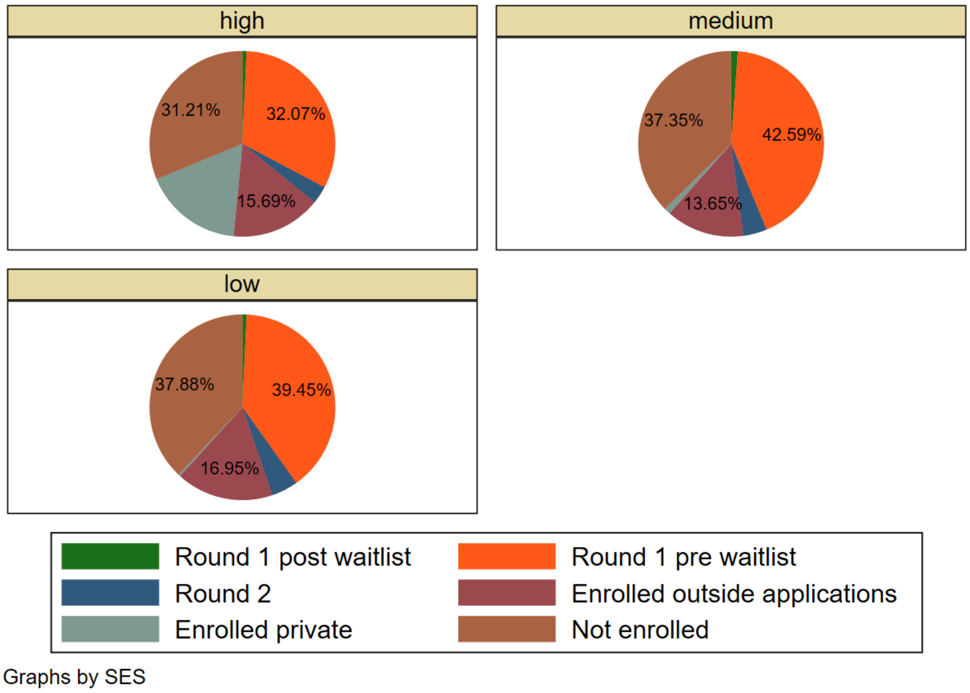
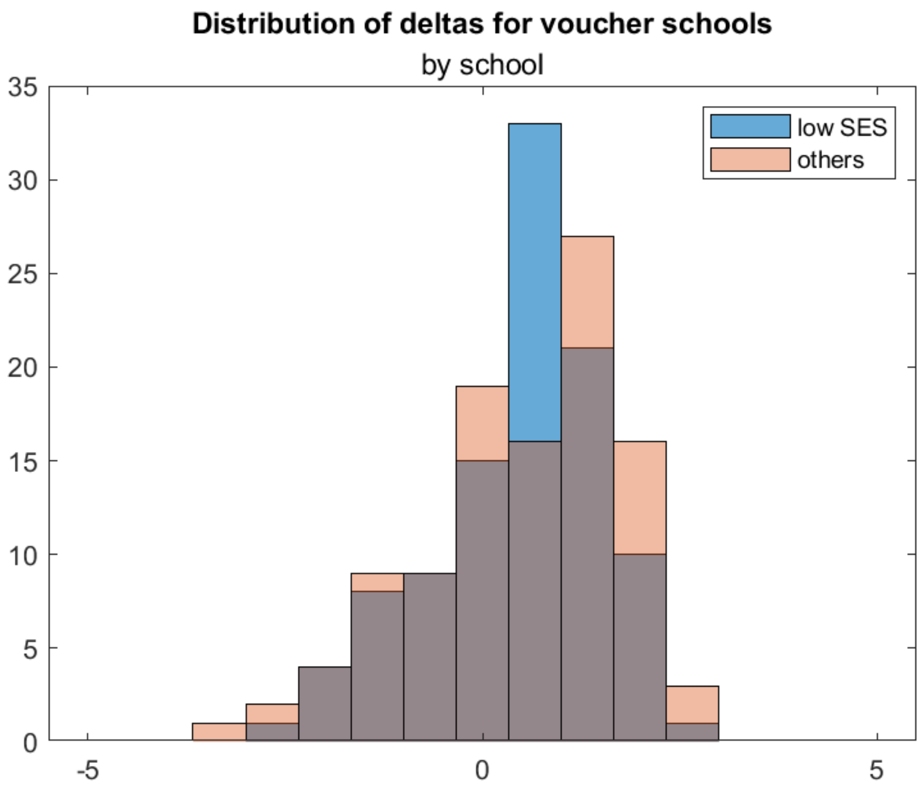


Figure 1.3: Enrollment outcome, wave 2 regions, 2018



Note: This figure uses all students enrolled in first grade in 2020 as the prospective population for preK in 2018.

Figure 1.4: School quality values for voucher schools, Low SES vs other students



Distribution of the estimates of school fixed-effects.
Each school has one fixed-effect for low-SES and one for others.
Deltas for public schools are forced to 0, and not shown in this graph.

Figure 1.5: Percentage of courses with a proportion of low-SES students below the quota, 2018

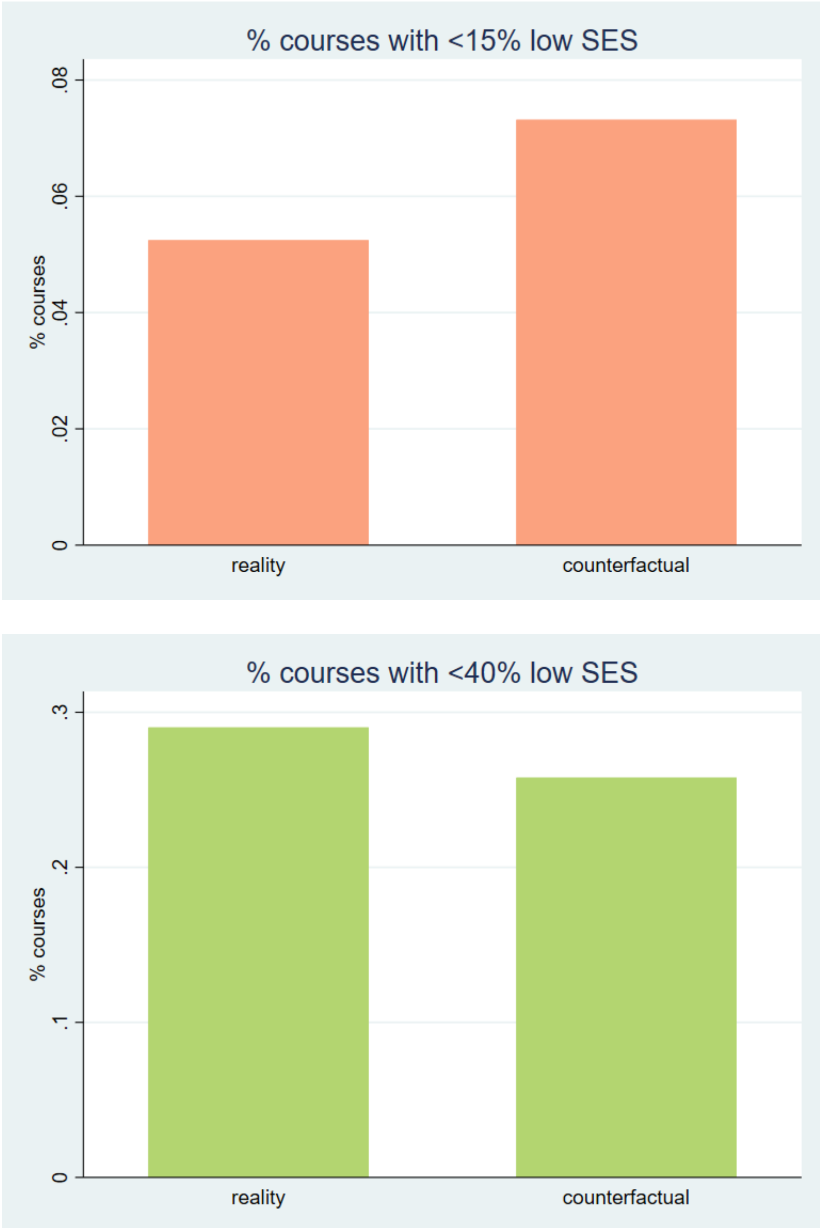


Figure 1.6: Distribution of the proportion of low-SES students across courses, 2018

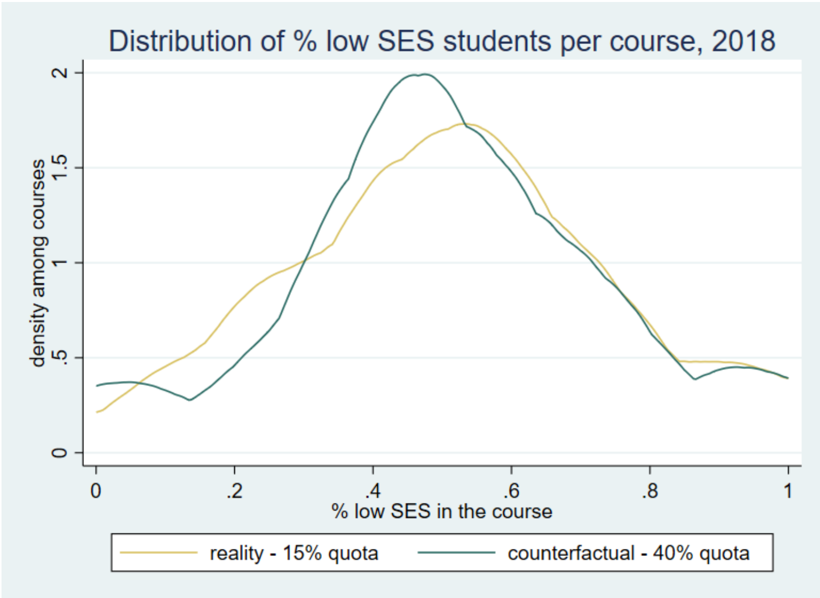


Figure 1.7: Unassigned students, 2018

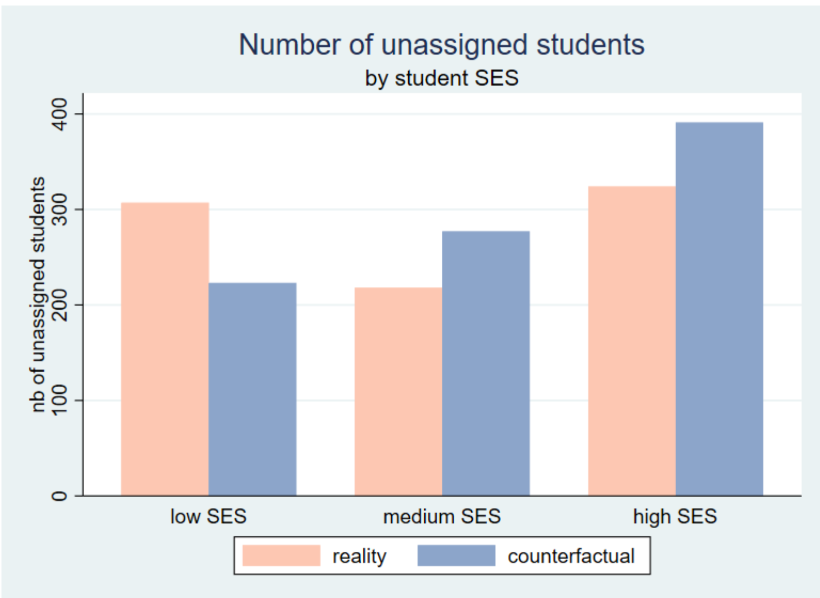


Figure 1.8: Distribution of the proportion of low-SES students across courses, counterfactual 10 ranks 2019

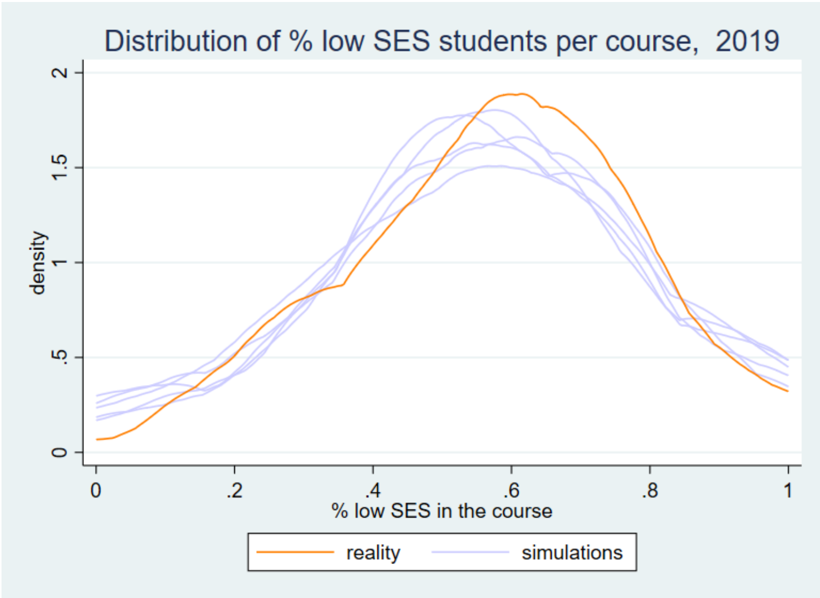


Figure 1.9: Students in voucher schools, counterfactual 10 ranks 2019

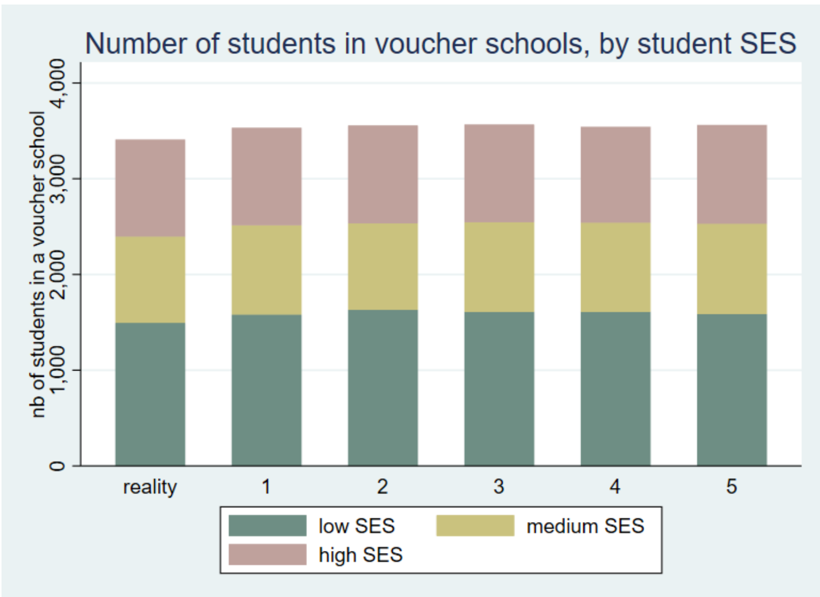


Figure 1.10: Unassigned students, counterfactual 10 ranks 2019

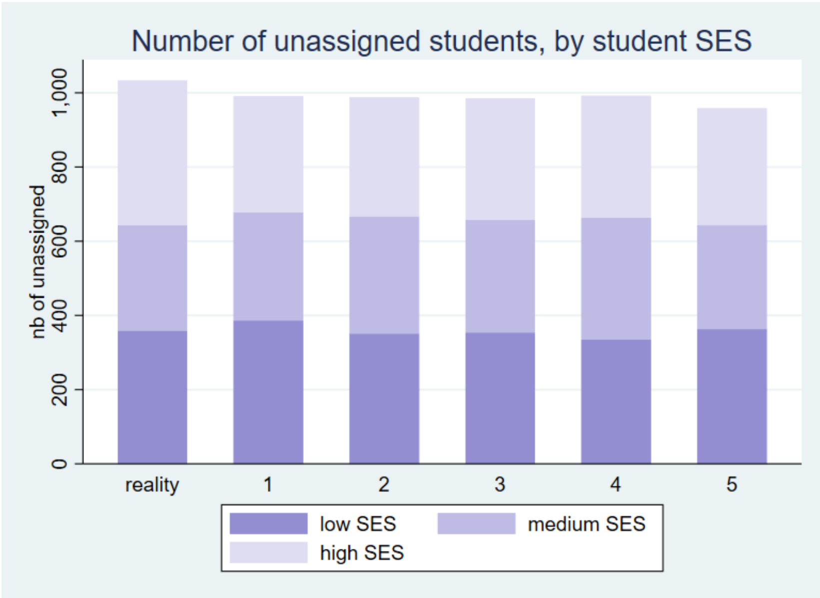


Figure 1.11: Quality of the assigned school, counterfactual 10 ranks 2019

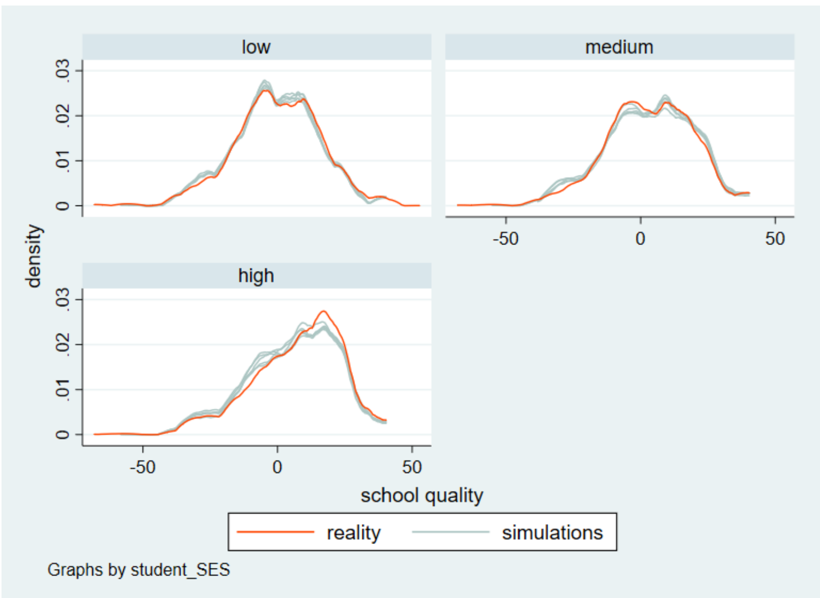


Figure 1.12: Number of unassigned students by list size, 2019

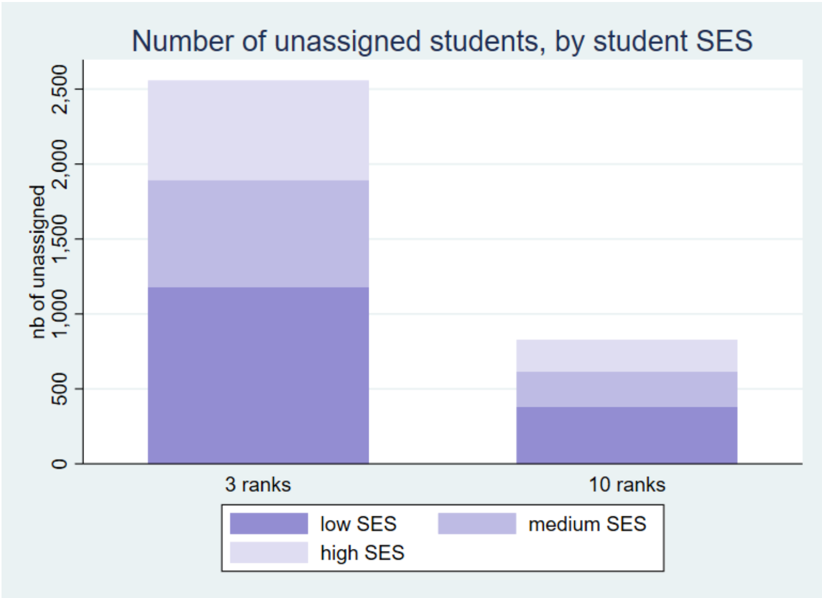


Figure 1.13: Distribution of % of low-SES students in the school by list size, 2019

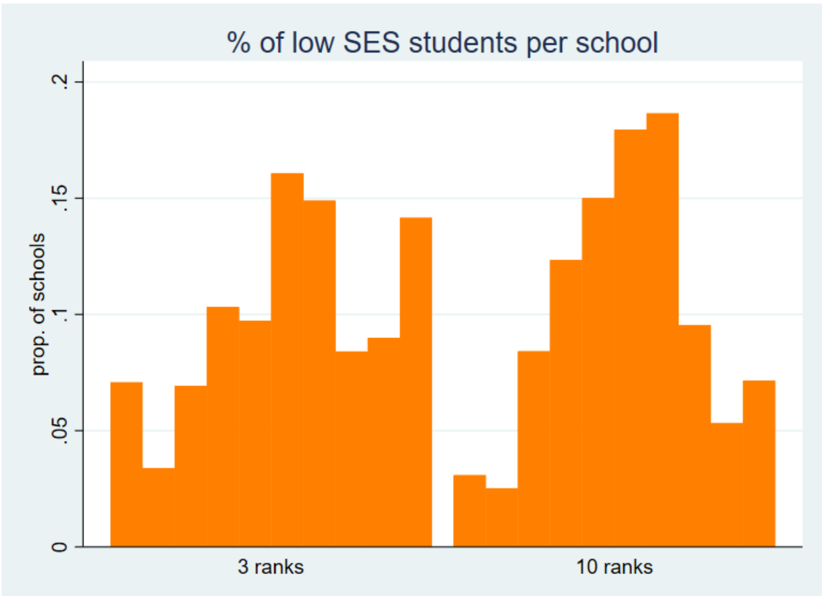
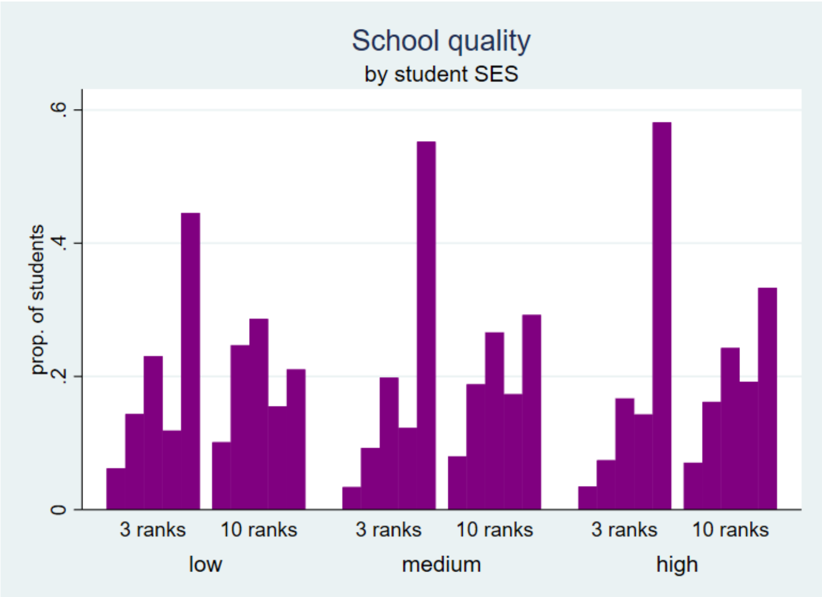


Figure 1.14: Distribution of school quality by list size
By student SES, 2019



Chapter 2

Applicant choice in the design of social housing allocations: Evidence from France

2.1 Introduction

Social housing is the provision of rental accommodation at subsidized rates. It represents 6% of the housing stock in developed countries [1], and a public expenditure of more than 50 billion dollars in the US.[2] Since rents are fixed, there is by definition no market for social housing and prices do not reflect market clearing. Units are allocated according to specific rules determining how applicants are selected, usually set by a housing authority. These rules vary greatly across countries and often stem from historical context or reflect public opinion.[3] However, matching theory shows how influential the choice of an allocation mechanism can be on final outcomes. In this light, assessing the welfare impact of a choice of allocation rules seems key. This is especially true since social housing is almost everywhere in excess demand, meaning that the applicant pool to choose from is large for any given unit.

Existing microeconomic theory has highlighted crucial tradeoffs inherent in the selection of allocation mechanisms for social housing (Arnosti and Shi (2020), Thakral (2016)) and compared theoretical properties of various existing mechanisms. However, the empirical literature is silent about the welfare impacts of a change in allocation rules.

This paper compares social housing allocation rules according to the household welfare they generate. It provides a framework to simulate social housing applications and assignments under any imposed rule. In order to simulate household application behavior under a modified allocation rule, I identify the determinants of household decision-making when applying for social housing. I model household behavior in the application process, disentangling preference over the location and features of social housing units from the household's chances of receiving an offer. Separately identifying preferences from expectations over future offers is made possible by estimating the model on a novel dataset that captures the geography of applicants' social housing requests in France as well as their reactions to the social housing offers they receive.

The first contribution of this paper is to construct a comprehensive dataset from French administrative records, encompassing all social housing applications and allocations nationwide over seven years. It includes exhaustive applicant information as viewed by housing authorities, and detailed data on the offers and acceptance decisions regarding thousands of units. This is a significant advancement, since housing authorities usually do not disclose application data. Without access

[1] OECD (2020), "Social housing: A key part of past and future housing policy", Employment, Labour and Social Affairs Policy Briefs

[2] "PH4.1 Public spending on support to social rental housing", OECD Affordable Housing Database, 2022

[3] The preference for veterans in US allocations is a good example. Source: OECD (2020), "Social housing: A key part of past and future housing policy", Employment, Labour and Social Affairs Policy Briefs

to information on the pool of applicants, we cannot inform the decision to apply to social housing, nor can we distinguish between the applicants' preference for particular units and their probability of obtaining an offer.

This paper also bridges a gap in the literature by modeling the choices of both social housing authorities and applicants in the allocation process. In the model, authorities allocate housing by choosing from the pool of applicants in each city using a scoring rule. When a unit becomes available, applicants are ranked according to a unit-specific score, and the unit is offered to the applicants with highest score. As for households decisions, in the model they are twofold: households first choose whether to apply to social housing and in which city to do so, and they then decide whether to accept an offer if they receive it. The first decision is to choose whether to apply to social housing. If they apply, they choose a list of cities and join their waitlists. This decision is modeled as a discrete choice depending on household and city characteristics, as well as on the net present value of potential offers in each city. Applicants subsequently decide whether to accept an offer they receive. This is represented using a continuous-time dynamic framework, where an offer is accepted only if its value surpasses the expected future offers' discounted value, considering the likelihood of receiving these offers. This part of the model builds on a framework from the organ transplant literature (Agarwal, Ashlagi, Rees, Somaini, and Waldinger (2021)) by allowing for offer rejections to impact future outcomes and for allocation rules to be unknown.

The third contribution of this work is to develop an estimation method to identify current allocation rules and applicants' decisions to participate, choose a city, and accept a unit offered. Parameters of the allocation rule governing housing authorities' behavior are estimated using data on offers made and available applicants. As for applicants, I estimate the parameters of their preferences for applying to cities using observed choices in the applications data. To estimate the primitives of their decision to accept an offer they receive, I use data on applicant response to social housing offers. Household valuation for social housing units and waitlists are represented as functions of conditional choice probabilities to accept an offer, drawing on the methodological approach of Hotz and Miller (1993). Using flexible estimates of this conditional choice probability, I can approximate applicants' value for waiting and their preference over social housing units.

Estimation results show a significant variation in allocation practices between cities. However, vulnerable groups such as low-income households, single mothers and foreigners seem to have a lower chance of receiving social housing everywhere, which aligns with official reports on France's social housing sector. Applicants prioritize lower rents, especially for less expensive homes. Households with lower incomes also prefer neighborhoods with quality schools. When deciding where to apply to social housing, households tend to favor the city where they currently live, especially when they have a permanent job contract. However, they give considerable weight to their social housing prospects in the city. Households with higher income tend to have a stronger preference for applying to cities with a higher median income, suggesting a form of peer preference. As for households, results show that they strongly account for expected chances to receive a unit when making decisions, although their taste for unit characteristics play a role. Low-income households, immigrants and single mothers are shown to have a higher valuation for social housing units on average, although the overall welfare they get from the social housing sector is lower than others due to their relatively lower probability to receive an offer.

Finally, this paper provides a framework to measure welfare under counterfactual allocation rules. I simulate allocations using estimation results to approximate household behavior. Instead of using the allocation rules as modeled and estimated, I impose an allocation rule and compare final outcomes with the current system. The challenge in running these simulations lies in the fact that applicants form expectations over their chances of receiving an offer when making decisions. These chances change when the allocation rule is modified, and therefore the expected probabilities to receive a unit need to be recomputed in the simulations. I assume these expectations follow the long-run steady state probability of receiving an offer, and build on an algorithm from Agarwal et al. (2021) to compute it. I depart from Agarwal et al. (2021) in that I incorporate households' decision to participate in the applications and their choice of a waitlist. I additionally incorporate the impact of an offer rejection on the applicant's future outcome on the waitlist. I simulate counterfactual allocation outcomes under various scoring rules focusing on applicant characteristics, or their waiting time. Results indicate that better targeting low-income applicants in the scoring rule has the potential to increase welfare very significantly. Since those applicants have a higher valuation for social housing units, increasing their chances of receiving a unit increases the average value that an offer holds for its recipient. This however raises incentives to participate in social housing allocations, leading to an increase in waitlist size. Simulation results show that the welfare increase coming from better match values outweighs the decrease coming from increased congestion. Simulations also show that tying the probability to receive an offer to the time spent on the waitlist like first-come-first-serve or last-come-first-serve would decrease welfare by completely decorrelating the probability to get an offer and the value of a match.

In the subsequent sections, after summarizing the existing literature and specifying my contributions in section 2.2, I will delve into the intricacies of the French social housing sector and the datasets employed in sections 3.2 and 3.3. The model is articulated in section 2.5, followed by an exposition of the estimation strategy in section 2.6. Results are analyzed in section 2.7, and counterfactual scenarios are examined in section 2.8.

2.2 Literature

This paper relates to the literature on social housing allocations, matching markets and housing demand.

Previous empirical work on social housing allocations has often used data on assignments (Geyer and Sieg (2013) ; Van Ommeren and Van der Vlist (2016) ; Sieg and Yoon (2020)) or aggregated applications (Ferdowsian, Lee, and Yap (2022)) to estimate demand for social housing. Van Dijk (2019) models application choices of a myopic household using data on applications, but her emphasis is on measuring the impact of social housing on socioeconomic outcomes.

Informing policy decisions on the choice of a rule to allocate social housing requires access to application data, and a focus on estimating the drivers of application decisions. Waldinger (2021) partly bridges this gap in the literature by modeling social housing applications within the city of Cambridge, Massachussets. He is however restricted by the fact that his access is limited to one city, and therefore needs to assume out the possibility for applicants to optimize across different cities. This spatial optimization matters since a change in social housing policy at the country-level would likely impact the distribution of social housing applicants across space. Additionally, unlike Cambridge, many countries [4] operate under non-transparent, locality-specific allocation rules that permit a degree of discretion among public officials. My paper estimate these rules and incorporate them in a model of household application decisions to providing guidance regarding allocation policies in these contexts. My paper also complements Waldinger (2021) by the application mechanism used for identification. Although Waldinger (2021) uses data from a mechanism where applicants directly pick a social housing development, I focus on a setting where applicants receive take-it-or-leave-it offers, as is relatively common across the OECD.[5] I therefore provide a method to identify preferences for social housing that can be replicated in many other countries displaying such an application mechanism.

My paper relates to the broad literature on matching markets. Most of this literature focuses on static mechanisms where many applicants match with one resource as in school or college assignments (Abdulkadiroğlu et al. (2017) ; Agarwal and Somaini (2018) ; Fack et al. (2019) to name only a few). Among the few papers that focus on dynamic one-to-one matching problems (Agarwal, Hodgson, and Somaini (2020), Ederer (2023)), Robinson-Cortes (2019) uses duration models

[4] including France, Ireland, Norway, New Zealand, Sweden and the UK

[5] France, Finland, UK... Source: OECD (2020), "Social housing: A key part of past and future housing policy", Employment, Labour and Social Affairs Policy Briefs

to represent foster care placements and aforementioned Waldinger (2021) uses a portfolio-choice model, whereas I focus on discrete-choice methods. Verdier and Reeling (2022) and Agarwal et al. (2021) also use a discrete-choice framework to model applicant decisions in the case of the allocation of bear-hunting licenses for the former, and deceased-donor kidney transplants for the latter. This paper contributes to this literature by incorporating the participation decision and the choice of a waitlist before entering the optimal stopping phase. Additionally, it considers a case where an applicant can only reject a limited amount of offers before getting removed from the waitlist.

This paper also builds on the housing literature focused on estimating demand for private housing. It specifically relates to estimates of housing demand with dynamic considerations (Bajari, Chan, Krueger, and Miller (2013) ; Bayer, McMillan, Murphy, and Timmins (2016) and more). Diamond and McQuade (2019) measures the impact of nearby affordable housing on private house prices. I complement this literature by estimating demand for social housing by social housing applicants.

2.3 Context

Social housing is widely spread in France. There are 5 million social housing units in France in which 18% of French households live - about 12 million people.[6] This is unevenly spread across the territory: social housing units represent between 10 and 25% of the housing stock depending on the region.[7] Those units are mostly apartments, but 16% of them are houses.[8] Overall, in September 2023, about half of the French population either was living in social housing or had lived there in the past.[9]

Although social housing serves a poorer population than the private market on average - in 2017, the poverty rate among social housing tenants was 35% and only 23% in the private rental market[10] - a wide range of the population is eligible. There are two conditions for eligibility: legally have the right to reside in France, and earn an income that falls below thresholds defined by law. These thresholds vary by region and household size, and are updated every year. As an example, in 2022, a single-person household living in Paris had to be earning below 24,316 euros per year to be eligible. In 2021, around 66% of the French population was eligible. [11]

[6] “Les Chiffres Clés du logement social”, Union Sociale pour l’Habitat, Août 2023
[7] InseeAnalyses #110, Hauts-de-France, 2020
[8] Ibid. see footnote 6
[9] “Chiffres et statistiques du logement social”, ministère de la Cohésion des territoires et des Relations avec les collectivités territoriales, septembre 2023
[10] Insee Références “Les conditions de logement en France”, édition 2017
[11] “Mieux connaître la demande de logement social pour mieux orienter les politiques publiques” USH pour le 81e congrès HLM, 2021

In France, as in most countries, the demand for social housing far exceeds its supply. At the end of 2022, about 2.4 million households were on the waitlist for social housing - about 480,000 had been allocated a unit that same year.[12]

Social housing applications can be done online or physically at a social housing developer's counter. Households are asked to list cities in which they apply by order of their preference. They then join the waitlist of all these cities and can receive an offer from anywhere within them. Every year, applicants need to renew their applications by providing some paperwork - their application is removed if they fail to do so.

Social housing developers manage social housing units and oversee their development. To finance new projects, they receive subsidies from multiple sources including the State, municipalities, workers' associations, and sometimes large corporations. In exchange, these sponsors are allocated a predetermined percentage of the housing units, which they can distribute based on their specific needs. Sponsors (henceforth referred to as Sponsoring Allocators) have the discretion to choose occupants, provided they meet legal criteria such as income limits and family situation. This mechanism allows them to cater to distinct demographics or necessities within their area of influence. When Sponsoring Allocators, State authorities identify a targeted subset from the pool of all applications and make selections when a unit under their quota becomes available.

When a unit becomes available, the Sponsoring Allocator makes a selection of a few applicants which is then discussed and ordered in a committee composed of representatives of the State, the municipality, and social housing developers. The first-ranked applicant is informed of the decision and has a few days to accept or reject the offer - up to a week, depending on the developer in charge of the unit. In case of a rejection the unit is offered to the next applicant on the list. In rare cases where the list is exhausted, the committee has to meet again to agree on another list.

Depending on the city and on the Sponsoring Allocator, the consequences for declining offers are variable. Applicants are commonly required to provide a rationale for their refusal, offering a broad margin for subjective judgment. Unofficial guidance for potential or existing social housing applicants universally underscores the cost associated with turning down offers, a cost that escalates with successive rejections.

Expanding the focus to the entire set of allocation rules reveals a significant lack of transparency. Statutory regulations allow for considerable leeway and are further complicated by multiple layers of local directives. Prior research in both economics and sociology has aimed to decipher and

[12] Ibid. see footnote 6

quantify these rules (Bonnal, Boumahdi, Favard 2012; Bourgeois 2018). However, applicants can usually access average waiting times by city for the previous few years.

2.4 Data

2.4.1 Datasets used

My principal dataset is sourced from the "Système National d'Enregistrement" (SNE), a compulsory portal for all social housing requests in France. Applicants submit a detailed form to the SNE, initiating a file that captures and updates their profile with any subsequent accepted social housing offer. These records demand annual renewal, and non-compliance results in application withdrawal. My tracking window spans from application entry to exit, regardless of outcome — securing a social housing unit or discontinuing the process by failing to renew. Upon application, households are asked to choose a list of cities for which their application will be considered. This list is stored in their SNE file together with the other information provided.

I have access to a rich dataset containing all information provided at the time of application and the list of listed cities. Although names are removed for privacy, it is comprehensive in scope: it includes applicant demographics (age, income, information on job contract), household size, and geographic variables (city of residence and employment). Initial submission dates alongside removal or assignment dates afford me the means to calculate waiting times. Furthermore, I observe accepted offer characteristics, including apartment size, layout, energy efficiency, and rent. I also access information on the unit's building such as apartment count, construction age, upkeep status, and crucially, geocoordinates.

However, the SNE falls short in capturing offers rejected by the applicants. To fill this gap, I leverage another dataset detailing offers extended by the State when acting as the Sponsoring Allocator[13] for years 2019 to 2021 - including those that were rejected by the applicant. This dataset, called "Système Priorité Logement" (SyPLo), includes the subset of applicants seen as more vulnerable and thus prioritized by the State, which serves as the pool from which the State selects applicants for offers. This subset of applicants and offers will be my main sample.

In order to estimate the determinants of the decision to apply to social housing or not, I need access to data covering households who are eligible but do not apply to social housing. I use a dataset combining housing and income tax records from 2011 to 2021 called FIDELI (FIchiers DÉmographiques sur les Logements et les Individus). It allows me to observe information on

[13] See section 3.2 for a definition of Sponsoring Allocator.

household composition and income for all households paying taxes or receiving benefits in France. Additionally, it contains the geolocation of their accommodation, together with housing characteristics such as apartment size and layout.

2.4.2 Descriptives

Table 2.1 compares units that become available in Syplo and SNE samples. SyPLo units allocated in 2019 and 2020 represent a bit under 25% of the total amount of allocated units. Their rents are lower on average, although units are of comparable size to the rest.

Following results focus on a specific French department in the Paris region surrounding the city of Versailles (henceforth refered to by its postcode, 78). This is the focus of current results, although later versions will include comparisons across departments. Figure 2.1 shows the number of applicants in each city in the SyPLo sample between 2019 and 2021. The variance across cities is very large, ranging from a few hundreds to about 6,000. Focusing on the number of active applications to choose from when allocating a specific unit, 2.2 shows that the pool of applicants for a given units is generally extremely large, rarely going below a few hundreds.

Table 2.3 shows summary statistics for the distance between applicants' current city of residence and cities containing social housing for applicants who listed at least one city in department 78 in their application. The large majority of applicants seem to apply within their department of residence. Table 2.4 shows summary statistics for applicant household income.

Table 2.2 shows household descriptives for my main sample (*syplo*), the universe of applicants (*sne*) and the universe of French taxpayers (*univ*) for the whole of France, and for the department 78. Social housing applicants are younger than the rest of the population, and less often retired. Households are larger, and their heads more often unemployed. Households already living in social housing apply more often to social housing than the rest of the population. The population of non European foreign citizens is vastly overrepresented. Notable differences when comparing my main sample and the universe of applicants are the underrepresentation of households living in social housing and the overrepresentation of unemployed and the non European applicants.

Table 2.5 shows descriptives on the rate of offer acceptance across the population. The first striking fact is that the acceptance rate is around 60%, which seems very low considering that few applicants receive more than one offer. As expected, more vulnerable populations are more likely to accept an offer: poorer applicants, students, non French citizens and younger people have significantly higher acceptance rates. Cheaper units are more likely to be accepted, but other unit variables do not seem to generate such a gap in acceptance rates.

Figure 2.3 shows the distribution of number of cities in applicants' rank-ordered lists in my main sample. Around 32% choose to rank only one city, and very few applicants rank more than eight.

2.5 Model

Households in a market decide whether to apply to social housing. If they do, they submit a rank-ordered list of cities where they want to apply. If they apply, they then join each city's waitlist until they are removed either by being assigned a unit or by voluntary departure (see details in section 2.5.4). Applicants on a city's waitlist are considered whenever a unit becomes available in this city. They can thus receive an offer for a unit in any of the cities that they listed in their application. I assume the applicant cannot modify the list of cities once it is submitted, and she cannot apply to social housing more than once in her lifetime.

During their wait, applicants can receive offers and reject them. If they accept an offer, they are matched with the unit and leave the waitlist. If they reject, they continue waiting for another offer. If they reject a second offer, they have to leave the waitlist. This reflects the fact that in practice applicants loose rights to the pool of state-run social housing units (see Context section) after multiple rejections.

When a unit becomes available, applicants in that city are ranked according to a scoring rule described in section 2.5.2. The unit is offered to the applicant with highest score.

The model is divided in three stages. First, the housing authorities' choice of who to offer a given unit is modeled as resulting from a score. Using this score, applicants make two types of decisions. First, the decision to apply to social housing and the choice of a ranked-ordered list of cities. Second, the decision of accepting or rejecting an offer when received. I set up a dynamic model defining the value of being on the waitlist of a given city as a determinant of the offer acceptance decision. This object will also take a part in the decision of whether and where to apply.

2.5.1 Setting

Unit arrival and assignment

Social housing units become available at rate λ_m in city m. The characteristics z_j of the incoming unit j are drawn independently from distribution F_m.

When a unit j becomes available in city m, applicants in the waitlist of this city are ranked according to a scoring rule. The applicant with the highest score receives the offer, and can accept or reject. If she accepts, both applicant and unit match and leave the market. If she rejects, the offer is made to the applicant with second-highest score etc. Applicants are not aware of potential refusals by others before they receive the offer.

Assumption 1 *All units are assigned, there is no vacancy. This means that the pool of applicants is always great enough so that at least one of them will accept.*

Assumption 2 *The supply of social housing units is fixed and units arrive at a constant rate λ_m in each city m.*

Assumption 3 *(i) Applicant departures prior to assignment and unit arrivals are governed by independent and exogenous Poisson processes.*
(ii) The voluntary exit rate of an agent i after waiting t is $\delta_i(t) \equiv \delta(t; x_i)$. Each agent has a terminal date $T_i < \infty$ at which exit occurs with probability 1.

Assumption 3 (ii) reflects the fact that some social housing applicants fail to renew their waitlist which causes their application to be removed.

Timing of applicant decisions

When a household i with characteristics x_i arrives on the market, she needs to choose whether to apply to social housing, and if so which cities' waitlists to join. When making this decision she does not know which units will become available in the future, nor which offer she may receive.

If she chooses to apply to social housing, the applicant submits a list of cities where she wants to apply and waits for an offer in one of those cities. If she receives an offer, she can accept it and be matched with the unit, or reject it and continue waiting. An applicant can only reject at most one unit in each city, otherwise she has to leave the waitlist. Time is continuous, such that the probability to receive two offers at the same time is zero.

2.5.2 Housing authority: applicant selection

As the decision rule of the housing authorities over applicant characteristics is not public, I model the selection of applicants as depending on a score. For each social housing unit, all applicants on its city's waitlist get a score. For social housing unit j, applicant i gets score $s_{ij}(t_{ij}, r_{ij})$, which depends on applicant characteristics, waiting time when j gets offered t_{ij}, and the number of offers she has rejected in the city before j becomes available r_{ij}.

Since housing authorities may judge applicants differently according to the characteristics of the unit they are trying to assign, I allow for parameters to vary with unit characteristics. This allows, for example, for household size to have a different impact on the score depending on the size of the housing unit to allocate. Additionally, parameters are allowed to vary according to some characteristics of the city where the unit to be assigned is located. This reflects the fact that assignment decisions are made at the local level.

The score of applicant i for unit j located in city m writes as follows:

$$s_{ij}(t_{ij}, r_{ij}) = \bar{s}_{ij} + \nu_{ij} = \zeta(x_i, t_{ij}, r_{ij}; \alpha_j^x, \alpha_j^t, \alpha_j^r) + \nu_{ij} \tag{2.1}$$

with

$$\alpha_j^k = \bar{\alpha}^k + \tilde{\alpha}^k z_j + \tilde{\tilde{\alpha}}^k w_m \quad \forall k \in \{x, t, r\}$$

where x_i are applicant i's characteristics and z_j are unit characteristics. ν_{ij} is an iid shock following a T1EV distribution and w_m are characteristics of city m where unit j is located. ζ is a flexible function known up to a set of parameters vectors $\alpha = (\alpha_j^x, \alpha_j^t, \alpha_j^r)$.

The shock ν_{ij} is assumed to depend only on i and j, and not on waiting time t and rejections r. The reason for this is that each unit is assumed to become available only once in the sample, and therefore a given applicant i will be considered for a given unit j at most once. It is therefore unnecessary to add a dependence on t and r.

I implicitly assume here that rejections in other cities do not affect an applicant's score. Similarly, as detailed in section 2.5.4, an application gets removed from a given city's waitlist after the applicant has rejected a second offer in the same city. Therefore, an applicant could technically reject more than two offers in different cities and still be on their waitlists. These assumptions tremendously simplify the representation of the value of being on a waitlist as defined in section 2.5.4, in a context where applicants can be on the waitlist of multiple cities at the same time.

For each social housing unit that becomes available, authorities rank all applications in the unit's city according to their unit-specific score. Unit j in city m is offered to applicant $i^\star$ such that:

$$s_{i^\star j} = \max_{i \in \mathcal{I}_j} s_{ij}$$

where $\mathcal{I}_j$ is the set of applicants in city m when unit j becomes available.

2.5.3 Applicant: decision of whether and where to apply

Whenever they arrive on the market, households need to decide whether they apply to social housing, and if so they need to choose a set of city waitlists to join. I assume applicants list cities according to their true preferences, and are evaluated as candidates in each of these cities, irrespective of the city's position on the applicant's list. They can list as many cities as they want. Rank-ordered lists are submitted upon arrival on the market and cannot be modified subsequently.

Applicants are specifically asked to rank truthfully on the application website, and they have no incentive to strategize one way or another while submitting their list of cities. Therefore, it is reasonable to assume they rank according to their true preference ordering.
When listing cities, I assume that applicants hold rational expectations over the frequency and nature of unit arrival in each city. Their beliefs over the probability to receive an offer conditional on its availability will be detailed in section 2.5.4.

A household chooses a rank-ordered list of waitlists $m \in \mathcal{M} \cup 0$, where $\mathcal{M}$ is the set of cities considered and $m = 0$ means they do not apply to social housing. Cities are ranked such that the first one chosen maximizes the applicant's utility, the second one yields second-highest utility etc. Applicants stop listing when the outside option $m = 0$ is preferred to unranked cities.

The utility from applying to city m is a function of city characteristics such as the proportion of social housing in the city or the proportion of unemployment.
Importantly, the utility for applying to a city m depends on the expected utility brought to applicant i by potential offers in this city. Let's define $V(t, r; x_i, m)$, the continuation value of being in the waitlist of city m for an applicant with characteristics x_i who has been waiting for t days on the list and has rejected r offers. Applicants compare this value to that of being matched with the offered unit when deciding whether to accept or reject an offer - more details in section 2.5.4. I allow the utility for applicant i to apply to city m, U_{im}, to depend on $V(t, r; x_i, m)$

(henceforth $V_{im}(t,r)$). Specifically, I let U_{im} depend on $V_{im}(t=0, r=0)$, since at the time of listing cities the applicant has not yet waited any day nor rejected any offer.
Additionally, applicants are allowed to have heterogeneous tastes for city characteristics based on their own characteristics.
The utility for applicant i of applying to city m writes:

$$U_{im} = w_m\gamma_1 + x_i\gamma_2 + \gamma_3\text{Dist}_{im} + \gamma_4 V_{im}(0,0) + \eta_{im}$$

where Dist_{im} is the distance between the city where household i currently lives and city m, x_i is a set of applicant characteristics, w_m is a set of city characteristics, η_{im} is an iid shock which is assumed to follow a Type 1 Extreme Value distribution. This shock is observed by applicants at the time of making their application. Note that η_{im} is not correlated to ν_{ij}, the shock to applicant i's score.

The value of the outside option of not applying to social housing is normalized to 0. I will further discuss this normalization in section 2.6.4. Note however that I allow for heterogeneity in the cost of applying (or taste for social housing) across households, which is captured by the component $\gamma_2 x_i$ of U_{im}. See section 2.5.4 for a discussion on how the identification of γ_2.
Ranking decisions are modeled as an exploded logit, where a household submits the rank m_1, m_2, m_3 if and only if

$$U_{im_1} + \eta_{im_1} \geq U_{im_2} + \eta_{im_2} \geq U_{im_3} + \eta_{im_3} \geq 0 \geq U_{im'} + \eta_{im'} \quad \forall m' \in \mathcal{M} \setminus \{m_1, m_2, m_3\}$$

As a household can submit a rank-ordered list of any size, the size of the list is determined by the amount of cities for which $U_{im_1} + \eta_{im} \geq 0$, zero being the value of the outside option.

2.5.4 Applicant: decision to accept an offer

The decision to accept or reject an offer is modeled as an optimal stopping framework in continuous time, where applicants are facing a trade-off between accepting the offer and waiting to get a preferred one in the future. The choice of continuous time stems from the advantages of avoiding simultaneous events for tractability - e.g simultaneous offers, or contiguous offer arrival and applicant exit.
Offer characteristics are drawn randomly from F_m, so an applicant does not know the characteristics z of upcoming offers. Additionally, applicants know the score function described in section 2.5.2, and are therefore aware that their time on the waitlist affects their probability to receive an

offer.

This part of the model follows closely Agarwal, Ashlaghi, Rees, Somaini and Waldinger (2021), which I henceforth refer to as AARSW. I complement their framework by adding the number of past rejected offers as a state variable, and introducing a limit for this variable above which a new rejection results in the applicant leaving the waitlist.

Payoffs

When deciding whether to accept an offer, an applicant i compares the value of matching with the offered unit $\Gamma_{ij}(t)$ with the continuation value of being on the waitlist of city m conditional on having waited t days and rejected r offers in the past, $V_{im}(t,r)$. If the applicant leaves the waitlist, they receive the payoff associated with the outside option, which is 0. Applicants discount future payoffs at rate ρ.

Assumption 4 *(i)* $\Gamma_{ij}(t) \equiv \Gamma(t, x_i, z_j, w_m) + \varepsilon_{ij}$*, where* x_i *is applicant characteristics,* t *the number of days on the waitlist,* z_j *unit characteristics and* w_m *are characteristics of city* m*, where unit* j *is located.*
(ii) $\varepsilon_{ij} \sim \mathcal{N}(0,1)$ *and iid, and* $\varepsilon_{ij} \perp\!\!\!\perp \nu_{ij}$

Note that assumption 4 (i) implicitly states that the value of a match does depend on the time waited on the list but not on the number of offers rejected. It additionally assumes idiosyncratic taste shocks enter the match value addititively. Assumption 4 (ii) states that the preference shock on the value of a match is independent from the shock on her score ν_{ij} and from the shock to her city tastes η_{im}. The former is violated if a variable correlated with an applicant's tastes for housing makes an individual more attractive to the housing authority.

Beliefs over future offers

Applicants know the score function and the value of observables for all assigned applicants in the sample. They have perfect foresight on the pool of available applications at every point in time. However, they do not observe ν_{ij}.

Assumption 5 *Agent* i *who has waited* t *periods on the waitlist of city* m *believes the probability an object with characteristics* z_j *will be available to her if it arrives is*

$$\pi(t_i, r_i; z_j, x_i, w_m) = P[s_{ij}(t_i, r_i) = \max_{i'} s_{i'j}(t_{i'}, r_{i'}) \, \forall i' \in \mathcal{I}_j] \tag{2.2}$$

where $\mathcal{I}_j$ *is the set of available applications in city* m *when unit* j *becomes available.*

Assumption 5 states that applicants know the scoring rule and the composition of the waitlist at every point in time, such that they know the pool of available applications when a unit becomes available. This is a strong assumption, but I believe it is a reasonable approximation of a reality where the pool of applicants is so large that the overall chances for an applicant to receive a unit changes very little over time. Applicants are likely to have a sense of the size and composition of the waitlist at a given point in time by word of mouth and thanks to the information provided on the websites of various social housing authorities. If variations on these metrics alter only marginally their chances of receiving a unit, π will be a good approximation of their beliefs. Such an assumption on beliefs is necessary to keep the model tractable and avoid letting beliefs depend on other applicants' actions. Alternative definitions of π can be tested in robustness checks.

Value functions

As the applicant may at most have one rejection, accepting an offer for an applicant with one past rejection does not involve a trade-off between leaving the waitlist or staying. This is however not the case for an applicant with no past rejection. Therefore, the computation of $V_{i,m}(t, r = 1)$ and $V_{i,m}(t, r = 0)$ are done separately.

Assumption 6 *When considering an offer in city m, applicants are myopic about their chances of receiving an offer in another city of their waitlist.*

Assumption 6 greatly simplifies the analysis by removing the dependency across value functions of different cities. Implications and robustness checks are discussed in section 2.6.4. For the rest of the analysis, $V_{i,m}(t, r)$ can be thought of as the option value of waiting in city m, fixing the value in other cities the applicant may have ranked to zero.

When receiving an offer, an applicant who has already rejected once chooses between exiting the waitlist or the value of accepting the offer. The continuation value for applicant i in waitlist m is thus defined by the following differential equation:

$$\left(\rho + \delta_i(t) + \int \pi_{iz}(t, 1)\, dF_m\right) V_{i,m}(t, 1) = \lambda_m \int \pi_{iz}(t, 1) \int \max\{0, \Gamma_{iz}(t)\} dG dF_m + \frac{\partial V_{i,m}(t, 1)}{\partial t} \tag{2.3}$$

where I abuse notation by writing $\pi_{iz}(t, r) = \pi(t, r, x_i, z, w_m)$ and $\Gamma_{iz}(t) = \Gamma(t, x_i, z, w_m) + \varepsilon_{ij}$. The first integral on the right hand side is taken with respect to the unit characteristics z of the incoming unit, which are drawn from F_m. The second integral is taken with respect to ε_{ij}, the unobserved component to the value of matching with unit j. G is therefore the cdf of a $\mathcal{N}(0,1)$. You can find more information on how I derive this equation in Appendix B.1.1.

When the applicant has never rejected an offer in the past, the value function writes:

$$\Big(\rho + \delta_i(t) + \int \pi_{iz}(t,0)\,dF_m\Big)V_{i,m}(t,0) = \lambda_m \int \pi_{iz}(t,0) \int \max\{0, \Gamma_{iz}(t) - V_{i,m}(t,1)\} dG dF_m + \frac{\partial V_{i,m}(t,0)}{\partial t} \tag{2.4}$$

Equation (2.4) can be rewritten as follows (see Appendix B.1.2 for the proof):

$$V_{im}(t,0) = \int_t^{T_i} e^{-\int_t^s [\rho + \delta(\tau;x_i) + \int \pi_{iz}(\tau,0)\,dF_m]\,d\tau} \Big(\lambda_m \int \pi_{iz}(s,0) \int \max\{V_{im}(s,1), \Gamma_{ij}(s)\} dG\,dF_m\Big) ds \tag{2.5}$$

Equation (2.3), together with end condition $V_{im}(T_i,1) = 0$, is a first-order linear differential equation. In order to simplify its expression, let me rewrite $\int \max\{0, \Gamma + \varepsilon, \}\,dG(\varepsilon)$ as a function of Γ.

$$\begin{aligned} \xi(\Gamma, V) &\equiv \int \max\{\Gamma + \varepsilon, V\}\,dG(\varepsilon) \\ &= \big(\Gamma + E[\varepsilon \,|\, \varepsilon > V - \Gamma]\big)\big(1 - P[\varepsilon < V - \Gamma]\big) + V\,P[\varepsilon < V - \Gamma] \\ &= \left(\Gamma + \frac{\Phi'(V-\Gamma)}{1 - \Phi(V-\Gamma)}\right)\big(1 - \Phi(V-\Gamma)\big) + V\,\Phi(V-\Gamma) \end{aligned} \tag{2.6}$$

where (2.7) comes from the expression for the conditional expectation of a standard Normal variable

$$E[\varepsilon | a < \varepsilon < b] = \frac{\Phi'(a) - \Phi'(b)}{\Phi(b) - \Phi(a)}$$

Equation (2.3) now writes:

$$\Big(\rho + \delta_i(t) + \int \pi_{iz}(t,1)\,dF_m\Big)V_{i,m}(t,1) = \lambda_m \int \pi_{iz}(t,1)\,\xi(\Gamma_{ij}(t),0)\,dF_m + \frac{\partial V_{i,m}(t,1)}{\partial t} \tag{2.7}$$

Its solution writes:

$$V_{im}(t,1) = \int_t^{T_i} e^{-\int_t^s [\rho+\delta(\tau;x_i)+\lambda_m \int \pi_{iz}(s,1)\,dF_m]\,d\tau} \Big(\lambda_m \int \pi_{iz}(s,1)\xi\big(\Gamma_{ij}(s),0\big)\,dF_m\,ds\Big) \tag{2.8}$$

You can find the proof in Appendix B.1.3.

2.6 Estimation

The estimation of the model is divided in three stages. I first estimate the score function using data on offers made to applicants. I then estimate parameters entering the decision to accept or reject an offer thanks to a CCP representation. I finish by estimating parameters of the decision of which rank-ordered list of cities to submit using rank-ordered lists submitted by applicants in the data.

2.6.1 Estimation of the score function

Parameters $(\bar{\alpha}^x, \bar{\alpha}^t, \bar{\alpha}^r, \tilde{\alpha}^x, \tilde{\alpha}^t, \tilde{\alpha}^r, \tilde{\tilde{\alpha}}^x, \tilde{\tilde{\alpha}}^t, \tilde{\tilde{\alpha}}^r)$ are estimated by maximum likelihood, following a multinomial logit formulation. The model is estimated on the sample of offers[14], using the first offer that was made for a given unit.

Since housing authorities choose one applicant over many hundreds of applications for each unit, probabilities to receive an offer are bound to be very small. This will lead to an approximation bias which can affect the estimation. In order to circumvent this issue, I randomly sample K non-chosen applicants from the city's waitlist for each apartment. Those will be the only ones considered as alternatives to the chosen candidate in the housing authorities' choice set when filling a given unit.

2.6.2 Estimation of parameters of offer acceptance decision

The primitives that we want to recover here are the match value $\Gamma(t, x_i, z_j, w_m)$ and the value functions $V_m(t, r; x_i, m)$. I rewrite them as functions of conditional choice probabilities (Hotz and Miller (1993)) and exogenous terms that can be estimated separately such as the rate of unit arrival λ_m and the rate of applicant departure $\delta_i(t)$. The discount rate ρ is fixed to 5% per year.

Applicants are assumed to stay a maximum of 15 years on the waitlist. They are assumed to die at age 100. Therefore, $T_i = \max\{100 - \text{age}_i, 15\} \times 365$, where age_i is the age of applicant i in years when they enter the market.

[14]Which covers offers made by state-owned social housing units only. See section 3.3 for more detail.

Rate of applicant departure and unit arrival

I estimate a censored Gompertz proportional hazards model with the following parametrization of the rate of departure:

$$\delta(t; x_i) = \delta_1 \exp(\delta_2 t) \exp(x_i \beta) \tag{2.9}$$

where $\delta_1 \exp(\delta_2 t)$ is the baseline hazard function. δ_1, δ_2 and β are estimated by MLE.
Note that the distribution of observed waiting times is censored to the right since some applicants are assigned a unit before they choose to voluntary exit. The analysis accounts for this right-censoring.

The unit arrival rate λ_m is modeled as a Poisson process with covariates w_m and estimated using maximum likelihood.

$$\begin{aligned} \log(\lambda) &= a_1 + a_2 w_m \\ \Leftrightarrow \lambda &= e^{a_1 + a_2 w_m} \end{aligned} \tag{2.10}$$

CCPs

In order to recover match values $\Gamma(t, x_i, z_j, w_m)$, we rewrite them as a function of conditional choice probabilities (CCP). We thus need a flexible estimate of the CCP to approximate $V_m(t, r; x_i, m)$.

I use a probit formulation to estimate the conditional choice probabilities. The probability of accepting an offer is written as follows:

$$P(t, r, x_i, z_j, w_m) = \Phi(\chi(x_i, z_j, t, r, w_m; \theta)) \tag{2.11}$$

where $\chi(.)$ is a flexible set of functions, θ its vector of parameters, and $\Phi(.)$ is the cdf of the standard normal distribution. Parameters θ are estimated by maximum likelihood.

Match value and continuation value

In the case where the applicant has rejected one offer, so $r = 1$. An applicant accepts the offer if and only if the match value is greater than the outside option:

$$\Gamma(t, x_i, z_j, w_m) + \varepsilon_{ij} \geq 0$$

Let $P^r_{ijtm} \equiv P(t, r, x_i, z_j, w_m)$ be the conditional probability that agent i accepts an offer of unit j given (t, x_i, z_j, w_m) and r rejections. Given that ε_{ij} is iid distributed with cdf G, we can write:

$$P^{r=1}_{ijtm} = P[\Gamma(t, x_i, z_j, w_m) + \varepsilon_{ij} \geq 0 \,|\, t, r = 1, x_i, z_j, w_m] \tag{2.12}$$

$$= 1 - G\big(- \Gamma(t, x_i, z_j, w_m)\big) \tag{2.13}$$

We can now write the match value in terms of the CCP:

$$\Gamma(t, x_i, z_j, w_m) = -G^{-1}(1 - P^{r=1}_{ijtm}) \tag{2.14}$$

Therefore, we can estimate Γ by replacing $P^{r=1}_{ijtm}$ with a flexible estimate $\hat{P}^{r=1}_{ijtm}$ based on observed offer acceptance decisions, while using the fact that G is the cumulative distribution function of a standard Normal distribution.

Using a CCP estimate, I recover $\hat{V}_{im}(t, 1)$ by using equation (2.8) and replacing Γ, δ and λ with their estimates. Conditional on an estimate for $\{\pi_{iz}(t, r = 1)\}_{\forall z \in Z_m}$, we can replace integrals over F_m, the distribution of unit characteristics in m, by sums over all units in city m. This would yield the following estimating equation:

$$\hat{V}_{im}(t, 1) = \int_t^{T_i} e^{-\int_t^s [\rho + \hat{\delta}(\tau; x_i) + \frac{\hat{\lambda}_m}{|J_m|} \sum_{j \in J_m} \hat{\pi}_{ij}(s,1)]\, d\tau} \Big(\hat{\lambda}_m \frac{1}{|J_m|} \sum_{j \in J_m} \hat{\pi}_{ij}(s, 1) \xi\big(\hat{\Gamma}_{ij}(s), 0\big)\, ds\Big) \tag{2.15}$$

where J_m is the set of social housing units in m that are allocated in my sample, $\xi(.,.)$ is defined by equation (2.6), and the estimates derived in equations (2.9), (2.10) and (2.11) . The integral over s needs to be numerically approximated.

The only object that is left to be estimated aside from the CCP is the believed probability to receive an offer π. I estimated it using data on received offers:

$$\hat{\pi}(t, r, x_i, z_j, w_m) = P[\text{applicant } (x_i, t, r) \text{ first receives offer } (z_j, w_m)] \tag{2.16}$$

$$= \frac{\exp(\bar{s}_{ij})}{\sum_{i' \in \mathcal{I}_j} \exp(\bar{s}_{i'j})} \tag{2.17}$$

Using $\hat{V}_{im}(t, 1)$, I can recover $\hat{V}_{im}(t, 0)$ using equation 2.5:

$$\hat{V}_{im}(t,0) = \int_t^{T_i} e^{-\int_t^s [\rho + \hat{\delta}(\tau; x_i) + \frac{\hat{\lambda}_m}{|J_m|} \sum_{j \in J_m} \hat{\pi}_{ij}(s,0)]\, d\tau} \Big(\hat{\lambda}_m \frac{1}{|J_m|} \sum_{j \in J_m} \hat{\pi}_{iz}(s,0) \xi \big(\hat{\Gamma}_{ij}(s), \hat{V}_{im}(t,1) \big)\, ds \Big) \tag{2.18}$$

2.6.3 Estimation of parameters of application decision and city choice

Estimation method

For the estimation of the parameters of U_{im} $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$, I use maximum likelihood, where the likelihood of an observation i who submitted a rank of 3 cities m_1, m_2, m_3 is:

$$\frac{\exp(U_{im_1})}{1 + \sum_{m \in \mathcal{M}} \exp(U_{im})} \times \frac{\exp(U_{im_2})}{1 + \sum_{m \in \mathcal{M} \setminus m_1} \exp(U_{im})} \times \frac{\exp(U_{im_1})}{1 + \sum_{m \in \mathcal{M} \setminus \{m_1, m_2\}} \exp(U_{im})} \times \frac{1}{1 + \sum_{m \in \mathcal{M} \setminus \{m_1, m_2, m_3\}} \exp(U_{im})}$$

I make use of listed cities for all applicants of my state-targeted subsample of applicants.[15]

Dealing with sample selection bias

My main sample consists of households who chose to apply to social housing, and therefore it is choice-based. This biases my estimates since by definition I observe applicants for whom U_{im} is greater than the outside option for at least one city. This means that my estimates may not be accurate in predicting the probability for an applicant to apply to social housing. I try to address this bias using two alternative methods: weighting observations by the inverse of their probability to be sampled, and adapting the likelihood to fit the observed sample.

My first strategy is to weigh observations by the inverse probability of being in the sample, following Manski & Lerman (1977). This method serves as a comparison to the initial exploded logit estimation, although it does require the determinants of sampling to be independent from those of choice - which is unlikely to be satisfied in my case.

[15]See section 3.3 for more details on the samples I use.

In order to compute the probability to be in the sample of applicants, I adapt a method from Waldinger (2021). I make use of a sample containing the universe of French households (henceforth Universe sample) to model and estimate the probability for a given household to apply to social housing at the period of interest and therefore be found in my main sample. I then use this probability to sample households from Universe and add them to my main sample of social housing applicants.

I model and estimate the probability $\mathcal{V}(c_i;\beta^{\text{univ}})$ for a French household in Universe with characteristics c_i to be found in my main sample of applicants (henceforth Main sample), that is among state-targeted[16] social housing applicants in 2019-2021. To estimate this probability, I use a minimum distance estimator that matches several aggregate statistics $m(.)$ between Main and Universe samples.

$$\min_{\beta^{\text{univ}}} (m_{univ}(\beta) - m_{main})'(m_{univ}(\beta) - m_{main})$$

where

$$m_{univ}(\beta) \equiv \sum_{i=1}^{I} \amalg(c_i;\beta)\, m_i$$

and

$$\amalg(c_i;\beta) = \Phi(c_i'\beta)$$

m_i is household i's contribution to statistic $m(.)$, and m_{main} is the aggregate value of the statistic for the Main sample. I is the number of eligible households in the portion of the Universe sample considered.

I group cities in 9 categories according to their population size and median income. For each city group, for each of the three years in my sample, my minimum distance estimator compares the following statistics across the two samples:

- Total number of households
- Moments of the distribution of income (mean, variance)
- Number of households with at least one member who was not born in France
- Number of households with at least one member who is unemployed

I obtain $\hat{\amalg}(c_i)$ for each household in my main sample, and weigh observations by $\frac{1}{\hat{\amalg}(c_i)}$ in the

[16]Ibid. note 15

maximum likelihood described in section 2.6.3.

My second strategy to deal with sample-selection bias is to adjust to likelihood formulation to fit the selection. Specifically, the likelihood in section 2.6.3 describes the probability for individual i to submit a rank R, which I write $P[R_i = R]$ - where R_i refers to the list of cities provided by applicant i. Instead, what I observe in the sample where all households are applicants is $P[R_i = R \mid \max_{m\in\mathcal{M}} U_{im} \geq 0]$. I can write this probability as follows:

$$\begin{aligned} P[R_i = R \mid \max_{m\in\mathcal{M}} U_{im} \geq 0] &= \frac{P[R_i = R \cap \max_{m\in\mathcal{M}} U_{im} \geq 0]}{P[\max_{m\in\mathcal{M}} U_{im} \geq 0]} \\ &= \frac{P[R_i = R]}{1 - P[0 \geq \max_{m\in\mathcal{M}} U_{im}]} \end{aligned}$$

where

$$P[0 \geq \max_{m\in\mathcal{M}} U_{im}] = \frac{1}{1 + \sum_{m\in\mathcal{M}} \exp(U_{im})}$$

2.6.4 Discussion

The interpretation of city-specific value functions $V_{im}(t, r)$ is not straightforward in a context where households can apply to multiple cities. They do not represent the whole value of waiting for applicant i if the applicant has ranked multiple cities since she may receive an offer from another city. Therefore, $V_{im}(t, r)$ can be described as the option value of being in city m, fixing the net present value of potential offers outside m to zero.

To relax assumption 6, we should consider the general case where the continuation value of being on the waitlist is the value of waiting given a list of cities R. The value of waiting after having submitted rank R should be represented as a function of an applicant's rejection status in all cities in her list, $\bar{r} \equiv [r_1, ..., r_m, ...r_{|R|}]$ for all cities $m \in R$. It can be written as follows:

$$\begin{aligned} \Bigg(\rho + \delta_i(t) + \sum_{m\in R: r_m = 0} \Big[\int \pi_{iz}(t, 0)\, dF_m\Big] + \sum_{m\in R: r_m = 1} \Big[\int \pi_{iz}(t, 1)\, dF_m\Big]\Bigg) V_{i,R}(t, \bar{r}) = \\ \sum_{m\in R: r_m = 0} \Big[\lambda_m \int \pi_{iz}(t, 0) \int \max\{0, \Gamma_{iz}(t) - V_{i,R}(t, \bar{r}'_m)\} dG dF_m\Big] + \\ \sum_{m\in R: r_m = 1} \Big[\lambda_m \int \pi_{iz}(t, 0) \int \max\{0, \Gamma_{iz}(t) - V_{i,R_{(-m)}}(t, \bar{r}_{(-m)})\} dG dF_m\Big] + \frac{\partial V_{i,R}(t, \bar{r})}{\partial t} \end{aligned} \tag{2.19}$$

where $\bar{r}'_m \equiv [r_1, ..., r_m + 1, ..., r_{|R|}]$, $R_{(-m)}$ is the ranking R where city m is removed and $\bar{r}_{(-m)} \equiv [r_1, \ldots, r_{m-1}, r_{m+1}, \ldots, r_{|R|}]$. This expression accounts for the dependence of the score on past rejections, and for the fact that an applicant is removed of the waitlist of city m when she has rejected more than one offer. The separation of sums into those over cities where the applicant has already rejected and others reflects the impact of a rejection on future prospects in a given city. The last term shows that when deciding whether to accept an offer, an applicant compares the offer value Γ with the continuation value of being on a waitlist where m is removed, $V_{i,R_{(-m)}}$, since rejecting a second time in m means exiting its waitlist. Note that the occurrence of multiple offers across different cities at once has probability zero given that time is continuous.

The computation of this object is complex, since the amount of ranked cities is unrestricted and ranges from 1 to 17 in the data. Computing V_{iR} for each applicant based on their ranked-order list R_i is therefore computationally intensive. Additionally, instead of separating the computation of V_{iR} in 2 cases, as we do for $V_{im}(t, r = 0)$ and $V_{im}(t, r = 1)$, I would need to compute it for each value of $\bar{r}$, of which there are many if $|R| = 17$. This is especially true since you need to consider the cases where there is only one m left in R, only 2, etc. The amount of cases to compute $V_{iR}(t, \bar{r})$ is $2^{|R|} + 2^{|R|-1} + 2^{|R|-2} + \cdots + 2^{|R|-|R|+1}$.

Although computationally intensive, the computation of $V_{iR}(t, \bar{r})$ is feasible. Starting from the case where m is the only city left in R, which we will refer to as the case where $R = R_m$, we can observe that $V_{iR_m} = V_{im}$ as defined and computed in sections 2.5 and 2.6. Moving to the case where there are two cities (m, m') left in R, referred to as $R_{m,m'}$, we can write value functions in terms of known objects. Consider the case where $r_m = 1$ and $r_{m'} = 1$, so that $\bar{r} = [1, 1]$. The equation defining $V_{i,R_{m,m'}}$ can be written as

$$\begin{aligned}\left(\rho + \delta_i(t) + \int \pi_{iz}(t,1)\, dF_m + \int \pi_{iz}(t,1)\, dF_{m'}\right) V_{i,R_{m,m'}}(t,\bar{r}) \\ = \lambda_m \int \pi_{iz}(t,1) \int \max\{0, \Gamma_{iz}(t) - V_{i,R_{m'}}(t,\bar{r}_{(-m)})\} dG dF_m \Big] \\ + \lambda_{m'} \int \pi_{iz}(t,1) \int \max\{0, \Gamma_{iz}(t) - V_{i,R_{(-m')}}(t,\bar{r}_{(-m')})\} dG dF_{m'} \Big] \\ + \frac{\partial V_{i,R_{m,m'}}(t,\bar{r})}{\partial t}\end{aligned} \tag{2.20}$$

where the right hand side is divided between cases where an applicant receives an offer in m, and in m'. The first term displays the choice between accepting the offer, of value Γ and rejecting it, of value $V_{i,R_{m'}}(t, \bar{r}_{(-m)})$ since it would mean being removed from city m and reverting back to the case where there is only one city in R. The second term mirrors the first for an offer in city m'. Since $V_{i,R_{m'}} = V_{i,m'}$, all terms of the right hand side are known functions of t except for $\frac{\partial V_{i,R_{m,m'}}(t,\bar{r})}{\partial t}$. Equation 2.20 is therefore an ordinary differential equation that can be rewritten using the same derivation as that detailed in appendix B.1.3 for writing $V_m(t, r = 1)$.

Following the same logic, I could write V_{iR} for all rankings observed in the data. This would however be a heavy computation, which is why I focus my analysis on the simpler case where applicants consider only their future outcomes in the city where they make the offer acceptance decision.

I believe this simplification may not significantly impact my qualitative result, as the main tradeoffs between match value and future offer prospects are represented in the myopic case where applicants focus on one city when accepting an offer. I can verify this in a robustness analysis where I replace V_{im} by V_{iR}. Endogenous model parameters (match values Γ and value functions V) would be estimated from applicants that rank one city only. This is possible since the model assumes that match values, scoring rules and value functions do not depend on initial cities ranked. It is also made feasible by how large my dataset is, meaning that I observe thousands of individuals who ranked only one city.
Appendix B.2 displays parameter estimates for conditional choice probabilities when estimated on the pool of applicants who ranked only one city in department 78. Results reassuringly show that, albeit for a reduction in precision, this sample change does not affect coefficient estimates.

$V_{im}(t, r = 0)$ is, however, the relevant object to use in the utility of applying to city m since it literally represents the option value of putting m on the waitlist as far as social housing offers are concerned. Additionally, conditional on parameter estimates, assumption 6 has no bite on counterfactual results because those simulate cases where applicants rank only one city (see Appendix for a discussion on this simplification).

2.7 Results

These results use data for the Yvelines department (henceforth department 78) which corresponds to the region of Versailles, in the South-East of Paris. Further results will include other French regions.

2.7.1 Score

Table 2.6 show estimates for the score. These confirm that the score increases with waiting time, especially at the beginning of the wait. The coefficients corresponding to the number of rooms per household member show that richer cities expect more spacious homes for a given household size.

Figure 2.4 shows the shape of the impact of the effort rate (e.g the ratio of rent to household income) on the score, net of interactions. Figure 2.5 and 2.6 plot the score on household income for two different units. Very low incomes seem strongly penalized, although estimates show that poorer people are less disadvantaged for units in poorer maintenance conditions. Consistently with a penalty for precarious households, single mothers and people with non-permanent jobs are at a disadvantage in the allocation rule. This is unexpected for social housing, although it was already flagged in reports by government institutions controlling social housing (Agence nationale de contrôle du logement social). Additionally, and also consistent with reports from public institutions, being French increases an applicant's score - less so when the unit is in a poorly-maintained building.
The penalty for offer rejections seem to exist in big cities only - although we should account for the fact that very few applicants have a past rejected offer. Authorities also seem to favor applicants who are not already living in social housing.

2.7.2 Other components of the value function computation

CCP results are shown in table 2.9. Results for the unit arrival rate are found in table 2.8, and those for the rate of voluntary departure are in table 2.7.

2.7.3 Match values and value functions

Figure B.1 shows how the match value $\Gamma_{ij}(t=0)$ of a given unit evolves with its rent. Reassuringly we see a downwards relationship between match value and rent. Figure B.2 shows the relationship between match value and neighborhood quality, as measured by the rate of distinctions (cum laude or above) in the Year 9 national examination for students at the middle school nearest to the unit. Although applicants with lower incomes seem to have a preference for a higher quality school, higher incomes unexpectedly display the opposite pattern. Figure B.3 reassuringly shows that households of different size target a different number of rooms.

Figure B.4 plots the histogram the value functions for different waiting times - distinguishing the value functions with one or no rejections. As expected, value functions seem higher without past rejections. As the waiting time increases, both value functions converge to zero - the graph for t=5475 days is not shown because all values are zero, since the time on the waitlist is limited to 15 years for everyone. This is what we would expect from the model, and is therefore reassuring.
Figure 2.7 compares the distribution of valuations for cities according to their median income. It

appears that the valuation for poorer cities is higher on average to that of richer ones. Results uncover heterogeneity across applicants, whereby poorer applicants seem to have a comparatively higher valuation for poorer cities.

Expectations over the chance to receive a unit form a big part of an applicant's value for being on the waitlist of a given city. The left subplot in figure 2.8 plots the distribution of match values Γ_{ij} in cities with low or high proportions of social housing. The plot indicates that units in cities with low proportion of social housing seem to be preferred on average. However, the reverse ordering prevails when it comes to applicants' continuation values for being on the waitlist of a city. The right subplot shows the distribution of value functions $V_m(t=0, r=0)$ across those two types of cities, and here cities with a high percentage of social housing seem to be preferred. The reason for this lays in the higher arrival rate of units in cities with a large social housing sector, making it more likely to be offered a unit.

2.7.4 City choice

Table 2.10 shows estimates for the value of a city U_{im}. These estimates are calculated without adjusting for choice-based sampling. Future versions will incorporate the robustness strategies mentioned in section .

The table shows that $V_{im}(t=0, r=0)$ weighs positively in a household's utility for a city's waitlist. This validates the use of $V_{im}(t=0, r=0)$ as a measure of a household's social housing prospects in city m. This positive weight is consistent across specifications, and seems little affected by changes in specification once controlling for basic city characteristics and for whether the applicant already lives there.

Households prefer applying to the city where they currently live. Older households are more averse to applying far from where they live.

Among the determinants of decision to apply, a city's median income is the most important. Applicants seem to avoid very poor or very wealthy cities, as the relationship is bell-shaped. Richer applicants and French ones have a higher preference for wealthier cities than the rest.

Additionally, applicants favor cities with a lower proportion of social housing and a growing population.

I use specification from Model 4 for the rest of the analysis.

In Table B.1, I investigate whether the value of being on the waitlist $V_{im}(t=0, r=0)$ weighs similarly on the decision of all applicants, and for all cities. Results indicate that "strategic" considerations over future offers matters less when applicants consider applying where they live and when considering going far from home. It appears that applicants account most for their social housing potential in the city when considering cities around where they currently live.

Table B.1 additionally shows that wealthier applicants and younger applicants weigh their social housing prospects more than others when making their decision. This could come from differences in information about their social housing prospects, or from relative importance of other factors in their choice of where to live.

2.7.5 Welfare

In this section, I define welfare as follows:

$$\text{Welfare}_i = E[\max_m U_{im}] = \log\big(\sum_m \exp(U_{im})\big)$$

Figures 2.9, 2.10 and 2.11 plot differences in welfare across applicant groups, as a percentage of the rightmost group's welfare. Figure 2.9 shows that welfare seems to be relatively higher for wealthier households. Interestingly, figure 2.10 shows that French people are significantly better off in terms of welfare as defined. Single mothers have relatively lower welfare than the rest of the population, as shown in figure 2.11.

2.8 Counterfactuals

2.8.1 Equilibrium definition

The model equilibrium is defined as follows:

1. Optimality: agents maximize their lifetime utility

2. Consistent beliefs: π is consistent with equilibrium offer probabilities

3. $\tilde{P}_x(m)$ corresponds to the equilibrium rates at which applicants with characteristics x join city m's waitlist. It is the probability that city m is preferred to the outside option.

   $$\tilde{P}_x(m) = P[U_{im} > U_{i0} | x_i = x]$$

   where $U_{im} = \bar{U}_{im} + \eta_{im}$ is the utility of applicant i for applying to social housing in city m and $U_{i0} = \eta_{i0}$.

4. Steady-state balance condition: let $\mu^\star_m(x, t, r)$ be the equilibrium probability distribution of characteristic x, waiting time t and number r of past rejections in the waitlist of city m, and

$p^\star_m$ the proportion of applicants who apply to city m. Assuming x is discrete, $p^\star_m$ writes:

$$p^\star_m = \sum_{x \in \mathcal{X}} h_x \, \tilde{P}_x(m)$$

where h_x is the rate at which applicants of type x arrive on the market before making city choices.

The proportion of applicants with characteristics $(x, t = 0, r)$ in m is such that it satisfies the following:

$$\mu_m(x, t = 0, r = 0) = \frac{\tilde{P}_x(m) h_x}{p^\star_m} \tag{2.21}$$

$$\& \quad \mu_m(x, t = 0, r = 1) = 0 \tag{2.22}$$

For each city m, for each discrete characteristics x, $\mu^\star_m$ satisfies:

$$\frac{\partial \mu_m(x, t, r = 0)}{\partial t} = -\mu_m(x, t, r = 0)\kappa_m(x, t, r = 0) \tag{2.23}$$

$$\frac{\partial \mu_m(x, t, r = 1)}{\partial t} = -\mu_m(x, t, r = 1)\kappa_m(x, t, r = 1) + \mu_m(x, t, r = 0)\tilde{\kappa}_m(x, t, r = 0) \tag{2.24}$$

where $\kappa_m(x, t, r)$ is the probability for applicant with attributes (x, t, r) to leave the waitlist of city m, be it for assignment of by voluntary departure. More specifically, it is the probability that they leave the "pool of applicants" with characteristics (x, t, r) of that waitlist - in that sense an applicant with $(x, t, r = 0)$ who rejects an offer leaves her pool of the waitlist and joins the pool with characteristics $(x, t, r = 1)$. $\tilde{\kappa}_m(x, t, r = 0)$ is the probability for an applicant to receive an offer and reject it in city m, therefore joining the pool $(x, t, r = 1)$.

The departure probability $\kappa_m(x, t, r)$ is the probability to either receive an exogenous shock and leave the waitlist, or receive an offer. Applicants with one rejection $(r = 1)$ indeed leave the waitlist whenever they receive an offer: if they accept it they get assigned, if they reject it they depart from the waitlist. Applicants with no rejections $(r = 0)$ also leave the pool of applicants with attributes $(x, t, r = 0)$ from the waitlist whenever they receive an offer: either they accept and get assigned or refuse and thereafter belong to the pool $(x, t, r = 1)$. Therefore, $\kappa_m(x, t, r)$ writes as follows:

$$\kappa_m(x,t,r) = P\left[\text{applicant } i \text{ receives an exogenous shock and leaves} \,|\, x_i = x\right] +$$
$$P\left[(\text{unit } j \text{ becomes available}) \cap (s_{ij}(t,r) > s_{i'j}\, \forall i' \in \mathcal{I}^\star_m \,|\, x_i = x) \mid j \in J_m\right] +$$
$$P\Big[\cup_{R\in\mathcal{R}} \Big(\Big(\cup_{m'\in R} \Big[(\text{unit } j \text{ becomes available}) \cap (s_{ij}(t,r) > s_{i'j}\, \forall i' \in \mathcal{I}^\star_{m'}) \mid j \in \tilde{J}_{m'}(x,t,r)\Big] \mid R_i = R\Big)$$
$$\cap\, (R_i = R)\Big) \,|\, x_i = x\Big]$$

where $\mathcal{I}^\star_m$ is the set of applicants in city m in equilibrium. Each applicant $i' \in \mathcal{I}^\star_m$ has attributes $(x_{i'}, t_{i'}, r_{i'})$. J_m the set of units in city m $s_{ij}(t,r)$ the score of applicant with attributes (x_i, t, r), and $\tilde{J}_m(x,t,r)$ is the set of such units that applicant would characteristics (x,t,r) would accept. R_i is the rank-ordered list of city that an applicant provides upon arrival on the market, and $\mathcal{R}$ represents the set of all possible ways to create such a list of ranked cities.

The last term in this expression, $P\left[\cup_{R\in\mathcal{R}} \ldots\right]$, refers to the probability that the applicant receives and accepts an offer from another city where she is on the waitlist.

Now let me write the expression defining $\tilde{\kappa}_m(x,t,r=0)$, the probability that and applicant with $r = 0$ rejects an offer in m. It writes:

$$\tilde{\kappa}_m(x,t) = \lambda_m \int \pi(t, r=0; z, x, w_m) \int P\big[\Gamma_{ij}(t) + \varepsilon < V_{i,m}(t, r=1)\big]\, dG(\varepsilon)\, dF_m(z) \tag{2.25}$$

2.8.2 Counterfactual analysis

The first set of counterfactuals I am interested in focuses on the impact of a change in the allocation rule on applicant welfare. I start by implementing a first-come-first-serve scoring rule, where only waiting time matters. I first compute the impact of a change in rule without accounting for applicants' response in their application behavior and then proceed to computing the equilibrium outcome of such a change. In order to compute the equilibrium outcome, I need to compute the probability of receiving an offer $\pi^\star$ in the new equilibrium. Equilibrium being in steady-state, this is the solution to a fixed-point computation detailed in Appendix B.3.

In order to make computations more tractable, I assume in counterfactuals that applicants can only list at most one city upon arrival on the market, and therefore can be on one city's waitlist maximum. Justifications for this simplification are detailed in Appendix B.3.3, but the intuition is that when applicants can be in multiple cities' waitlist at once, the composition of the pool of applicants in each city's waitlists depends on the probability of receiving an offer in another city. This itself depends on the initial rank-ordered list submitted by the applicant.

The algorithm to compute counterfactuals can be summarized as follows. For each iteration k:

1. For city m: compute $V_m(t,r,x,z,w_m)^k$ given beliefs π_m^{k-1}. This yields accept/reject strategies and therefore probabilities $\kappa_m^k(x,t,r)$ and $\tilde{\kappa}_m^k(x,t)$.
2. Compute utilities for each city $\{U_m(x)\}_{\forall m}$. this yields city choices and therefore $\{\tilde{P}_x(m)\}_{\forall m}$
3. Compute queue composition $\{\mu_m^k\}_{\forall m}$ & length $\{N_m^k\}_{\forall m}$
4. Compute $\{\pi_m^k(t,x,w)\}_{\forall m}$ using $\{\mu_m^k\}_{\forall m}$, $\{N_m^k\}_{\forall m}$ and $\{V_m(t,r,x)\}_{\forall m,\forall x,\forall t,\forall r}$
5. Repeat until convergence of $\{V_m(t,r,x)\}_{\forall m,\forall x,\forall t,\forall r}$, $\{\mu_m^k\}_{\forall m}$ and $\{N_m^k\}_{\forall m}$

Given that the algorithm to compute equilibrium counterfactual is computationally demanding, I simulate a department for the counterfactual analysis. I choose n cities which will be the only ones available in the department. I randomly sample applicants within the department and discretize their characteristics so that I have a fixed number of applicant types, which includes applicant income, age, citizenship, job contract, household size and composition, but also whether they currently live in social housing, and the distance between the city where they currently live and the n cities in the simulated department. I compute allocations under the benchmark scenario, keeping allocation rules as they are estimated, and compare it to various scenarios.

Results are computed for $n = 2$, where both cities display median values for almost[17] all characteristics except that one (henceforth called Impoverished) has a median income of $15,000$, placing her at the bottom of the distribution for department 78, and the other (henceforth called Affluent) has a median income of $40,000$ placing her at the top. Social housing units are selected at random from poorer/wealthier cities in the data, and the rate of arrival of units corresponds to the median for poor/wealthy cities in the data. I draw a random sample of 100 applicants and infer their types to compute simulations. The rate of arrival of each type corresponds to that in the data for the department.

I first simulate allocations under the current scoring rules, and then compare it various alternative ones. First-come-first-serve (FCFS) ranks applicants solely based on their waiting time, where a

[17] Cities are assumed to have a large population

longer wait is ranked higher. "No weight on waiting time" uses the current allocation rule and removes its dependency on time. "No weight on past rejections" removes the penalty for the first rejection, although supplementary rejections still lead to an exclusion from the waitlist. Last-come-first-serve (LCFS) incentivizes acceptance by ranking first applicants that have just arrived on the waitlist. "Linear waiting time" considers the current scoring rule but adds a linear premium on waiting time. "Favor the Poor" weighs negatively household income and considers no other applicant characteristic, nor waiting time. Future versions of the paper will consider the case where a rejection no longer leads to an exclusion and is not penalized.

Figures B.5 to B.11 show the average probability to receive an offer for applicants of various types as a function of their waiting time, number of past rejections and the city where they apply. This is the outcome from the algorithm and reassuringly corresponds to expectations for each scoring rule. As expected, when relaxing the dependence of the score on applicant characteristics, this probability is the same for everyone, as in the FCFS. In this scenario, offers are made only to applicants who have waited almost 10 years, the maximum amount of time allowed in the simulations.
Table 2.11 shows that tying scoring rules to waiting time as in a FCFS or LCFS leads to a substantial reduction in welfare. For the former, it is partly due to an increase in waiting times. This is also due to worse matching of households to social housing units, in a context where preferences are horizontal. This is visible in figure B.20, showing expected match values in each city, weighted by the probability to get an offer when it's available.
Adding weighting time linearly in the score has a negative impact on welfare. Figure B.13 and B.15 show the proportion of each type (x, t, r) in the waitlist of each city. We see that the portion of applicants with a long waiting time among the population who received and rejected an offer is high in the Linear waiting time scenario. This comes from the fact that the relative weight of waiting time among other characteristics determining the score get so high after a certain wait, and the demand is in such excess, that many applicants receive an offer after a very long wait.

Figures B.5, B.6, B.7 and B.8 show how the probability to receive a given available offer $\pi_{ij}(t, r)$ varies across scenarios for different applicants in both cities.

Figure B.22 shows that households are more likely to apply to the Impoverished city in most scenarios, due to the higher arrival of social housing units. This is no longer true in the LCFS.

The "Favor the Poor" scenario is of interest because it focuses entirely on rewarding applicants whose value for social housing is highest, as shown in figure B.18 which plots average match values Γ_{ij} across applicant groups. Lower income households tend to have a higher value for social housing unit, and therefore favoring them in the allocation rule should increase the average match value of a unit. This is what we see in Figure B.20, which compares the value of $E[\Gamma_{ij}\pi_{ij}]$ across applicant groups and counterfactual scenarios. This is higher for all groups except the rich in the "Favor the Poor" scenario, which explain why it increases welfare by more than 20%. Therefore, targeting low income households is overall beneficial on welfare for a planner that has a utilitarian perspective.

2.9 Conclusion

This paper aims at informing decision-makers in the choice of a selection rule to allocate social housing. It bridges a gap in the literature by providing a framework to estimate household preferences over social housing units, and modeling the decision-making process of an applicant to social housing, allowing for geographical relocation. Findings indicate that a first-come-first-serve mechanism as sometimes in use in the US is welfare decreasing for all applicants. Instead, results points towards favoring low-income households in the allocation to improve welfare.

The methods developed in this paper can be applied to other countries where social housing is allocated by take-it-or-leave-it offers at the local level, of which there are many in the OECD (Finland, UK, US...). Furthermore, the refined optimal-stopping framework I employ to elucidate preferences regarding social housing unit features boasts universal applicability in dynamic settings where declining an offer has downstream consequences. This is particularly relevant in constrained decision-making systems like certain unemployment benefit schemes (e.g. the Danish one), where the subsidy level is contingent upon the rejection of job offers and may even be revoked after a pre-determined number of such rejections.

This study successfully identifies the impact of a social housing allocation rule on household welfare. However, it focuses on the welfare obtained from the social housing sector, abstracting from other forms of government interventions. Future research should investigate the interplay between social housing allocation rules and other government subsidies like housing allowances, which often exist in conjunction with social housing.

Tables and Figures

Table 2.1: Comparing working sample (SyPLo) with universe of applicants (SNE)

Year	Variable	Value	
		SNE	SyPLo
2019	N	457132	103218
	Average rent	507	431
	Average size	64	63
2020	N	386232	95808
	Average rent	443	404
	Average size	64	62

Figure 2.1: Number of applicants in a city's waitlist

Department 78

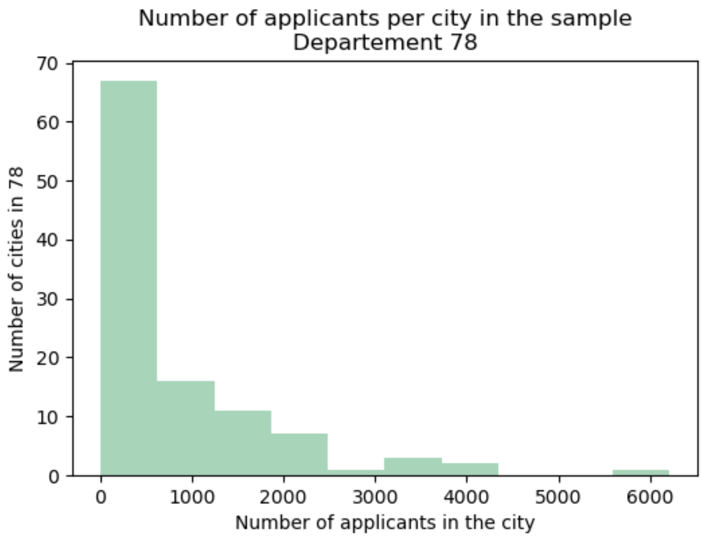


Figure 2.2: Number of applications per unit

Department 78

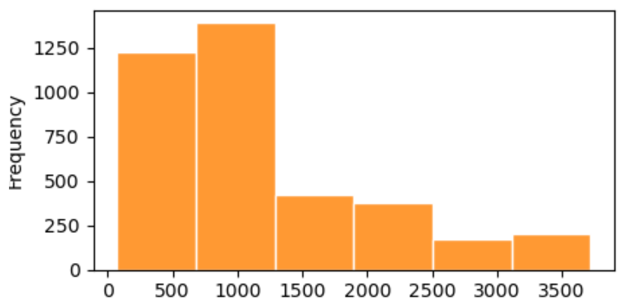


Table 2.2: Comparison of household characteristics across samples

	Variable	syplo(1)				sne(2)				univ(3)			
		Mean	25%	50%	75%	Mean	25%	50%	75%	Mean	25%	50%	75%
All observations													
	household size	2.34	1	2	3	2.24	1	2	3	1.83	1	1	2
	currently in soc. hous.	0.17				0.31				0.16			
	age*	44.95	34	43	54	45.72	34	43	55	59.14	45	59	73
	retired*	0.06				0.10				0.34			
	unemployed*	0.15				0.14				0.11			
	French*	0.67				0.78				0.85**			
	non European*	0.28				0.18				0.04**			
	European*	0.05				0.04				0.04**			
	N syplo	441,896											
	N sne	3,475,221											
	N univ	35,307,696											
Department 78													
	household size	2.47	1	2	3	2.36	1	2	3.00	2.01	1	2	3
	currently in soc. hous.	0.23				0.33				0.22			
	age*	44.88	35	43	53	44.49	34	42	53	58.07	45	57	71
	retired*	0.06				0.07				0.29			
	unemployed*	0.12				0.11				0.09			
	French*	0.60				0.69				0.78**			
	non European*	0.33				0.26				0.06**			
	European	0.06				0.05				0.06**			
	N syplo	15,999											
	N sne	86,890											
	N univ	717,173											

* Values refers to the household member who is the primary tax filer ("référent fiscal" in French).
** Since tax files do not record an individual's citizenship, this refers to the country where they were born
(1) Sample of applicants pre-selected by the State to be assigned a unit
(2) Universe of social housing applicants
(3) Universe of tax filers

Table 2.3: Average distance between current city and cities in the department among applicants

Department 78

	mean	std	25th percentile	50th percentile	75th percentile
All SH applicants to the department	36.42	274.90	11.68	19.69	29.12
Low income (<1k/month)	32.91	236.35	12.15	20.37	29.84
Medium income (1k-2k/month)	32.83	225.76	11.68	19.72	29.05
High income (>2k/month)	43.18	345.94	11.41	19.29	28.90
Non EU citizenship	22.74	25.84	12.11	20.16	29.03
EU citizenship	21.82	24.73	11.21	19.25	28.48
French citizenship	46.55	360.87	11.37	19.44	29.26
Single mother household	33.50	242.82	11.34	19.28	28.65
Not single mother household	39.22	302.39	11.89	20.03	29.50

Table 2.4: Average income among applicants

Department 78

	mean	std	25th percentile	50th percentile	75th percentile
All	1819.90	890.87	1222.0	1682.0	2273.00
Low income (<1k/month)	686.84	261.40	520.0	774.0	900.00
Medium income (1k-2k/month)	1517.68	268.31	1300.0	1515.0	1740.00
High income (>2k/month)	2790.46	687.65	2240.0	2590.0	3186.00
Non EU citizenship	1624.55	752.57	1143.5	1515.0	2019.00
EU citizenship	1868.26	844.45	1298.0	1753.0	2388.25
French citizenship	1938.10	952.33	1299.0	1787.0	2440.00
Single mother household	1745.83	827.61	1197.0	1627.0	2140.50
Not single mother household	1890.76	942.01	1260.0	1737.0	2417.75

Table 2.5: Probability for an offer to be accepted

	Variable	Value	Acceptance Rate (%)
Applicant characteristics			
	age	Low	62.36
		High	50.80
	citizenship	EU	61.68
		French	49.41
		Non_EU	67.47
	household size	1	54.93
		2-4	54.44
		5+	54.87
	income	Low	63.95
		High	49.20
	job contract	permanent	57.98
		unemployed	57.19
		student	68.52
		other	50.73
Unit characteristics			
	building age	Low	56.32
		High	57.15
	building maintenance state	Good	55.46
		Normal	55.37
		Poor	57.62
	neighborhoold school quality	Low	57.49
		High	55.66
	rent	Low	60.25
		High	52.91
	size	Low	59.09
		High	54.03

Figure 2.3: Number of cities in applicants' rank-ordered lists

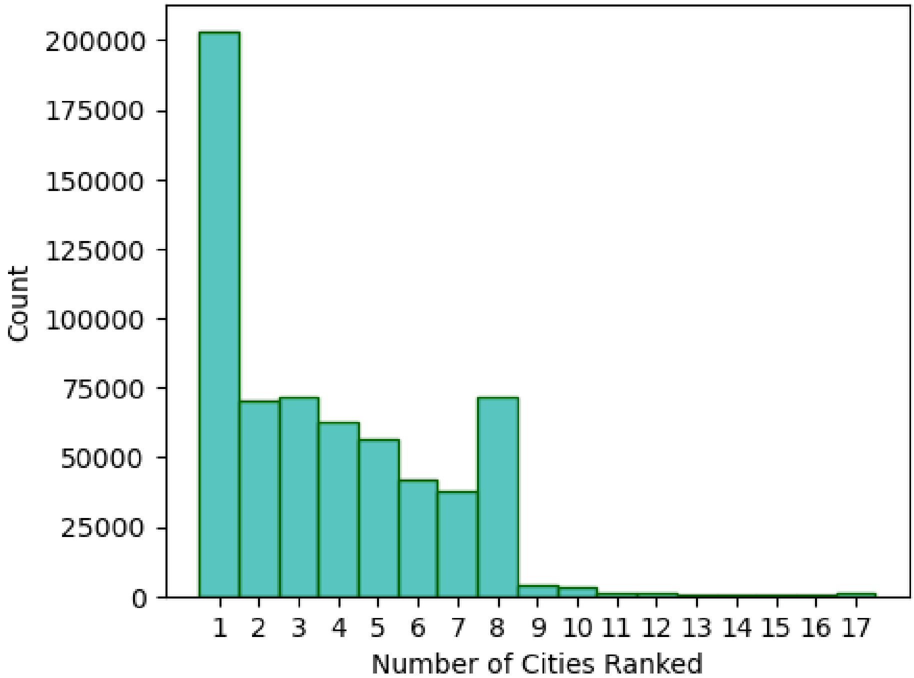


Figure 2.4: Impact of effort rate on the score, net of interactions

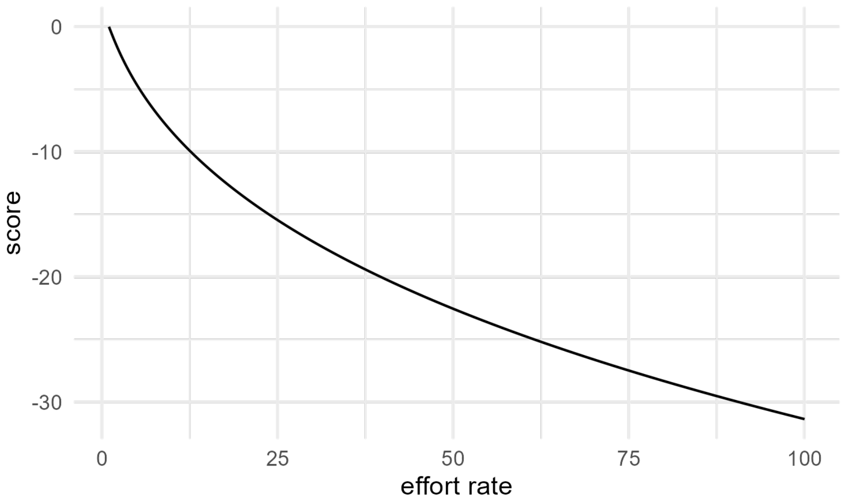


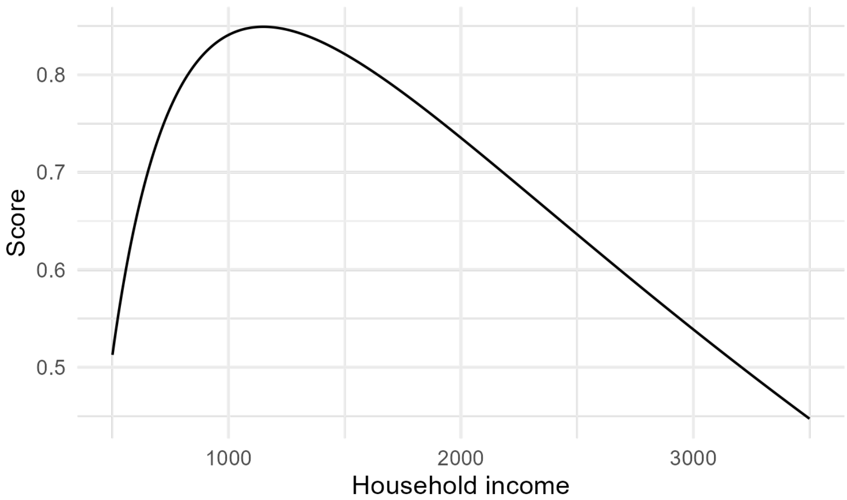
Table 2.6: Score Results

waiting time (months)	
spline 0 - 12	0.054***
	(0.004)
spline 12+	0.001
	(0.001)
waiting time × big city	0.001
	(0.001)
waiting time × quality of nearest secondary school	−0.003*
	(0.002)
effort rate	
log(effort rate)	−1.453***
	(0.321)
log(effort rate)ˆ2	−0.754***
	(0.108)
log(effort rate)ˆ3	−0.089***
	(0.014)
effort rate × city median income	−0.257*
	(0.139)
single mother	
single mother	−0.241**
	(0.103)
single mother × city median income	−0.0189
	(0.031)
past rejections	
past rejection	0.005
	(0.036)
past rejection × big city	−0.052
	(0.050)
citizenship	
French	0.194***
	(0.038)
French × (building maintenance = poor)	−0.130***
	(0.048)
currently in social housing	
already in soc. hous.	−0.084***
	(0.028)
soc. hous. × city median income	0.021
	(0.042)
number of rooms per person	
nb rooms/hh size	12.230***
	(0.340)
nb rooms/hh size ˆ2	−7.562***
	(0.142)
nb rooms/hh size × (building maintenance state = new)	0.771***
	(0.205)
nb rooms/hh size × (building maintenance state = poor)	−0.215*
	(0.127)
Controls: applicant age, job contract type and interactions	
Pseudo R-squared	0.199
Observations	10,678
Log Likelihood	−20,518.350

Note: Quality of nearest secondary school refers to the % of students who got a cum laude distinction or above at the “Brevet des Collèges” (Year 9 examination) in the school closest in distance to the social housing unit. The reference category for the variable “building maintenance state” is ”normal”. “City median income” is the median yearly income in the city expressed in 10,000.

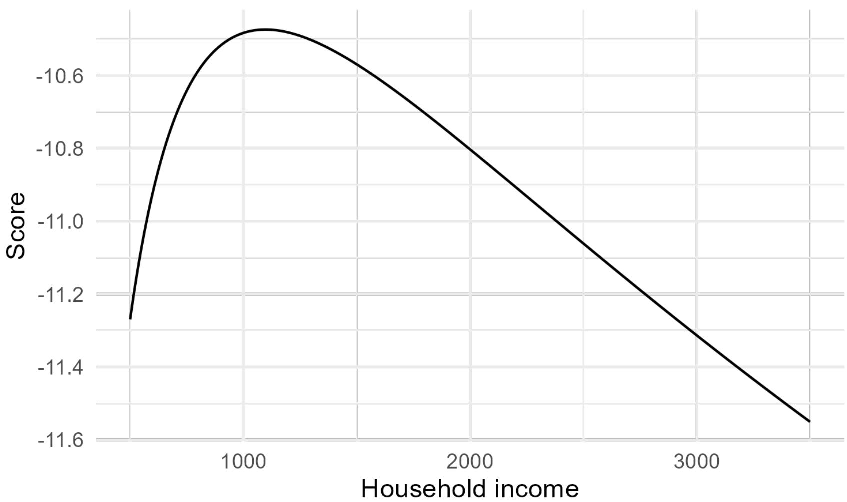
See Online Appendix for a more detailed table.

Figure 2.5: Impact of income on the score unit



Unit with rent = 350€, poor maintenance, in small city with median yearly income = 16990

Figure 2.6: Impact of income on the score unit



Unit with rent = 500€, maintenance = new, in big city with median yearly income = 40000

Table 2.7: Rate of voluntary departure results ($\delta_i(t)$)

	δ
age (years)	
age	-0.027***
	(0.004)
age^2	0.000***
	(0.000)
monthly income per consumption unit (in thousands)	
income	0.012*
	(0.005)
income^2	-1.51 e^{-08}
	(1.029 e^{-08})
citizenship - ref: French	
EU	-0.362***
	(0.056)
Non EU	-0.722***
	(0.031)
job contract type - ref: unemployed	
other	-0.153**
	(0.049)
permanent	0.319***
	(0.042)
student	0.381***
	(0.110)
woman	-0.072**
	(0.026)
already in soc. hous.	-0.195***
	(0.032)
constant	-8.114***
	(0.104)
ancillary param	7.6 e^{-06}
	(1.15 e^{-05})
N	27141

* p<0.05, ** p<0.01, *** p<0.001

Table 2.8: Rate of unit arrival (λ_m)

	λ
social housing increase in the city - ref: negative or 0	
low (<20% increase 2012-2019)	-0.006
	(0.039)
high (≥20% increase 2012-2019)	-0.003
	(0.046)
social housing prevalence in the city - ref: < 20%	
medium (20-40%)	-0.005
	(0.041)
high (>40%)	0.091
	(0.051)
population in 2017	
pop	$7.92e^{-6}$
	$(6.12e^{-6})$
pop^2	$-1.38e^{-10}$
	$(1.80e^{-10})$
pop^3	$1.28e^{-15}$
	$(1.36e^{-15})$
median income	
median income	0.00005
	(0.0000601)
median income ^{2}	$-2.01e^{-9}$
	$(2.24e^{-9})$
median income ^{3}	$2.36e^{-14}$
	$(2.64e^{-14})$
constant	-0.291
	(0.505)
N	5910

* p<0.05, ** p<0.01, *** p<0.001

Table 2.9: CCP Results

monthly income/uc spline (thousands of €)	
0 - 1.458	0.0101
	(0.0699)
1.458 - 2.09	-0.570***
	(0.0961)
2.09+	0.418***
	(0.0629)
job contract type - ref: other	
permanent	0.116***
	(0.0299)
unemployed	0.129**
	(0.0403)
citizenship - ref: French	
EU	0.203***
	(0.0480)
Non EU	0.244***
	(0.0256)
waiting time (months)	
0 - 26	-0.000655
	(0.00178)
26+	-0.00280
	(0.00188)
nb rejections	0.108***
	(0.0246)
rent spline (hundreds of €)	
0 - 541	-0.00367
	(0.0176)
541+	-0.0386
	(0.0243)
floor - ref: medium (2-4)	
low (0-1)	0.0329
	(0.0646)
high (4+)	0.110*
	(0.0520)
Controls: rent, number of rooms per person, building age, building maintenance state, unit in urban priority area, woman,large city, social housing increase in the city, social housing prevalence in city, and interactions.	
N	16,693

Note: "Quality of nearest secondary school" refers to the % of students who got a cum laude distinction or above at the "Brevet des Collèges" (Year 9 examination) in the school closest in distance to the social housing unit. "Urban priority areas" are a translation of "Quartier Prioritaire de la Ville" (QPV). They are neighborhoods identified as poorer and less developed. "Social housing increase in the city" refers to the amount of social housing built in the city between 2012 and 2019 as a percentage increase. "Social housing prevalence in the city" refers to the percentage of housing units that are social housing in the city in 2017.

See Online Appendix for a more detailed table.

Table 2.10: City choice results (U_{im})

VARIABLES	Model 1	Model 2	Model 3	Model 4
$V_{im}(t=0, r=0)$	13.42***	8.842***	8.043***	8.065***
	(0.0694)	(0.0758)	(0.0768)	(0.0770)
distance to current city (log(km))		0.536***	1.230***	0.240***
		(0.00560)	(0.0138)	(0.0202)
distance to current city (log) ^{2}			-0.213***	
			(0.00319)	
distance to current city (log) × applicant age				-0.00105***
				(0.000408)
distance to current city (log) × (job = permanent)				-0.0161
				(0.0147)
distance to current city (log) × (job = unemployed)				-0.0683***
				(0.0207)
same city		1.415***	1.865***	2.557***
		(0.0111)	(0.0144)	(0.0328)
same city × (job = permanent)				0.0700**
				(0.0331)
same city × (job = unemployed)				-0.0952**
				(0.0475)
city median income (in 10,000)		0.930***	3.802***	4.493***
		(0.00518)	(0.0273)	(0.0304)
city median income^2			-0.617***	-0.737***
			(0.00529)	(0.00552)
city median income × same city				-0.770***
				(0.00922)
city median income × applicant income				0.0743***
				(0.00694)
city median income × (citizenship = French)				0.0358***
				(0.0118)
city population (in 1,000)		-0.00728***	-0.00263***	-0.000933***
		(0.000126)	(0.000343)	(0.000341)
city population^2			-4.81e-05***	-5.31e-05***
			(2.16e-06)	(2.13e-06)
city population growth		1.130***	0.263***	0.288***
		(0.0343)	(0.0350)	(0.0348)
social housing prevalence (proportion)		2.817***	-3.667***	-3.262***
		(0.0337)	(0.135)	(0.147)
social housing prevalence^2			6.554***	6.961***
			(0.207)	(0.207)
social housing prevalence × (job = permanent)				-0.381***
				(0.0778)
social housing prevalence × (job = unemployed)				-0.280**
				(0.110)
Observations	187,656	187,373	187,373	187,373
Number of groups	25,399	25,399	25,399	25,399

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 2.7: Value function comparisons across city wealth

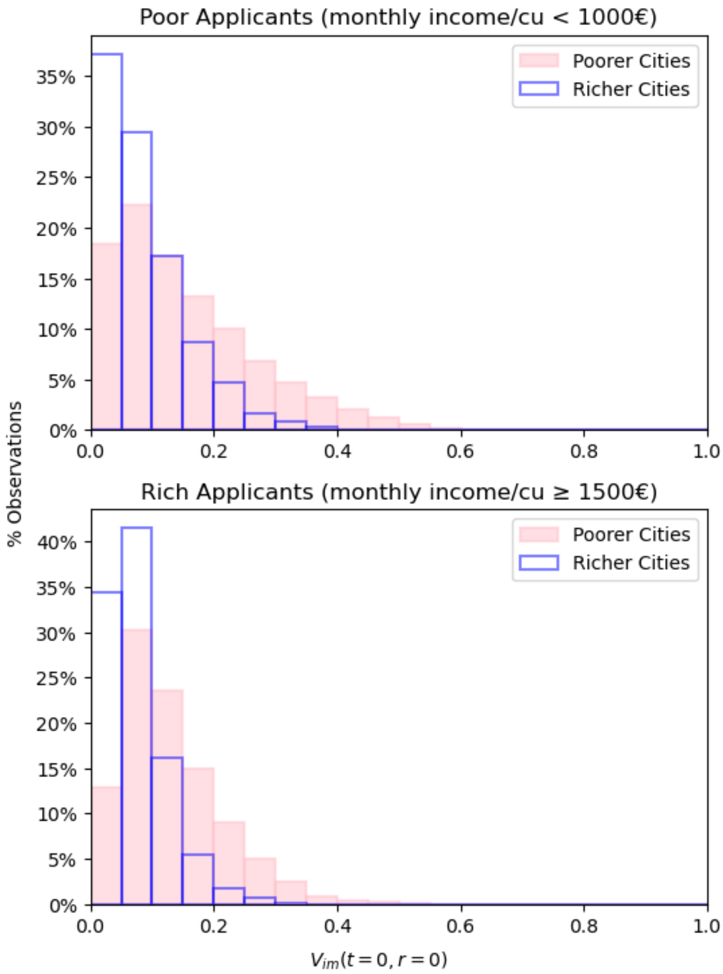
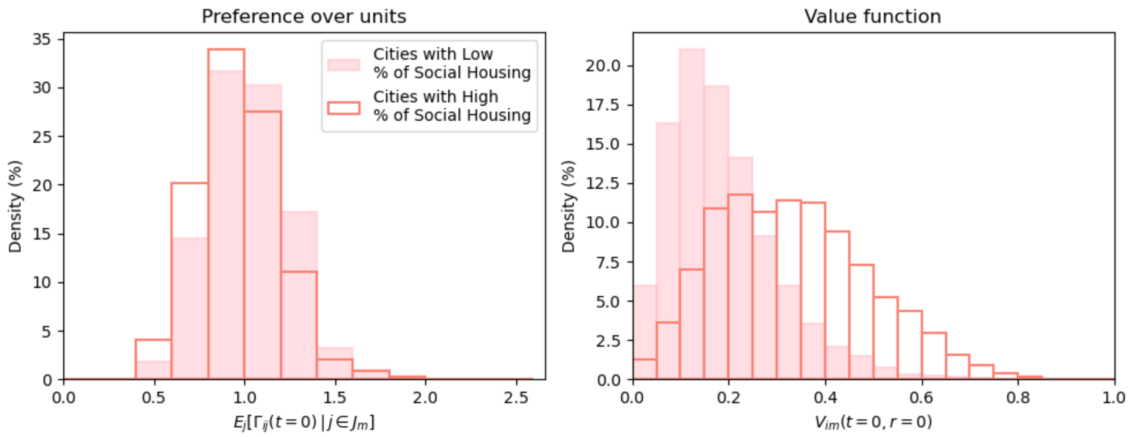


Figure 2.8: Compare match values and value functions



Values are computed for all applicants in the department of Yvelines (78). In the left subplot, each histogram pools all units within the group of cities. In the right hand side, there is just one value for each city-applicant pair.

Figure 2.9: Welfare comparisons across applicant income

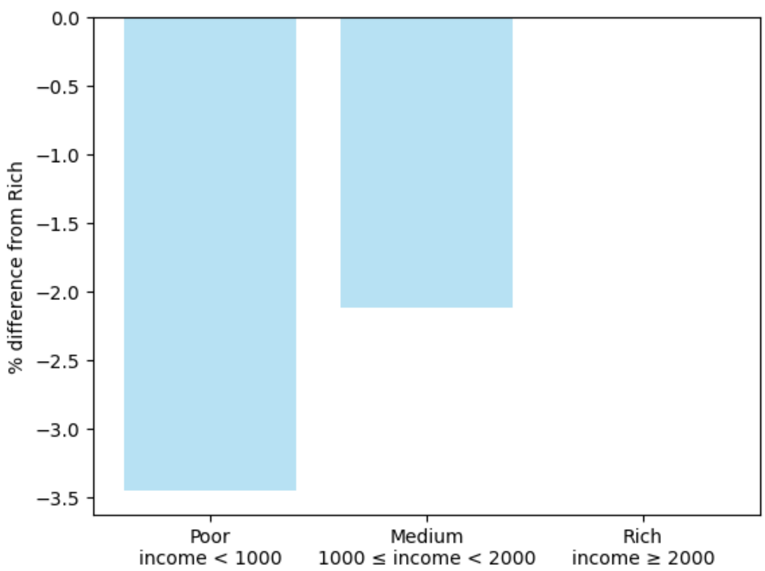


Figure 2.10: Welfare comparisons across applicant citizenship

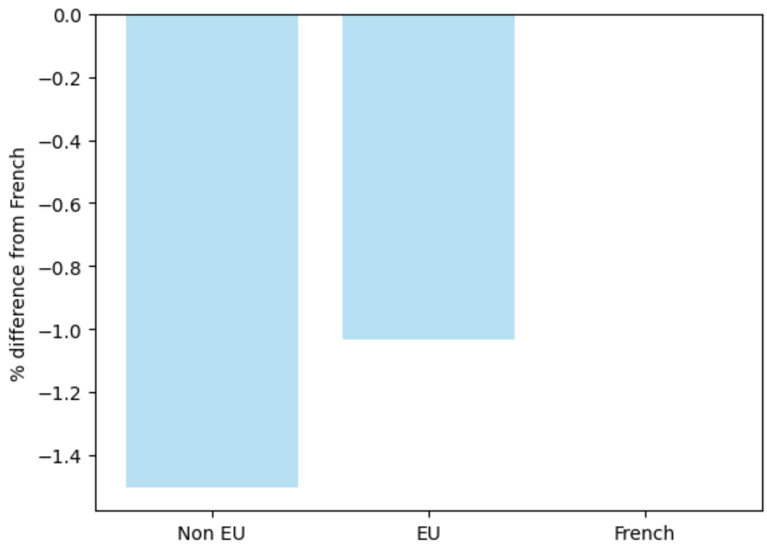


Figure 2.11: Welfare comparisons for single mothers

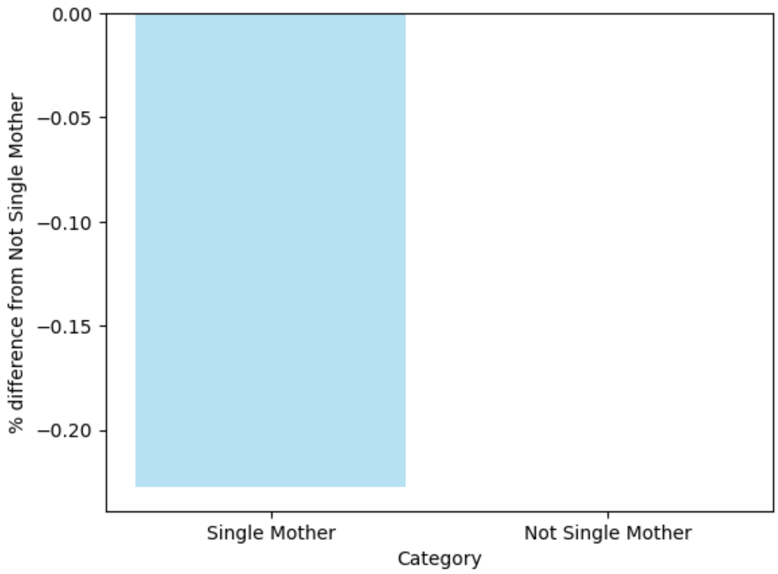


Table 2.11: Impact of a change in allocation rule on welfare

Equilibrium results (in % ΔW)

hh group	FCFS	no waiting time	no rejection	LCFS	linear waiting time	favor the Poor
all	-19.1687	0.0423	0.0269	-16.2894	-3.6666	20.3185
low income	-19.0647	0.0161	0.0266	-15.9931	-3.4206	53.7705
high income	-19.2978	0.0747	0.0272	-16.6570	-3.9717	-21.1802
non EU citizen	-22.0511	-0.0330	0.0326	-19.2290	-3.2523	35.9595
EU citizen	-23.3170	0.0931	0.0327	-19.8196	-4.5813	22.8745
French citizen	-16.4631	0.0817	0.0220	-13.6676	-3.7589	9.5524
single-mother hh	-18.5062	0.0359	0.0271	-15.6230	-3.4924	21.8723
non single-mother hh	-20.0946	0.0511	0.0266	-17.2207	-3.9099	18.1472

Chapter 3

The impact of local political cycles on the allocation of social housing

3.1 Introduction

The allocation of public goods is often subject to accusations of distortion by decision-makers for political reasons. This issue is particularly prevalent in the case of local public goods whose allocation is decentralized and opaque, such as social housing. In France, where social housing represents 18% of the housing stock and 60% of the population is eligible to apply, municipality officials have significant influence over the attribution of social housing units. The process is frequently at the center of public debate due to its opacity, and local politicians are sometimes accused of abusing their power to strategically direct social housing units to individuals that would otherwise not receive any[1].

In this paper, I investigate whether French city officials opportunistically distort social housing allocations when elections are coming. Schmutz and Verdugo (2023) document the fact that the political leaning of French mayors affects the way in which social housing is allocated. Specifically, they find that left-wing mayors tend to allocate social housing to immigrants more than others. Their results are compatible with left-wing mayors being ideologically motivated by immigrant welfare, although some of their evidence points towards a strategy to gain voters. I complement their findings by investigating whether mayors use social housing opportunistically by modifying their behavior when elections are coming. My objective is to understand whether social housing allocations follow a political cycle, like other public goods have been shown to (Brender and Drazen (2005), Alt and Lassen (2006)). I make use of detailed data on social housing applications and assignments to identify whether the 2020 French municipal elections cause changes in the population receiving social housing units. I use variation in time to elections as well as in electoral competition to identify how the characteristics of assigned applicants are impacted by upcoming elections.

The theory that incumbent politicians may manipulate public policy to enhance their re-election prospects was first proposed by Nordhaus (1975) who introduced political cycles. This hypothesis has spawned extensive research confirming the impact of political cycles on public goods provision, as evidenced in studies by Brender and Drazen (2005) and Alt and Lassen (2006). This body of work primarily examines distortions in levels of public revenue or expenditure, with some studies identifying a link between these distortions and increased re-election probabilities for incumbents, as found by Balaguer-Coll, Brun-Martos, Forte, and Tortosa-Ausina (2015). Other research focuses on the composition of public expenditures rather than their overall level, suggesting that

[1] Le Figaro "À Paris Habitat, l'attribution d'un logement cossu à une élue du CSE déclenche une guerre syndicale" (2023), La Voix du Nord "Du favoritisme dans l'attribution des logements sociaux dans la métropole lilloise? On démêle le vrai du faux"

local politicians strategically target public spending to benefit their electoral base, as explored by Drazen and Eslava (2010) and evidenced further by Ejdemyr, Kramon, and Robinson (2018) which showed politicians sometimes strategically target the assignment of local public goods before elections. In the context of social housing, where new construction is a long-term endeavor, local officials might attempt to sway voters in the short-term by modifying allocation rules, potentially favoring certain demographic groups to align with the political ideology of their base or strategically position voters within electoral districts.

Few studies have examined how the allocation of local public goods through matching markets is influenced by political cycles. Drometer and Méango (2020) demonstrate that the approval rates for naturalization requests in the U.S. can be swayed by electoral concerns. In the realm of social housing, since the overall number of assigned units is unchanged in the short term, the focus shifts to who receives the available units rather than the number of units provided. Gay (2017) and Eriksen and Rosenthal (2010) show that politics and voting behavior affect the distribution of affordable housing credits in the US (the LIHTC program). However, they do not investigate whether these partisan motives can lead to cyclical allocation patterns.
This paper complements these findings by analyzing whether affordable housing is subject to a political cycle around municipal elections. Its focus on a market where public officials allocate out housing units through a waitlist system in novel in this literature. Yet, such a process is common across developed countries that have social housing[2], and better understanding its determinants is key to optimizing the allocation. As such, this paper also contributes to the strand of literature aiming at understanding the factors influencing social housing allocations by housing authorities (Cai (2018), Nelson, Borg, Nieuwenhuis, and Alm (2023), and Botbol (2024)).

Understanding whether local politicians distort the allocation of local public goods for electoral motives is crucial to the debate over the centralization versus decentralization of local public goods provision. Theories by Besley and Coate (2003) and Joanis (2014) suggest that centralization might reduce the potential for partisan distortions by distancing decision-making from local political pressures. However, centralizing the allocation of local public goods could also lead to a disconnect from local contextual needs, potentially resulting in less effective governance. Answering this question is particularly pressing to social housing policy since recent announcements by French government officials indicate a willingness to further decentralize its allocation to mayors[3].

[2]Social housing represents 6% of the housing stock in OECD countries, and up to 20% in some countries like Austria and the Netherlands.

[3]Public Senat, "Attribution des logements sociaux : Gabriel Attal annonce un pouvoir renforcé des maires" January 2024, https://www.publicsenat.fr/actualites/politique/attribution-des-logements-sociaux-gabriel-attal-annonce-un-pouvoir-renforce-des-maires

My empirical analysis uses both an event study and a difference in difference framework to study the dependency of assignee characteristics on time to municipal elections. First, I compare assigned applicants across time, controlling for municipality specifics and time trends, and allowing for heterogeneous treatment effects depending on the mayor's political leaning. Crucially, I follow Aidt, Veiga, and Veiga (2011) in assuming that municipalities where the winner candidate gets a shorter margin are more affected by elections than others, and make use of this variation to identify the effect of elections. My results do no show an effect of the local political cycle on the allocation. I then run a difference-in-differences estimation where some municipalities are assumed to be unaffected by local elections since they are mostly run by state officials. I test this assumption by comparing allocation criteria in and out of election periods for these state-run units, using data on all offered issues for these units, included those that were rejected. My results confirm that of the event study in that I do not find evidence of a distortion of allocations around election times. However, they support the fact that a mayor's political leaning strongly influences who receives social housing - a finding that corroborates results from Schmutz and Verdugo (2023).

Evidence of mayor electoral opportunism in social housing allocations would advocate for more centralization in the allocations. Although my results do not support this claim, the fact that they confirm a strong impact of a mayor's political views on who receive social housing could lead policymakers to wish for more centralization to achieve a more uniform allocation of social housing across cities.

3.2 Context

Social housing plays a crucial role in France, housing approximately 18% of French households across some 5 million units, thereby accommodating around 12 million people[4]. Predominantly configured as apartments—comprising 84% of the total—social housing also includes a smaller proportion of houses, which make up 16%[5]. The distribution of these units is not uniform across the country, varying from 10 to 25% of the housing stock by region[6]. As of September 2023, nearly half of the French population has either resided in social housing at some point or continues to do so[7].

The demographic group served by social housing typically has a lower average income compared

[4] "Les Chiffres Clés du logement social", Union Sociale pour l'Habitat, August 2023
[5] "Les Chiffres Clés du logement social", Union Sociale pour l'Habitat, August 2023
[6] InseeAnalyses #110, Hauts-de-France, 2020
[7] "Chiffres et statistiques du logement social", ministère de la Cohésion des territoires et des Relations avec les collectivités territoriales, September 2023

to the private rental market. For instance, in 2017, the poverty rate among social housing tenants stood at 35%, in contrast to 23% within the private rental sector[8]. Eligibility for social housing is determined by two main criteria: lawful residency in France and an income below certain legal thresholds, which are revised annually and vary by region and household size. For example, a single person in Paris in 2022 needed an income under €24,316 to qualify. In 2021, approximately 66% of the French population met the eligibility requirements for social housing[9]. The supply of social housing does not meet demand: as of the end of 2022, about 2.4 million households were on the waitlist, yet only about 480,000 were allocated housing during that year [10].

Applications for social housing can be submitted either online or in person at a developer's office. Applicants must list their preferred cities in order of priority and are then placed on a waitlist for all listed locations, eligible to receive offers from any of them. Applications must be renewed annually with the requisite documentation, failing which they are removed from the list.

The development and management of social housing units are overseen by developers who fund new projects through subsidies from various sources including the state, municipalities, workers' associations, and occasionally large corporations. In return, these funding entities, known as Sponsoring Allocators, receive a certain quota of housing units which they distribute based on specific needs within their constituencies. Sponsoring Allocators have the discretion to select occupants from the pool of applicants, provided they meet the established criteria such as income limits and family situation.

When a unit becomes available, the Sponsoring Allocator shortlists a few applicants, and their selection is reviewed by a committee that includes representatives from the state, the municipality, and the housing developers. The top-ranked applicant is then notified and given a few days—typically up to a week—to accept or reject the offer. If declined, the offer passes to the next in line. Should the list be exhausted without a successful allocation, the committee reconvenes to draw up a new list.

While the overall rules governing allocations lack transparency, compounded by layers of local directives, average waiting times by municipality for recent years are typically accessible to applicants. Prior research has endeavored to decode and quantify these opaque regulations (Bonnal, Boumahdi, Favard 2012; Bourgeois 2018), shedding some light on the intricate process of social housing allocation in France.

French municipal elections, held every six years, are based on a two-round system. Voters elect

[8] Insee Références "Les conditions de logement en France", edition 2017

[9] "Mieux connaître la demande de logement social pour mieux orienter les politiques publiques" USH for the 81st HLM Congress, 2021

[10] "Les Chiffres Clés du logement social", Union Sociale pour l'Habitat, August 2023.

municipal councilors who, in turn, choose the mayor. In municipalities with fewer than 1,000 inhabitants, a single-round, plurality-at-large voting system is used, with voters selecting as many candidates as there are council seats. Since this system makes it difficult to measure electoral competition, and because social housing is almost never located in such small municipalities, I focus on municipalities with at least 1,000 inhabitants. In larger municipalities, a two-round proportional representation system with a majority bonus is employed. If no list secures an outright majority in the first round, a second round is held. The list with the highest number of votes in the second round receives a majority of the council seats, with the remainder distributed proportionally among all lists passing a threshold.

3.3 Data

3.3.1 Sources

My primary dataset originates from the "Système National d'Enregistrement" (SNE), which is the mandatory portal for all social housing applications in France. Applicants initiate their application by submitting a comprehensive form to the SNE, which creates a file that continuously updates to reflect any accepted social housing offers. These records require annual renewal, and failure to comply results in the withdrawal of the application. My analysis spans the entire duration from when an application is entered to when it exits the system, regardless of whether the applicant secures a social housing unit or exits due to non-renewal. As part of the application process, households must select a list of preferred municipalities where they wish their application to be considered; this list, along with other information, is maintained within their SNE file.

The dataset at my disposal contains exhaustive information collected at the time of application, including the list of municipalities each applicant has selected. While names are omitted for privacy reasons, the dataset is expansive, encompassing demographic details (such as age and income, including employment contract information), household size, and geographic variables (current city of residence and employment). It also includes the dates of initial submission and either removal or assignment, enabling the calculation of waiting times. Additionally, I have data on the characteristics of any accepted offers, which cover apartment size, layout, energy efficiency, and rent. Information about the housing unit's building, including the number of apartments, age of construction, maintenance status, and geocoordinates, is also available.

One limitation of the SNE is its failure to record offers that applicants reject. This means that I do not observed the first applicant that was chosen for a given unit if she rejected, which prevents me from properly estimating the criteria of choice[11]. To address this deficiency, I use an additional dataset called "Système Priorité Logement" (SyPLo), which details offers extended by the State when acting as the Sponsoring Allocator[12] from 2019 to 2021, including those rejected by applicants. This dataset specifically includes a subset of applicants classified as particularly vulnerable and thus prioritized by the State, forming the primary pool from which the State extends offers.

In order to get information on municipality officials and election results I make use of an open-access dataset on election results provided by the French Ministry of Internal Affairs. This gives me, for each municipality, the political leaning of the mayors elected in 2014 and 2020. It also displays information on election results such as vote counts and participation rates in each municipality.

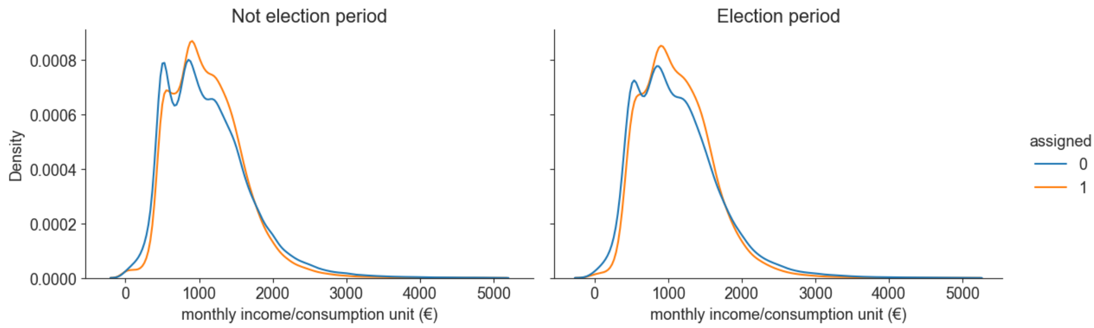
3.3.2 Descriptives

Table 3.1 displays characteristics of social housing applicants and compares them to the French population as a whole. Applicant households tend to be larger, have younger household heads, and are much more likely to be foreigners.
Table 3.2 shows descriptive statistics on the result of French municipal elections of 2014 and 2020. In 2020, candidates from all municipalities of less than 9,000 inhabitants were released from the obligation of declaring their political affiliation, resulting in a large increase in the number of winner with an unstated political color. Political colors of incumbents used in this paper come from the 2014 elections and therefore do not suffer from this lack of information. Miscellaneous candidates are usually affiliated to regionalist parties.

Figure 3.1 shows the distribution of household income among applicants to social housing, separating those that got assigned a unit in my sample from those that did not. The left subgraph represents applicants who left the waitlist outside of an election period, and the right one those who left during an election period. I consider the election period to be the 6 months preceding the 2020 municipal elections. I restrict my sample to start in January 2019 and it stops in December 2021. Comparing the two curves shows little differences between assigned applicants' income in and out of an election period. However, the effect could be heterogeneous across cities, and the targeted applicant population could be different depending on the political color of the incumbent mayor. Clearing out this possibility requires further empirical analysis.

[11] See the estimation of allocation criteria in section 3.4.2
[12] See section 3.2 for a definition of Sponsoring Allocator.

Figure 3.1: Distribution of income/household according to assignment status
By election period



Note: This figure uses all applications starting from 2019.

Table 3.1: Comparison of household characteristics across samples

Variable	Social housing applicants				All tax filers			
	Mean	25%	50%	75%	Mean	25%	50%	75%
household size	2.24	1	2	3	1.83	1	1	2
currently in soc. hous.	0.31				0.16			
age*	45.72	34	43	55	59.14	45	59	73
retired*	0.10				0.34			
unemployed*	0.14				0.11			
French*	0.78				0.85**			
non European*	0.18				0.04**			
European*	0.04				0.04**			
N sne	3,475,221							
N univ	35,307,696							

* Values refers to the household member who is the primary tax filer ("référent fiscal" in French).
** Since tax files do not record an individual's citizenship, this refers to the country where they were born

Table 3.2: Election descriptives

	2014	2020
Number of municipality with a winner of political color:		
right	3968	1215
left	3101	927
misc	1746	425
unstated	484	6696
center	315	456
far right	106	13
far left	12	0
Munic. with a Second Round	1696	1300
Munic. with Only a First Round	8036	8432
Munic. where the incumbent wins		3768
Munic. where the winner has a different political color than incumbent		7733
Total Number of Municipalities	9732	9732

3.4 Empirical strategy

My objective is to understand whether social housing is sometimes allocated following opportunistic motivations by the housing authorities. This is particularly likely to occur around municipal elections, given the allocation is mostly decided at the local level and mayors are involved. In order to understand whether social housing is allocated differently when municipal elections are coming, I use the observed pool of assigned social housing units and characteristics of the applicant receiving the unit from 2019 to 2021, and the occurrence of municipal elections in March 2020. I investigate changes in the average characteristics of assigned applicants as elections approach, using variation in electoral competitiveness across cities. Following Aidt et al. (2011), I posit that incumbent mayors in municipalities with a more competitive electoral environment are more likely to behave opportunistically, and I therefore use electoral competitiveness as a measure of treatment intensity - where an observed assignment is considered treated when it occurred less than 6 months prior to a municipal election.

3.4.1 Event study with varying treatment intensity

I study the effect of proximity to elections on the characteristics of applicants selected to receive a unit. Let's consider Y_{it}, a characteristic on the applicant receiving unit i at time t and in municipality $m(i)$, such as their age, income, or family size. Let me represent the dependence of this characteristic on proximity to elections as follows:

$$Y_{it} = \alpha + \beta_1 D_t + \lambda_{m(i)} + \varepsilon_{it} \tag{3.1}$$

Where D_t is the treatment dummy, equal to one when t is within the 6 months preceding the 2020 municipal elections, $\lambda_{m(i)}$ a municipality fixed-effect and ε_{it} an iid error term.

In this equation, β_1 captures seasonal effects if there are any. In order to be able to separate the effect of treatment from that of time, I assume that incumbents facing stronger competition from rivals are more likely to distort the allocation of social housing so as to gain voters. I therefore represent the effect of treatment as that of electoral competition during treatment. Additionally, I allow for incumbent mayors of different political leaning to respond differently to treatment. I rewrite equation representing the link between applicant characteristic Y_{it} and treatment as follows:

$$Y_{it} = \alpha^E + \beta_1 D_t\, Z_{m(i)} + \beta_2 D_t\, Z_{m(i)}\, X_{m(i)} + \beta_3 X_{m(i)}\, D_t + \theta_1 W_i^U + \theta_2 W_{m(i)}^M\, D_t + \lambda_{m(i)}^E + \delta_t^E + \varepsilon_{it}^E \tag{3.2}$$

Where $Z_{m(i)}$ represents treatment intensity in municipality $m(i)$, $X_{m(i)}$ a vector of dummy variables reflecting the political leaning of the incumbent mayor in municipality $m(i)$, W_i^U a set of unit-specific controls, W_i^M a set of municipality-specific controls, and δ_t a time fixed-effect.

In order to capture the level of electoral competition in a municipality, I use the margin of victory of the 2020-elections winner. Since this is endogenous to the incumbent's behavior regarding social housing allocation, I also use the 2014-election margin of victory as a proxy for 2020 electoral competition in a robustness check. Treatment intensity $Z_{m(i)}$ is the inverse of the margin of victory used, such that a small margin of victory (reflecting high electoral competition) leads to high treatment intensity.

The main identifying assumption here is that when fixing the incumbent's political leaning, municipalities with different electoral competitiveness follow parallel trends when it comes to characteristics of applicants receiving social housing in units of type W_i. This assumption means that for any time period couple (t, s) and any treatment intensity value tuple (z_1, z_2)

$$\begin{aligned} &E[Y_{it}(0) \mid Z_{m(i)} = z_1, X_{m(i)}, W_i] - E[Y_{it}(0) \mid Z_{m(i)} = z_2, X_{m(i)}, W_i] \\ &\qquad = E[Y_{is}(0) \mid Z_{m(i)} = z_1, X_{m(i)}, W_i] - E[Y_{is}(0) \mid Z_{m(i)} = z_2, X_{m(i)}, W_i] \end{aligned}$$

Where $Y_{it}(0)$ is the characteristic of the applicant assigned social housing unit i at time t outside of an election period.

As an example, this assumption is violated if municipalities with more electoral competition are also the ones facing higher inflow of immigrants, leading their social housing applicant pool to evolve differently in time that other municipalities. This is because it would probably lead the characteristics of their assigned applicants to evolve differently over time than that of other municipalities, even absent elections.

Figures C.1 to C.24 in the Online Appendix plot the evolution of the main applicant characteristics of interest aggregated for cities of similar electoral competitiveness and incumbent political leaning before the start of the election period.

In this design, the coefficients identifying an effect of treatment are β_1 an β_2. A large β_1 in magnitude indicates that incumbent mayors have a tendency to allocate social housing differently when elections are coming, and more so when facing fiercer competition from rival candidates. If we take Y_{it} to be a dummy taking the value one when the applicant is of French or EU citizenship

(and therefore able to vote for municipal elections), then a positive β_1 means that on average incumbent mayors tend to increasingly target potential voters in election times and when facing stronger competition. This could be a strategy to gain votes in the upcoming election.
In order to see whether incumbents with certain political leanings are more likely to use social housing allocations for opportunistic purposes, we should look at the magnitudes of β_2 coefficients. Large magnitudes indicate stronger differences in behavior across political families.

3.4.2 Difference-in-differences framework

When social housing developments are created, developers receive subsidies from entities such as the state and the municipality. These entities then become Sponsoring Allocators for a share of the units in the development[13]. This gives them significant influence over which applicants receive these units. Units for which the municipality is the Sponsoring Allocator are more likely to be assigned opportunistically, as mayors have greater control over the allocation process. Unfortunately, there is no reliable national-level information on which entity is the Sponsoring Allocator for each unit. The only exceptions are units where the State is the Sponsoring Allocator, as these are clearly labeled. Given these circumstances, I cannot focus my analysis on units for which the municipality is the Sponsoring Allocator. However, I can identify State-sponsored units, which are less likely to be influenced by the upcoming municipal election.

Using the approach developed in Botbol (2024), I model the probability to receive a social housing unit as a function of applicant characteristics in and out of the pre-election period. I estimate the applicant characteristics associated with higher or lower probability to receive a unit using data on all offers issued for state-sponsored social housing units in the department of Yvelines. The results, shown in Table 3.3, show little difference in the impact of applicant characteristics in and out of election periods for most variables. This suggests that State-sponsored units aren't strongly affected by municipal elections.
I therefore consider those units to be a control group, unaffected by the 2020 municipal elections, and estimate the effect of treatment using a difference-in-differences approach. The dependency of the characteristics of the assigned applicant for unit i at time t over treatment is modeled as

[13]See section 3.2 for more detail.

Table 3.3: Allocation criteria by time period, Yvelines Department

	No Election	Election
Marginal effect job contract = permanent	0.091164	0.102247
Marginal effect job contract = unemployed	0.043471	0.066778
Marginal effect citizenship = Non EU	-0.025357	0.039320
Marginal effect citizenship = EU	-0.025202	0.014809
Marginal effect has rejected	-0.004846	0.019199
Marginal effect single mother	0.009283	0.015048
Elasticity effort rate	-0.006611	-0.006189
Elasticity nb room per hh member	-0.850951	-0.882052
Elasticity waiting time	-0.015702	0.034287

Note: Figures represent marginal effects for discrete and categorical variables, and elasticities for continuous variables.

follows:

$$Y_{it} = \alpha^D + \gamma_1 D_{it} + \gamma_2 D_{it} Z_{m(i)} + \gamma_3 D_{it} X_{m(i)} + \gamma_4 D_{it} Z_{m(i)} X_{m(i)} + \gamma_5 D_{it} X_{m(i)} + \kappa_2 \text{State}_i + \lambda^D_{m(i)} + \delta^D_t + \varepsilon^D_{it} \tag{3.3}$$

Where State_i is a dummy taking value 1 when the unit is state-owned and $D_{it} = 0\,\forall t$ for state-owned units.

As in section 3.4.1, I allow for treatment to vary in intensity as a function of the level of electoral competition in the municipality. However, unlike in section 3.4.1, treatment intensity is no longer required for identification, and I can identify an effect of treatment separately from its relation with electoral competition. The main identifying assumption is that outcomes follow a parallel trend in control and treatment groups, which can be formalized as follows:

$$\begin{aligned} \forall (t,s), \quad & E[Y_{it}(0)\,|\,\text{State} = 0, X_{m(i)}, W_i] - E[Y_{it}(0)\,|\,\text{State} = 1, X_{m(i)}, W_i] \\ & = E[Y_{is}(0)\,|\,\text{State} = 0, X_{m(i)}, W_i] - E[Y_{is}(0)\,|\,\text{State} = 1, X_{m(i)}, W_i] \end{aligned}$$

Figures C.25 to C.36 in the Online Appendix plot the evolution of the main applicant characteristics of interest for state-owned units and others, aggregated for cities of similar incumbent political leaning before the start of the election period.

Coefficient γ_1 represents the average effect of treatment in municipalities with the lowest possible treatment intensity and for the reference value of $X_{m(i)}$. Taking Y_{it} to be a dummy taking the value one when the applicant is of a citizenship that can vote for municipal elections[14] (as in section 3.4.1), a positive γ_1 suggests incumbent mayors tend to favor applicants who can vote in social housing allocations around election times. To understand the effect of treatment on incumbents of other political leanings than the reference category, one should consider $\gamma_1 + \gamma_3$, capturing treatment effect net of electoral competition for municipalities according to their value of $X_{m(i)}$. Coefficient γ_2 and γ_4 provide the additional effect of treatment brought by electoral competition. As a result, $\gamma_1 D_{it} + \gamma_2 D_{it} Z_{m(i)} + \gamma_3 D_{it} X_{m(i)} + \gamma_4 D_{it} Z_{m(i)} X_{m(i)}$ is the additional proportion of assigned applicants with French or EU citizenship brought by incumbents' opportunistic behavior.

3.5 Results

3.5.1 Event study results

Table 3.4 shows the main results. See full table of results in Table C.1 of the online appendix.

I run the regression for 4 different dummy outcomes: whether the applicant household is poor (income per consumption unit below the national median among social housing applicants, 1051€), young (below 30 years old), able to vote for municipal elections (French or EU citizens only), and a family (3 or more household members). Political leanings are categorized as customary by the french Ministry of the Interior.

Results show almost no evidence that the population assigned around election times in municipalities with more electoral competition. Coefficients for β_1 and β_2 are insignificant for all outcomes and political leanings, at the exception of far right mayors who seem to have a tendency to influence social housing assignments towards poorer households than usual whenever facing more electoral competition around elections. However, the coefficient estimates for β_3 attest that a mayor's political leaning seems to be associated with different time trends around elections regarding the income and age of assigned households. Should the trends be the same for all municipalities, this could be interpreted as an effect of proximity to elections on the selection of social housing applicants, given average time trends are captured by month $\times$ year fixed effects δ_t^E.

3.5.2 Difference-in-differences results

Table 3.4 shows the main results. See Table C.2 in the online appendix for full results.

I run the regressions for the same dummy outcomes as described in section 3.5.1. Estimates are very noisy and do not indicate any effect of elections of treatment intensity on the assigned population for social housing.

[14] Citizenship from any EU country grants the ability to vote for French municipal elections.

Table 3.4: Event study estimates

	poor	young	can vote	family
D × Z (β_1)	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Center × D × Z (β_2)		reference category		
X = Far Right × D × Z (β_2)	-0.004*	0.002	0.000	-0.001
	(0.002)	(0.002)	(0.001)	(0.002)
X = Left × D × Z (β_2)	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Misc × D × Z (β_2)	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Right × D × Z (β_2)	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Unstated × D × Z (β_2)		omitted		
X = Center × D (β_3)	0.060	-0.071*	0.016	-0.056
	(0.032)	(0.030)	(0.023)	(0.031)
X = Far Right × D (β_3)	0.095*	-0.100**	0.020	-0.040
	(0.040)	(0.038)	(0.029)	(0.039)
X = Left × D (β_3)	0.068*	-0.080**	0.017	-0.050
	(0.031)	(0.030)	(0.023)	(0.031)
X = Misc × D (β_3)	0.058	-0.063*	0.015	-0.056
	(0.032)	(0.030)	(0.023)	(0.031)
X = Right × D (β_3)	0.059	-0.075*	0.020	-0.052
	(0.031)	(0.030)	(0.023)	(0.031)
X = Unstated × D (β_3)	0.056	-0.068*	0.019	-0.044
	(0.032)	(0.030)	(0.023)	(0.031)
N	525,147	525,147	525,147	525,147
adj. R^2	0.082	0.034	0.088	0.032

* p<0.05, ** p<0.01, *** p<0.001

Table 3.5: DiD estimates

	poor	young	can vote	family
D (γ_1)	-0.008	-0.000	0.001	-0.005
	(0.007)	(0.006)	(0.005)	(0.007)
D × Z (γ_2)	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Center × D (γ_3)	reference category			
X = Far Right × D (γ_3)	0.046	-0.047	-0.008	-0.019
	(0.030)	(0.028)	(0.022)	(0.029)
X = Left × D (γ_3)	0.009	-0.011*	-0.003	0.000
	(0.006)	(0.005)	(0.004)	(0.006)
X = Misc × D (γ_3)	0.008	0.011	-0.001	0.014
	(0.009)	(0.008)	(0.006)	(0.009)
X = Right × D (γ_3)	-0.001	-0.003	-0.003	-0.004
	(0.006)	(0.005)	(0.004)	(0.006)
X = Unstated × D (γ_3)	0.005	-0.002	-0.005	0.002
	(0.008)	(0.008)	(0.006)	(0.008)
X = Center × D × Z (γ_4)	reference category			
X = Far Right × D × Z (γ_4)	-0.002	0.002	0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
X = Left × D × Z (γ_4)	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Misc × D × Z (γ_4)	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Right × D × Z (γ_4)	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Unstated × D × Z (γ_4)	omitted			
N	525,147	525,147	525,147	525,147
adj. R^2	0.082	0.034	0.088	0.032

* p<0.05, ** p<0.01, *** p<0.001

3.6 Conclusion

This study tries to identify evidence of opportunistic manipulation of social housing allocations by mayors. It uses a rich dataset of social housing allocations along with variation coming from both time to municipal elections and levels of competition between local candidates across municipalities. Additionally, the analysis leverages specifics of social housing administration in France to identify a group of social housing units close to unaffected by municipal elections in order to identify the effect of elections using a difference-in-differences framework. Results show little to no evidence of electoral cycles in social housing allocations, although results are little precise. Further research should look for heterogeneity in these results by splitting the sample in more homogeneous municipalities and populations.

Bibliography

Abdulkadiroğlu, A., N. Agarwal, and P. A. Pathak (2017). The welfare effects of coordinated assignment: Evidence from the new york city high school match. *American Economic Review 107*(12), 3635–89.

Abdulkadiroglu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2017). Do parents value school effectiveness? Technical report, National Bureau of Economic Research.

Abdulkadiroğlu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2020). Do parents value school effectiveness? *American Economic Review 110*(5), 1502–39.

Agarwal, N., I. Ashlagi, M. A. Rees, P. Somaini, and D. Waldinger (2021). Equilibrium allocations under alternative waitlist designs: Evidence from deceased donor kidneys. *Econometrica 89*(1), 37–76.

Agarwal, N., C. Hodgson, and P. Somaini (2020). Choices and outcomes in assignment mechanisms: The allocation of deceased donor kidneys. Technical report, National Bureau of Economic Research.

Agarwal, N. and P. Somaini (2018). Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica 86*(2), 391–444.

Aidt, T. S., F. J. Veiga, and L. G. Veiga (2011). Election results and opportunistic policies: A new test of the rational political business cycle model. *Public choice 148*, 21–44.

Alt, J. E. and D. D. Lassen (2006). Transparency, political polarization, and political budget cycles in oecd countries. *American Journal of Political Science 50*(3), 530–550.

Arnosti, N. and P. Shi (2020). Design of lotteries and wait-lists for affordable housing allocation. *Management Science 66*(6), 2291–2307.

Bajari, P., P. Chan, D. Krueger, and D. Miller (2013). A dynamic model of housing demand: Estimation and policy implications. *International Economic Review 54*(2), 409–442.

Balaguer-Coll, M. T., M. I. Brun-Martos, A. Forte, and E. Tortosa-Ausina (2015). Local governments' re-election and its determinants: New evidence based on a bayesian approach. *European Journal of Political Economy 39*, 94–108.

Barahona, N., C. Dobbin, and S. Otero (2023). Affirmative action in centralized college admission systems.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica 84*(3), 893–942.

Besley, T. and S. Coate (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of Public Economics 87*(12), 2611–2637.

Böhlmark, A., H. Holmlund, and M. Lindahl (2016). Parental choice, neighbourhood segregation or cream skimming? an analysis of school segregation after a generalized choice reform. *Journal of Population Economics 29*(4), 1155–1190.

Botbol, L. (2024). Applicant choice in the design of social housing allocations: Evidence from france. *WP*.

Brender, A. and A. Drazen (2005). Political budget cycles in new versus established democracies. *Journal of Monetary Economics 52*(7), 1271–1295. Political economy and macroeconomics.

Cai, X. (2018, July). Determinants Of Affordable Housing Allocation: Common Perspectives From Local Officials. *Regional Science Inquiry 0*(2), 227–237.

Concha-Arriagada, C. (2023). *Should I Stay, or Should I Go? Strategic Responses to Improve College Admission Chances*. WP.

Diamond, R. and T. McQuade (2019). Who wants affordable housing in their backyard? an equilibrium analysis of low-income property development. *Journal of Political Economy 127*(3), 1063–1117.

Drazen, A. and M. Eslava (2010). Electoral manipulation via voter-friendly spending: Theory and evidence. *Journal of Development Economics 92*(1), 39–52.

Drometer, M. and R. Méango (2020). Electoral cycles, partisan effects and us naturalization policies. *Public Choice 183*(1), 43–68.

Ederer, T. (2023). Labor market dynamics and teacher spatial sorting.

Ejdemyr, S., E. Kramon, and A. L. Robinson (2018). Segregation, ethnic favoritism, and the strategic targeting of local public goods. *Comparative Political Studies 51*(9), 1111–1143.

Epple, D. and R. E. Romano (1998). Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, 33–62.

Eriksen, M. D. and S. S. Rosenthal (2010). Crowd out effects of place-based subsidized rental housing: New evidence from the lihtc program. *Journal of Public Economics 94*(11), 953–966.

Escobar, J. F. and L. Huerta (2021). School choice and segregation. Technical report, Tech. rep. mimeo.

Fack, G., J. Grenet, and Y. He (2019). Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review 109*(4), 1486–1529.

Ferdowsian, A., K. H. Lee, and L. Yap (2022). Build to order: Endogenous supply in centralized mechanisms. Technical report, Working Paper.

Gay, C. (2017). A room for one's own? the partisan allocation of affordable housing. *Urban Affairs Review 53*(1), 40–70.

Gazmuri, A. (2018). School segregation in the presence of student sorting and cream-skimming: Evidence from a school voucher reform. *Working Paper*.

Geyer, J. and H. Sieg (2013). Estimating a model of excess demand for public housing. *Quantitative Economics 4*(3), 483–513.

Haeringer, G. and F. Klijn (2009). Constrained school choice. *Journal of Economic theory 144*(5), 1921–1947.

Hanushek, E. A. and L. W ößmann (2006). Does educational tracking affect performance and inequality? differences-in-differences evidence across countries. *The Economic Journal 116*(510), C63–C76.

Hastings, J., T. J. Kane, and D. O. Staiger (2009). Heterogeneous preferences and the efficacy of public school choice. *NBER Working Paper 2145*, 1–46.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Hsieh, C.-T. and M. Urquiola (2006). The effects of generalized school choice on achievement and stratification: Evidence from chile's voucher program. *Journal of public Economics 90*(8-9), 1477–1503.

Idoux, C. (2022). *Integrating New York City schools: The role of admission criteria and family preferences*. MIT Department of Economics.

Joanis, M. (2014). Shared accountability and partial decentralization in local public good provision. *Journal of Development Economics 107*, 28–37.

McEwan, P. J., M. Urquiola, E. Vegas, R. Fernandes, and F. A. Gallego (2008). School choice, stratification, and information on school performance: Lessons from chile [with comments]. *Economia 8*(2), 1–42.

Nelson, K., I. Borg, R. Nieuwenhuis, and S. Alm (2023). The political determinants of housing benefits. *European Sociological Review 39*(1), 104–117.

Nordhaus, W. D. (1975). The political business cycle. *The review of economic studies 42*(2), 169–190.

Robinson-Cortes, A. (2019). Who gets placed where and why? an empirical framework for foster care placement. Technical report, Working Paper.

Rothstein, J. M. (2006). Good principals or good peers? parental valuation of school characteristics, tiebout equilibrium, and the incentive effects of competition among jurisdictions. *American Economic Review 96*(4), 1333–1350.

Schmutz, B. and G. Verdugo (2023). Do elections affect immigration? evidence from french municipalities. *Journal of Public Economics 218*, 104803.

Sieg, H. and C. Yoon (2020). Waiting for affordable housing in new york city. *Quantitative Economics 11*(1), 277–313.

Thakral, N. (2016). The public-housing allocation problem. Technical report, Technical report, Harvard University.

Tincani, M. M., F. Kosse, and E. Miglino (2023). *College Access When Preparedness Matters: New Evidence from Large Advantages in College Admissions.* Centre for Economic Policy Research.

Van Dijk, W. (2019). The socio-economic consequences of housing assistance. *University of Chicago Kenneth C. Griffin Department of Economics job market paper, 0–46 i–xi 36.*

Van Ommeren, J. N. and A. J. Van der Vlist (2016). Households' willingness to pay for public housing. *Journal of Urban Economics 92*, 91–105.

Verdier, V. and C. Reeling (2022). Welfare effects of dynamic matching: An empirical analysis. *The Review of Economic Studies 89*(2), 1008–1037.

Waldinger, D. (2021). Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation. *American Economic Review 111*(8), 2660–2696.

Appendix A

Appendix for Chapter 1

Appendix: Additional Tables

Table A.1: Distribution of monthly tuition for voucher schools, country

	Montly tuition (USD)					
	Mean	P1	P25	P50	P75	P99
2014	14.8	0.0	0.0	2.2	19.1	106.4
2015	15.5	0.0	0.0	1.6	18.6	110.8
2016	13.6	0.0	0.0	0.0	14.7	113.0
2017	13.6	0.0	0.0	0.0	13.9	113.4
2018	12.2	0.0	0.0	0.0	9.5	112.6

Table A.2: DiD on school quality without private schools except voucher who switched, by student SES

	Low SES	Medium SES	High SES
treat	-3.821***	-2.502***	-4.388***
	(0.207)	(0.394)	(0.283)
round 1 only	5.676***	3.006***	6.083***
	(0.200)	(0.377)	(0.285)
both rounds	2.754***	-0.739	2.986***
	(0.393)	(0.536)	(0.515)
round 2 only	0.290	-3.639***	-5.113***
	(0.344)	(0.672)	(0.608)
2015	-0.162		-0.190
	(0.0903)		(0.111)
2016	0.0269		1.189***
	(0.0900)		(0.125)
2017	-0.703***	1.232***	1.992***
	(0.0901)	(0.174)	(0.128)
2018	-0.764***	0.186	0.801***
	(0.0965)	(0.168)	(0.128)
2019	-1.184***	0.129	0.590***
	(0.124)	(0.204)	(0.155)
_cons	-1.430***	2.450***	4.357***
	(0.0630)	(0.127)	(0.0814)
N	406370	96041	228357

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

the label Medium SES (parental income between 40 and 80% of the distribution) appeared in 2016 only

Appendix: Additional Figures

Figure A.1: Distribution of the proportion of low-SES students across courses, replication 2019

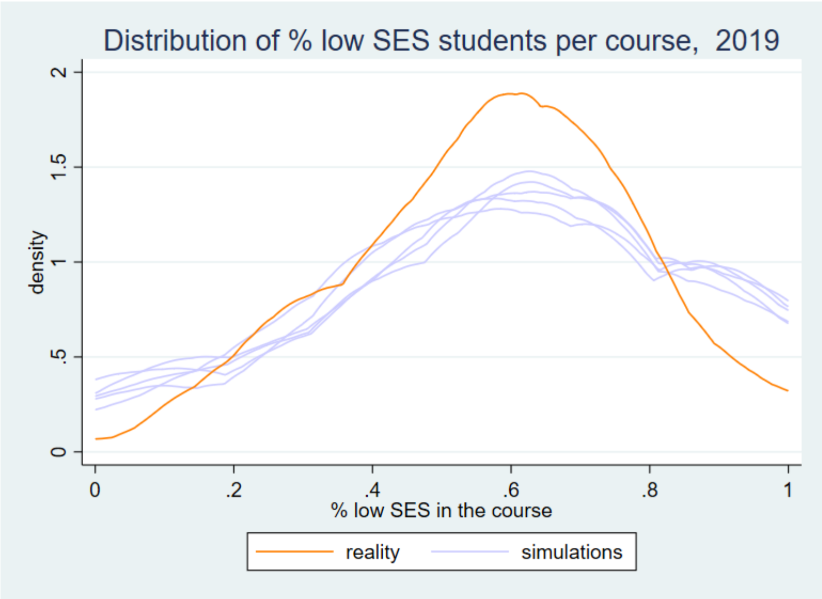


Figure A.2: Students in voucher schools, replication 2019

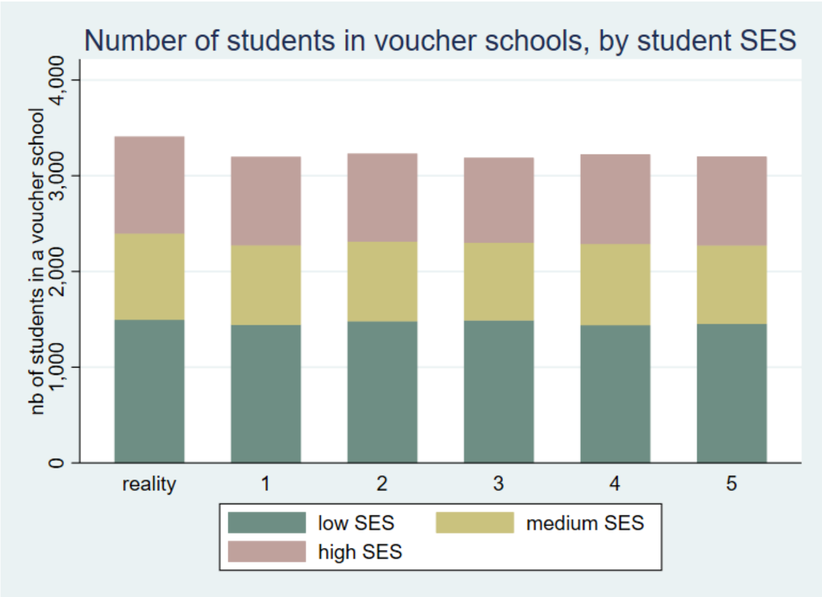


Figure A.3: Unassigned students, replication 2019

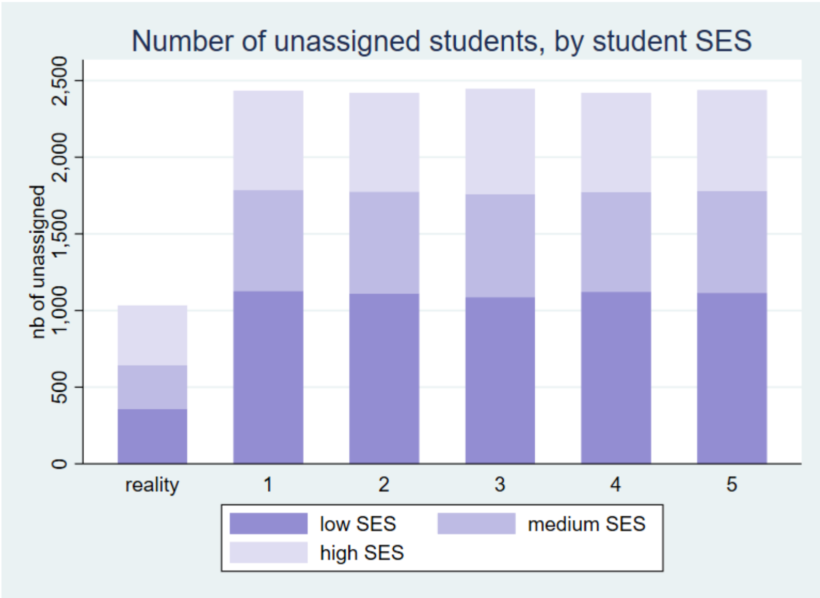
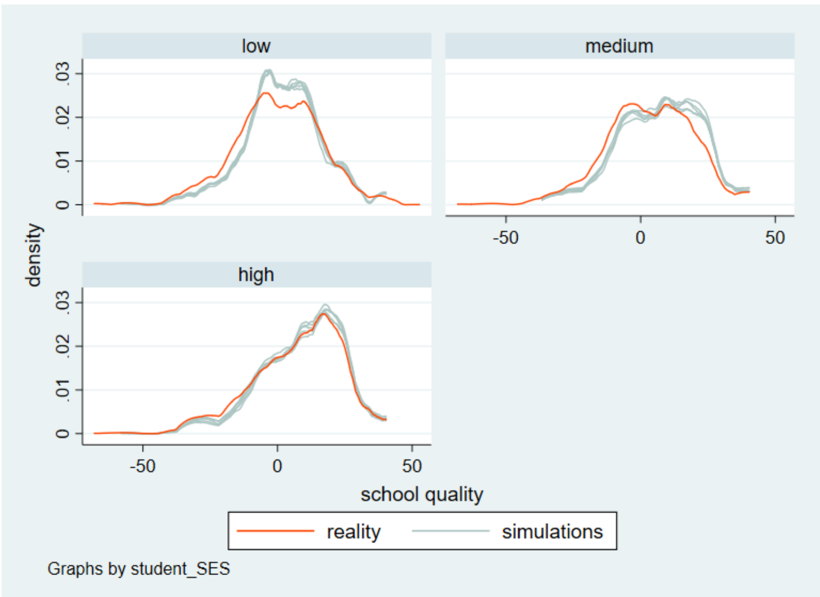


Figure A.4: Quality of the assigned school, replication 2019



Appendix B

Appendix for Chapter 2

B.1 Model and Estimation

B.1.1 Derivation of equation (2.3)

Let's consider a very small unit of time h. Removing subscripts for clarity, we can write the value of $V(t, r = 1)$ as follows:

$$
\begin{aligned}
V(t,1) =& \frac{1}{1+\rho h}\Big[\lambda h \int \pi(t,1) \int \max\{\Gamma(t),0\}\, dG\, dF + \big(1-(\delta+\lambda \int \pi(t,1)\, dF)h\big)V(t+h,1)\Big] \\
\Leftrightarrow& (1+\rho h)V(t,1) - V(t+h,1) = \lambda h \int \pi(t,1) \int \max\{\Gamma(t),0\}\, dG\, dF - (\delta+\lambda \int \pi\, dF)hV(t+h,1) \\
\Leftrightarrow& V(t,1) - V(t+h,1) + \rho h V(t,1) + (\delta + \lambda \int \pi(t,1)\, dF)hV(t+h,1) = \lambda h \int \pi(t,1) \int \max\{\Gamma(t),0\}\, dG\, dF \\
\Leftrightarrow& -\frac{V(t+h,1)-V(t,1)}{h} + \rho V(t,1) + (\delta + \lambda \int \pi(t,1)\, dF)V(t+h,1) = \lambda \int \pi(t,1) \int \max\{\Gamma(t),0\}\, dG\, dF
\end{aligned}
$$

Now if we let h tend to zero we get:

$$
(\rho + \delta + \lambda \int \pi(t,1) dF)V(t,1) = \lambda \int \pi(t,1) \int \max\{\Gamma(t),0\}\, dG\, dF + \dot{V}(t,1)
$$

where $\dot{V}(t,1) \equiv \frac{\partial V(t,r)}{\partial t}$.

B.1.2 Proof of equation (2.5)

I will show that equation (2.4) and (2.5) are equivalent.

Let's rewrite equation (2.5), removing indexation in i, j, m as well as dependence on x_i and w_m for readability:

$$
V(t,0) = \int_t^T e^{-\int_t^s [\rho+\delta(\tau)+\int \pi(\tau,0)\, dF]\, d\tau}\Big(\lambda \int \pi(s,0) \int \max\{V(s,1),\Gamma(s)\} dG\, dF\Big) ds \tag{B.1}
$$

If we define the following objects

$$
\begin{aligned}
a(t) &\equiv \rho + \delta(t) + \int \pi(t,0) \\
H(t) &\equiv \lambda \int \pi(s,0) \int \max\{V(s,1),\Gamma(s)\} dG\, dF
\end{aligned}
$$

we can rewrite:

$$
V(t,0) = \int_t^T e^{-\int_t^s a(\tau)\, d\tau} H(s) ds \tag{B.2}
$$

Let's derive both sides with respect to t:

$$\begin{aligned}\frac{\partial V(t,0)}{\partial t} &= -H(t) + \int_t^T H(s)\frac{\partial}{\partial t}(e^{-\int_t^s a(\tau)\,d\tau})\,ds \\ &= -H(t) + \int_t^T H(s)\,a(t)\,e^{-\int_t^s a(\tau)\,d\tau}\,ds \\ &= -H(t) + a(t)V(t,0)\end{aligned}$$

Which leads us to equation 2.4:

$$a(t)V(t,0) = \frac{\partial V(t,0)}{\partial t} + H(t)$$

B.1.3 Proof of equation (2.8)

Let's rewrite equation (2.8) removing indices and unnecessary dependencies:

$$\tilde{a}(t)V(t,1) = \tilde{H}(t) + \frac{\partial V(t,1)}{\partial t} \tag{B.3}$$

where

$$\begin{aligned}\tilde{a}(t) &\equiv \rho + \delta(t) + \int \pi(t,1)\,dF \\ \tilde{H}(t) &\equiv \lambda \int \pi(s,1)\xi(\Gamma_{ij}(t),0)\,dF\end{aligned}$$

This, together with initial condition $V(T,1) = 0$, forms an ordinary differential equation. The solution of its homogeneous equation $\tilde{a}(t)V(t,1) = \frac{\partial V(t,1)}{\partial t}$ is $V(t,1) = c\exp\int_{t_0}^t \tilde{a}(\tau)\,d\tau$, where c and t_0 are constants. We now replace $V(t,1) = c(t)\exp\int_{t_0}^t \tilde{a}(\tau)\,d\tau$ in equation (B.3) and solve for $c(t)$. We get:

$$\begin{aligned}\tilde{a}(t)c(t)\exp\int_{t_0}^t \tilde{a}(\tau)\,d\tau &= \tilde{H}(t) + \frac{\partial c(t)}{\partial t}\exp\int_{t_0}^t \tilde{a}(\tau)\,d\tau + c(t)\tilde{a}(t)\exp\int_{t_0}^t \tilde{a}(\tau)\,d\tau \\ \Leftrightarrow \tilde{a}(t)c(t) &= \tilde{H}(t)\exp\big(-\int_{t_0}^t \tilde{a}(\tau)\,d\tau\big) + \frac{\partial c(t)}{\partial t} + c(t)\tilde{a}(t) \\ \Leftrightarrow \frac{\partial c(t)}{\partial t} &= -\tilde{H}(t)\exp\big(-\int_{t_0}^t \tilde{a}(\tau)\,d\tau\big) \\ \Leftrightarrow c(t) &= -\int_{t_1}^t \tilde{H}(s)\exp\big(-\int_{t_0}^s \tilde{a}(\tau)\,d\tau\big)\,ds\end{aligned}$$

We report this value of $c(t)$ in $V(t,1) = c(t)\exp\int_{t_0}^t \tilde{a}(\tau)\,d\tau$ and fix t_1 to T using initial condition $V(T,1) = 0$ and get equation (2.8).

B.2 Results

Table B.1: Heterogeneous strategies in city choice

VARIABLES	Model 4 + interactions $V_{im}(t=0, r=0)$
$V_{im}(t=0, r=0)$	15.02***
	(0.410)
$V_{im}(t=0, r=0)$ × applicant income	0.676***
	(0.0908)
$V_{im}(t=0, r=0)$ × applicant age	-0.104***
	(0.00738)
$V_{im}(t=0, r=0)$ × (job = permanent)	-1.450***
	(0.229)
$V_{im}(t=0, r=0)$ × (job = unemployed)	-2.336***
	(0.289)
$V_{im}(t=0, r=0)$ × same city	-4.750***
	(0.160)
$V_{im}(t=0, r=0)$ × distance to current city (log)	-1.117***
	(0.0613)
distance to current city (log)	0.297***
	(0.0204)
distance to current city (log)^2	-0.00154***
	(0.000407)
distance to current city (log) × applicant age	-1.117***
	(0.0613)
distance to current city (log) × (job = permanent)	-0.00603
	(0.0149)
distance to current city (log) × (job = unemployed)	-0.0570***
	(0.0208)
same city	2.608***
	(0.0332)
same city × (job = permanent)	0.148***
	(0.0335)
same city × (job = unemployed)	-0.0104
	(0.0481)
same city × city median income	-0.707***
	(0.00957)
Additional controls	city median income, population, population growth and interactions
Observations	187,373
Number of groups	25,399

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figure B.1: Match value as a function of rent

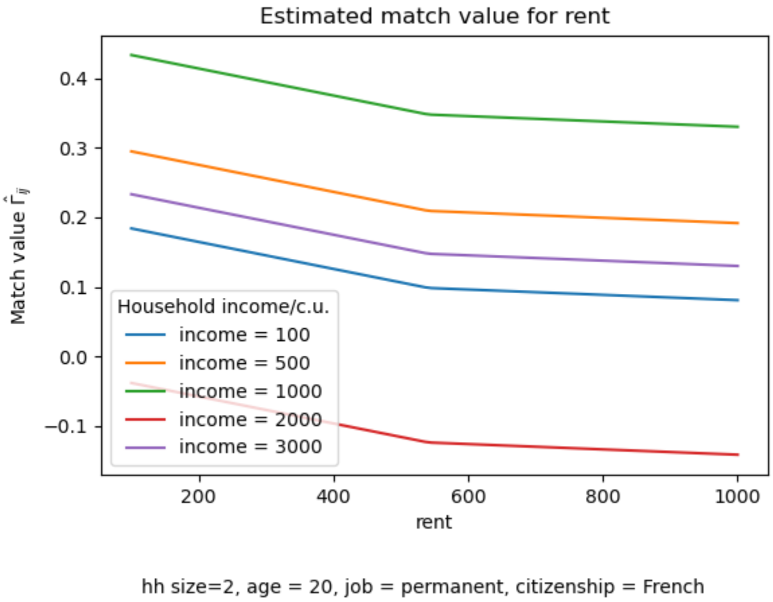


Figure B.2: Match value as a function of neighborhood quality and applicant income

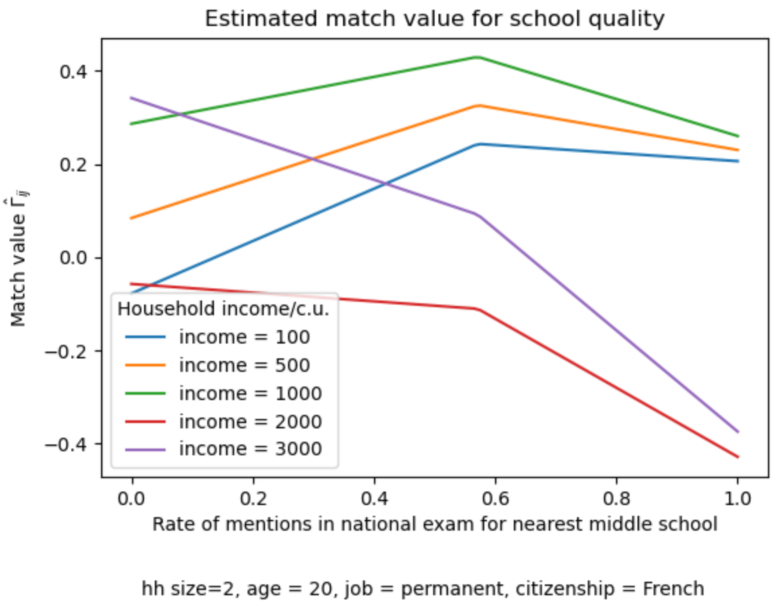


Figure B.3: Match value as a function of the number of rooms

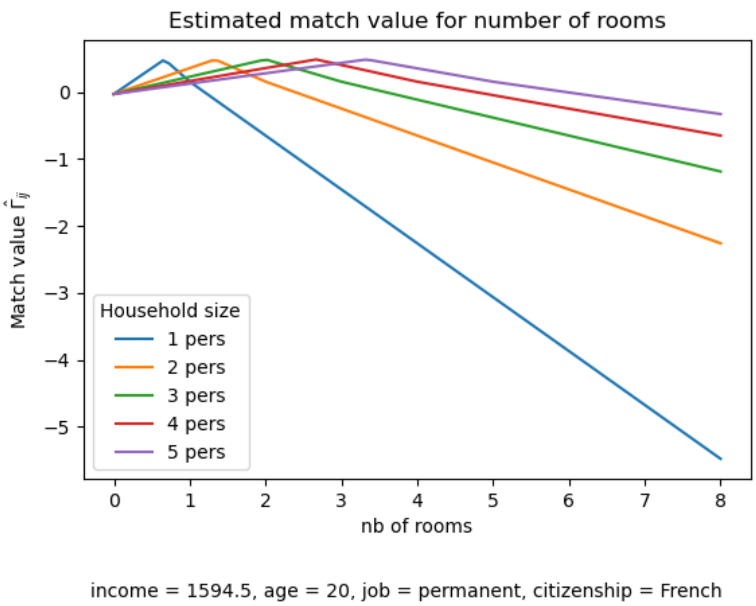
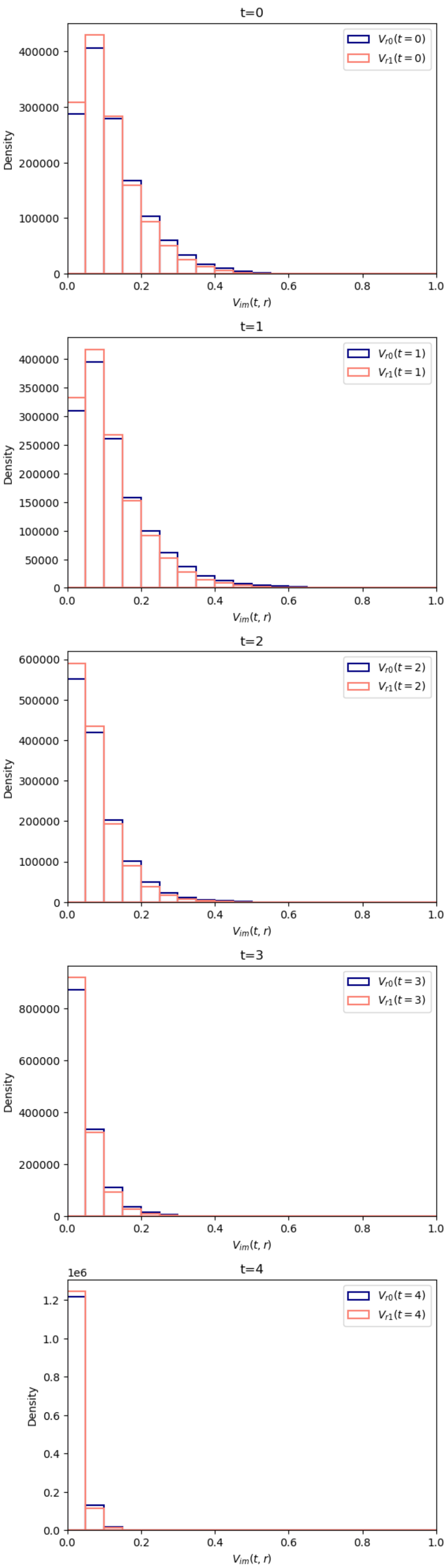


Figure B.4: Value function comparisons across r and t



B.3 Counterfactuals

B.3.1 Counterfactual analysis equilibrium

For simplicity, I assume in counterfactuals that applicants apply to social housing in only one city maximum. Points 3 and 4 of the equilibrium definition are therefore slightly altered. Here is the equilibrium definition that will be used for the counterfactual analysis.

1. Optimality: agents maximize their lifetime utility

2. Consistent beliefs: π is consistent with equilibrium offer probabilities

3. $\tilde{P}_x(m)$ corresponds to the equilibrium rates at which applicants with characteristics x choose city m.

   $$\tilde{P}_x(m) = P[U_{im} > U_{im'}\, \forall m' \in \mathcal{M} \cup 0 | x_i = x]$$

   where U_{im} is the utility of applicant i for applying to social housing in city m, and U_{i0} is the value of the outside option.

4. Steady-state balance condition: let $\mu^{\star}_m(x, t, r)$ be the equilibrium probability distribution of characteristic x, waiting time t and number r of past rejections in the waitlist of city m, and N_m the equilibrium number of applicants who apply to city m. Assuming x is discrete, N_m writes:

   $$N_m = \sum_{x \in \mathcal{X}} h_x \, \tilde{P}_x(m)$$

   where h_x is the rate at which applicants of type x arrive on the market before making city choices.

   The proportion of applicants with characteristics $(x, t = 0, r)$ in m is such that it satisfies the following:

   $$\mu_m(x, t = 0, r = 0) = \frac{\tilde{P}_x(m) h_x}{N_m} \tag{B.4}$$

   $$\& \quad \mu_m(x, t = 0, r = 1) = 0 \tag{B.5}$$

   For each city m, for each discrete characteristics x, $\mu^{\star}_m$ satisfies:

$$\frac{\partial \mu_m(x,t,r=0)}{\partial t} = -\mu_m(x,t,r=0)\kappa_m(x,t,r=0) \tag{B.6}$$

$$\frac{\partial \mu_m(x,t,r=1)}{\partial t} = -\mu_m(x,t,r=1)\kappa_m(x,t,r=1) + \mu_m(x,t,r=0)\tilde{\kappa}_m(x,t,r=0) \tag{B.7}$$

where $\kappa_m(x,t,r)$ is the probability for applicant with attributes (x,t,r) to leave the waitlist of city m, be it for assignment of by voluntary departure. More specifically, it is the probability that they leave the "pool of applicants" with characteristics (x,t,r) of that waitlist - in that sense an applicant with $(x,t,r=0)$ who rejects an offer leaves her pool of the waitlist and joins the pool with characteristics $(x,t,r=1)$. $\tilde{\kappa}_m(x,t,r=0)$ is the probability for an applicant to receive an offer and reject it in city m, therefore joining the pool $(x,t,r=1)$.

The departure probability $\kappa_m(x,t,r)$ is the probability to either receive an exogenous shock and leave the waitlist, or receive an offer. Applicants with one rejection ($r=1$) indeed leave the waitlist whenever they receive an offer: if they accept it they get assigned, if they reject it they depart from the waitlist. Applicants with no rejections ($r=0$) also leave the pool of applicants with attributes $(x,t,r=0)$ from the waitlist whenever they receive an offer: either they accept and get assigned or refuse and thereafter belong to the pool $(x,t,r=1)$. Therefore, $\kappa_m(x,t,r)$ writes as follows:

$$\kappa_m(x,t,r) = \delta_m(x,t) + \lambda_m \int \pi(t,r;z,x,w_m)\,dF_m(z) \tag{B.8}$$

The probability $\tilde{\kappa}_m(x,t,r=0)$ writes:

$$\tilde{\kappa}_m(x,t) = \lambda_m \int \pi(t,r=0;z,x,w_m) \int P\big[\Gamma_{ij}(t) + \varepsilon < V_{i,m}(t,r=1)\big]\,dG(\varepsilon)\,dF_m(z) \tag{B.9}$$

B.3.2 Algorithm details

Start from a discrete time grid $t = t_0, ..., t_l, ...T$, a set of cities $\mathcal{M}$, a vector of arbitrary initial beliefs $\{\pi^0\}_{\forall(x,t_l,r,z,w)}$, a sample of N applicants taken randomly from the population, a sample of units taken randomly in each city $\{\mathcal{J}_m\}_{\forall m}$, parameter estimates $\hat{\Gamma}, \hat{\delta}, \hat{\lambda}$, and a discount factor ρ.

Value function for each city $\{V\}_{\forall x,t_l,r,x,m}$

In order to compute $\{V^{r=1}\}_{\forall x,t_l,m}$ I use the following equation:

$$(\rho+\delta_i(t))V_{i,m}(t,1) = \lambda_m \int \pi_{iz}(t,1) \int \max\{0,\Gamma_{ij}(t)\}dGdF_m + \frac{\partial V_{i,m}(t,1)}{\partial t} \tag{B.10}$$

Note: $\Gamma_{ij}(t) = \Gamma_{xzw}(t) + \varepsilon$.

The differential equation form has the advantage of allowing for a computation of value functions by backward induction, using the fact that $V(T,r) = 0$. Practically, for iteration k of the algorithm, I compute V as follows, for t_l starting from T to t_0:

$$V^{(k)}_{x,m,r=1}(t_l) = \frac{\hat{\lambda}_m}{(\rho+\hat{\delta}_x(t))} \int \pi^{(k-1)}_{xzw_m,r=1}(t_l)\,\xi(\hat{\Gamma}_{xzw_m}(t_l))\,dF_m + \frac{V^{(k)}_{x,m,r=1}(t_{l+1}) - V^{(k)}_{x,m,r=1}(t_l)}{t_{l+1}-t_l} \tag{B.11}$$

$$\Leftrightarrow \Big(1+\frac{1}{t_{l+1}-t_l}\Big)V^{(k)}_{x,m,r=1}(t_l) = \frac{\hat{\lambda}_m}{(\rho+\hat{\delta}_x(t))} \int \pi^{(k-1)}_{xzw_m,r=1}(t_l)\,\xi(\hat{\Gamma}_{xzw_m}(t_l))\,dF_m + \frac{V^{(k)}_{x,m,r=1}(t_{l+1})}{t_{l+1}-t_l} \tag{B.12}$$

$$\Leftrightarrow V^{(k)}_{x,m,r=1}(t_l) = \frac{t_{l+1}-t_l}{t_{l+1}-t_l+1}\left(\frac{\hat{\lambda}_m}{(\rho+\hat{\delta}_x(t))} \int \pi^{(k-1)}_{xzw_m,r=1}(t_l)\,\xi(\hat{\Gamma}_{xzw_m}(t_l))\,dF_m + \frac{V^{(k)}_{x,m,r=1}(t_{l+1})}{t_{l+1}-t_l}\right) \tag{B.13}$$

where $\xi(\hat{\Gamma}_{xzw_m}(t_l)) = \int \max\{0,\Gamma_{ij}(t)\}\,dG$ since G is the pdf of a standard Normal.
I replace the integration over z by a sum over all units in $\mathcal{J}_m$.

$$V^{(k)}_{x,m,r=1}(t_l) \approx \frac{t_{l+1}-t_l}{t_{l+1}-t_l+1}\left(\frac{\hat{\lambda}_m}{(\rho+\hat{\delta}_x(t))} \sum_{j\in\mathcal{J}_m} \pi^{(k-1)}_{xz_jw_m,r=1}(t_l)\,\xi(\hat{\Gamma}_{xz_jw_m}(t_l)) + \frac{V^{(k)}_{x,m,r=1}(t_{l+1})}{t_{l+1}-t_l}\right) \tag{B.14}$$

I can then recover $V^{(k)}_{x,m,r=0}(t_l)$ by backward induction by using $V^{(k)}_{x,m,r=0}(T) = 0$ and the following:

$$\begin{aligned}(\rho+\hat{\delta}_x(t_l))V^{(k)}_{x,m,r=0}(t_l) = \hat{\lambda}_m \int \pi^{(k-1)}_{xzw_m,r=0}(t_l) \int &\max\{0,\hat{\Gamma}_{xzw_m}(t_l)+\varepsilon \\ &-V^{(k)}_{x,m,r=1}(t_l)\}dGdF_m + \frac{V^{(k)}_{x,m,r=0}(t_{l+1}) - V^{(k)}_{x,m,r=0}(t_l)}{t_{l+1}-t_l}\end{aligned} \tag{B.15}$$

$$\Leftrightarrow (\rho + \hat{\delta}_x(t_l) + \frac{1}{t_{l+1} - t_l})V^{(k)}_{x,m,r=0}(t_l) = \hat{\lambda}_m \int \pi^{(k-1)}_{xzw_m,r=0}(t_l)\, \xi\big(\hat{\Gamma}_{xzw}(t_l) - V^{(k)}_{x,m,r=1}(t_l)\big)\, dF_m + \frac{V^{(k)}_{x,m,r=0}(t_{l+1})}{t_{l+1} - t_l} \tag{B.16}$$

$$\Leftrightarrow V^{(k)}_{x,m,r=0}(t_l) = \frac{t_{l+1} - t_l}{(t_{l+1} - t_l)(\rho + \hat{\delta}_x(t_l)) + 1}\Bigg(\hat{\lambda}_m \int \pi^{(k-1)}_{xzw_m,r=0}(t_l)\, \xi\big(\hat{\Gamma}_{xzw_m}(t_l) - V^{(k)}_{x,m,r=1}(t_l)\big)\, dF_m + \frac{V^{(k)}_{x,m,r=0}(t_{l+1})}{t_{l+1} - t_l}\Bigg) \tag{B.17}$$

$$\Leftrightarrow V^{(k)}_{x,m,r=0}(t_l) \approx \frac{t_{l+1} - t_l}{(t_{l+1} - t_l)(\rho + \hat{\delta}_x(t_l)) + 1}\Bigg(\hat{\lambda}_m \sum_{j \in \mathcal{J}_m} \pi^{(k-1)}_{xz_jw_m,r=0}(t_l)\, \xi\big(\hat{\Gamma}_{xz_jw_m}(t_l) - V^{(k)}_{x,m,r=1}(t_l)\big) + \frac{V^{(k)}_{x,m,r=0}(t_{l+1})}{t_{l+1} - t_l}\Bigg) \tag{B.18}$$

Departure probability of the waitlist $\{\kappa\}_{\forall x, t_l, m, r}$

For someone with no past rejections, the probability to exit the waitlist is the sum of the probability to exit voluntarily and the probability to receive and accept an offer. However, we are considering here the departure of the pool of applicants with characteristics (x, t_l, m, r). When they receive and reject an offer, applicants with characteristics $(x, t_l, m, r = 0)$ leave the pool of applicants with such characteristics to join that with $(x, t_l, m, r = 1)$. Therefore, whenever they receive an offer, applicants with $r = 0$ leave their pool of the waitlist. For someone with $r = 1$, applicants also leave the waitlist whenever they receive an offer, since they are forced to leave if they reject.

$$\kappa_m(x, t, r) = \delta_m(x, t) + \lambda_m \int \pi(t, r; z, x, w_m)\, dF_m(z) \tag{B.19}$$

Therefore, for iteration k, $\{\kappa\}_{\forall x, t_l, m, r}$ can be computed as follows:

$$\kappa^{(k)}_{xw,r}(t_l) = \hat{\delta}_x(t_l) + \hat{\lambda}_m \frac{1}{|\mathcal{J}_m|} \sum_{j \in \mathcal{J}_m} \pi^{(k-1)}_{xzw,r}(t_l) \tag{B.20}$$

Now let me write down the rate $\tilde{\kappa}_m(x, t)$ at which applicants with type

$(x, t, r = 0)$ enters the pool of applicants with type $(x, t, r = 1)$:

$$\tilde{\kappa}_m(x,t) = \lambda_m \int \pi(t, r = 0; z, x, w_m) \int P\big[\Gamma_{ij}(t) + \varepsilon < V_{i,m}(t, r = 1)\big]\, dG(\varepsilon)\, dF_m(z) \qquad \text{(B.21)}$$

$$\Rightarrow \tilde{\kappa}^{(k)}_{xw}(t_l) \approx \hat{\lambda}_m \frac{1}{|\mathcal{J}_m|} \sum_{j \in \mathcal{J}_m} \pi^{(k-1)}_{xz_j w, r=0}(t_l)\, \Phi\big(V^{(k)}_{x,m,r=1}(t_l) - \hat{\Gamma}_{xz_j w_m}(t_l)\big) \qquad \text{(B.22)}$$

Utility for each city $\{U\}_{\forall m,x}$ and probability to choose each city $\{\tilde{P}\}_{\forall m,x}$

Let me write the utility for city m. I simplified notations by dropping the bar on U_{im} and replacing applicant i by its type x :

$$U^{(k)}_{xm} = \hat{\gamma}_1 w_m + \hat{\gamma}_2 \text{Dist}_{xm} + \hat{\gamma}_3 V^{(k)}_{x,m,r=0}(0)$$

Probability to rank city m for applicant with characteristics x:

$$\tilde{P}^{(k)}_{xm} = P[U^{(k)}_{xm} + \eta > U^{(k)}_{xm'} + \eta',\ \forall m' \in \mathcal{M} \cup 0] \qquad \text{(B.23)}$$

$$= \frac{\exp(U^{(k)}_{xm})}{1 + \sum_{m' \in \mathcal{M}} \exp(U^{(k)}_{im'})} \qquad \text{(B.24)}$$

Probability to choose the outside option:

$$\tilde{P}^{(k)}_{x0} = \frac{1}{1 + \sum_{m' \in \mathcal{M}} \exp(U^{(k)}_{im'})} \qquad \text{(B.25)}$$

Waitlist size $\{p\}_{\forall m}$ and composition $\{\mu\}_{\forall x, t_l, r, m}$

As specified in the equilibrium definition, the queue composition in each city m is defined by a pdf μ_m, where $\mu_m(x, t, r)$ is the proportion of applicants in the queue that has characteristics (x, t, r).

$$\frac{\partial \mu_m(x,t,r=0)}{\partial t} = -\mu_m(x,t,r=0)\kappa_m(x,t,r=0) \qquad \text{(B.26)}$$

$$\frac{\partial \mu_m(x,t,r=1)}{\partial t} = -\mu_m(x,t,r=1)\kappa_m(x,t,r=1) + \mu_m(x,t,r=0)\tilde{\kappa}_m(x,t,r=0) \qquad \text{(B.27)}$$

with

$$\mu_m(x, t = 0, r = 0) = \frac{\tilde{P}_x(m)\hat{h}_x}{N^{(k-1)}_m} \qquad \text{(B.28)}$$

$$\& \quad \mu_m(x, t = 0, r = 1) = 0 \qquad \text{(B.29)}$$

where $\hat{h}_x$ is the proportion of applicants of type x in the sample and $N_m^{(k-1)}$ is the size of the waitlist in iteration $k-1$.

This system of differential equations is solved numerically using an explicit Runge-Kutta method, as is standard in the literature (Boyce & DiPrima, "*Elementary differential equations and boundary value problems*" Wiley 2020).

The size of the waitlist writes:

$$N_m^{(k)} = \sum_{x \in \mathcal{X}} \sum_{t=0}^{T} \big(\mu_m(x,t,r=1) + \mu_m(x,t,r=0)\big)$$

Since μ is a probability measure, $\sum_{x\in\mathcal{X}} \int_0^T \big(\mu_m(x,\tau,r=1) + \mu_m(x,\tau,r=0)\big)\, d\tau = 1$. I therefore need to scale the output of $\{\mu_{xrm}^{(k)}(t_l)\}_{\forall x,r,t_l}$ after the fact so that it sums to one. Concretely, I divide $\{\mu_{xrm}^{(k)}(t_l)\}_{\forall x,r,t_l}$ by $N_m^{(k)}$.

Offer probabilities $\{\pi\}_{\forall x,t_l,r,z,w}$

For each city m, for each unit j in m, and for each set (x,t,r) such that $\mu_{xmr}^{(k)}(t_l) \neq 0$, compute $\bar{s}_{xtrzm}$ for each (x,t,r). Break ties randomly.

$$\pi_{xzwr}^{(k)}(t_l) = P[\bar{s}_{xt_lrzm} \geq \bar{s}_{x't_l'r'zm}\, \forall (x',t_l',r') \text{ s.t. } \mu_{x'mr'}^{(k)}(t_l') \neq 0] \times \frac{1}{N_m^{(k)} \mu_{xmr}^{(k)}(t_l)} \tag{B.30}$$

$$= \frac{\exp(\bar{s}_{xt_lrzm})}{\exp(\bar{s}_{xt_lrzm}) + \sum_{(x',t_l',r') \in \mathcal{I}_m} N_m^{(k)} \mu_{xmr}^{(k)}(t_l) \exp(\bar{s}_{x't_l'r'zm})} \tag{B.31}$$

where $\mathcal{I}_m$ is the set of (x,t_l,r) such that $\mu_{xmr}^{(k)}(t_l) \neq 0$.

B.3.3 Assumption that applicants apply to one city maximum

Assuming that applicants apply to one city greatly simplifies computations of the algorithm. The reason is that defining κ when applicants can be in multiply cities' waitlists at a time is difficult, since the probability for an applicant to depart from pool (x,t,r) of the waitlist includes the probability that she is assigned in another city that it has ranked. The problem is that the probability to be assigned a unit in another city m' is also correlated with that of being assigned in another city m''. The probability to leave the waitlist of m because I'm assigned in another city writes as follows:

$$\cup_{R\in\mathcal{R}} P\,[\text{assigned (and accepts) in } R \setminus m\,|R_i = R] \times P[R_i = R\,|\,m \in R_i]$$

where R_i is the list of cities ranked by i and $\mathcal{R}$ the set of all possible ROL of cities. The probability to be assigned in R is a combination between the probabilities of being assigned in each city in R, but those probabilities are not independent from one another.

The computatio of $P[R_i = R \,|\, m \in R_i]$ is also complex, since $P[U_{im} > U_{i0}]$ and $P[U_{im'} > U_{i0}]$ are correlated through $U_{i0} = \eta_{i0}$.

B.3.4 Counterfactual results

Figure B.5: Probability of receiving a given offer across applicant groups

Current scoring rule

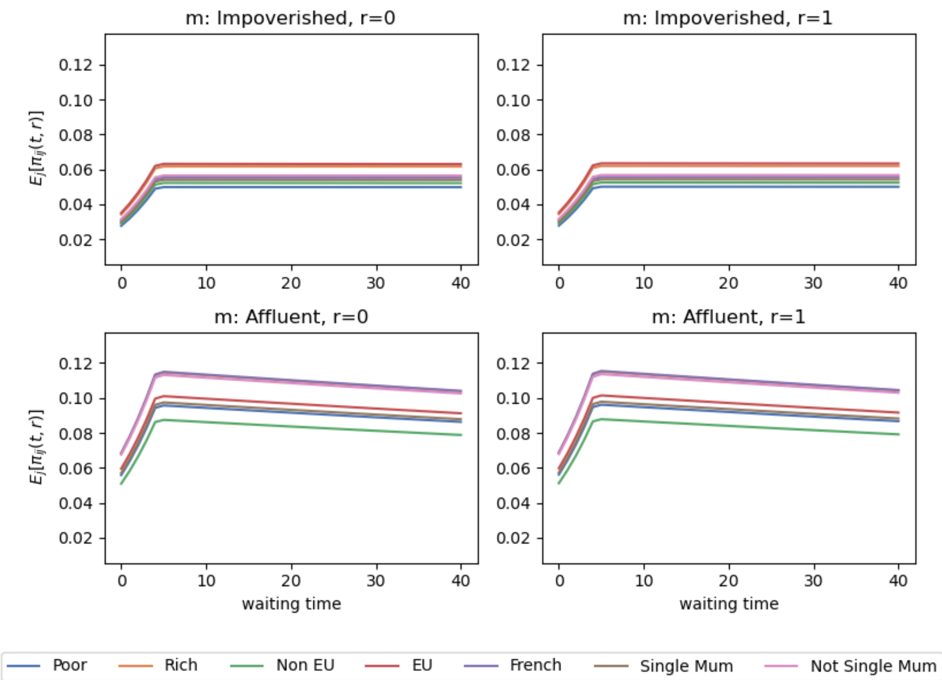


Figure B.6: Probability of receiving a given offer across applicant groups

FCFS

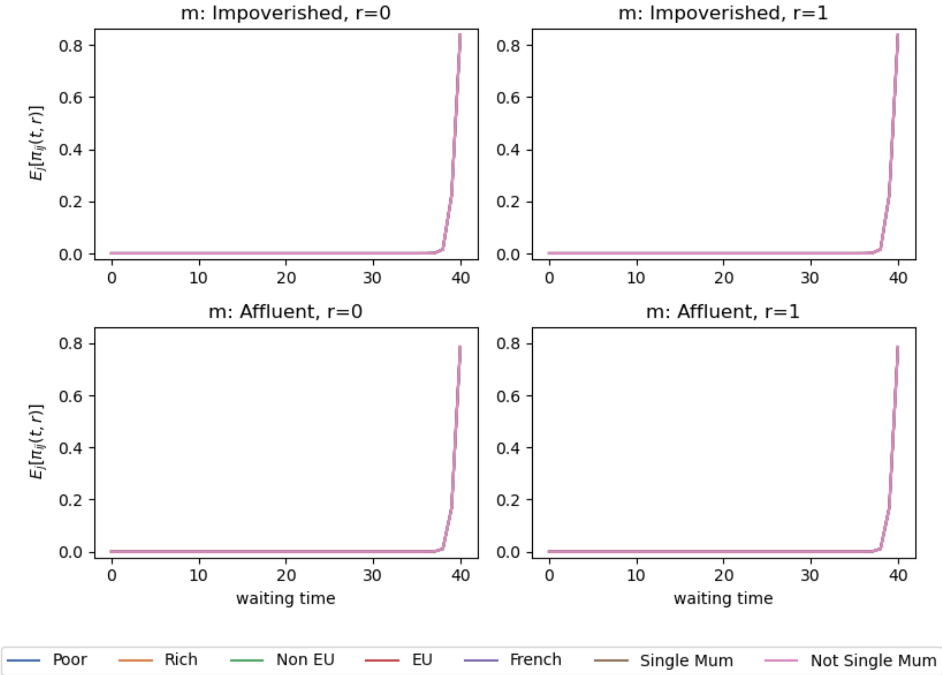


Figure B.7: Probability of receiving a given offer across applicant groups

No weight on waiting time

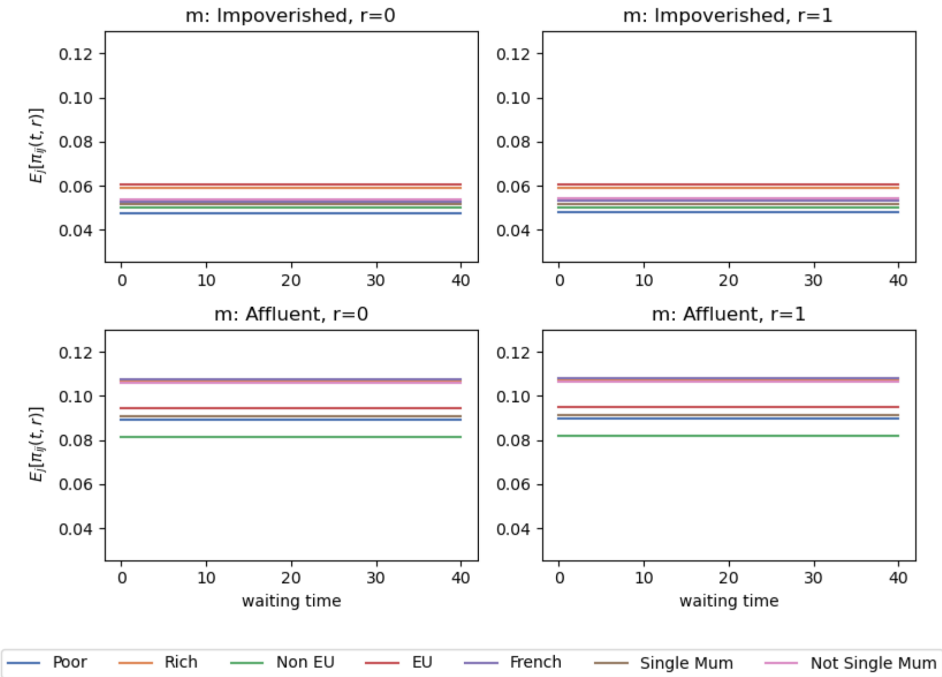


Figure B.8: Probability of receiving a given offer across applicant groups

No weight on past rejections

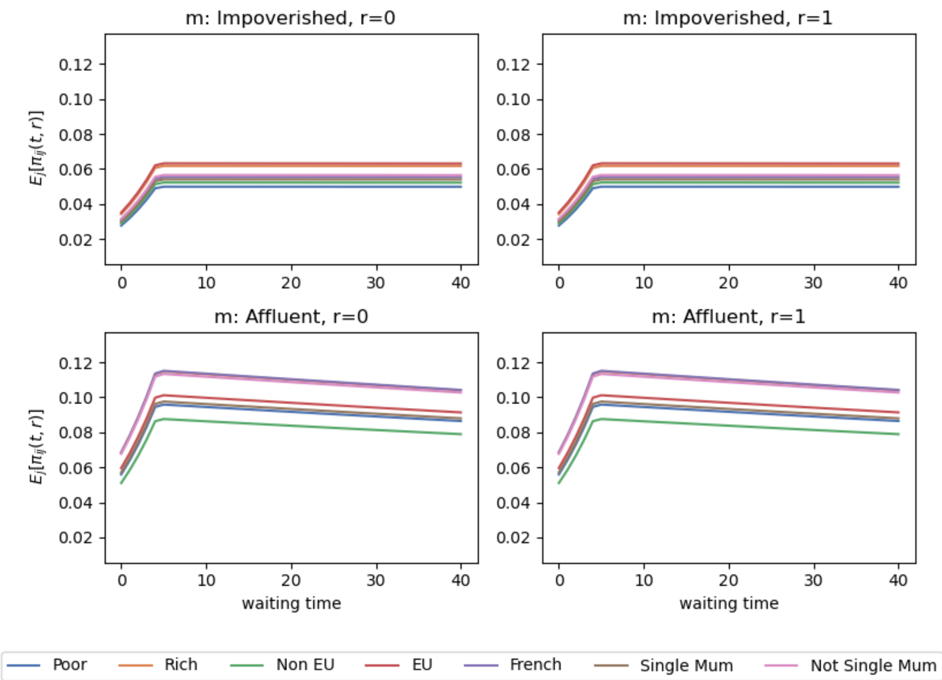


Figure B.9: Probability of receiving a given offer across applicant groups

LCFS

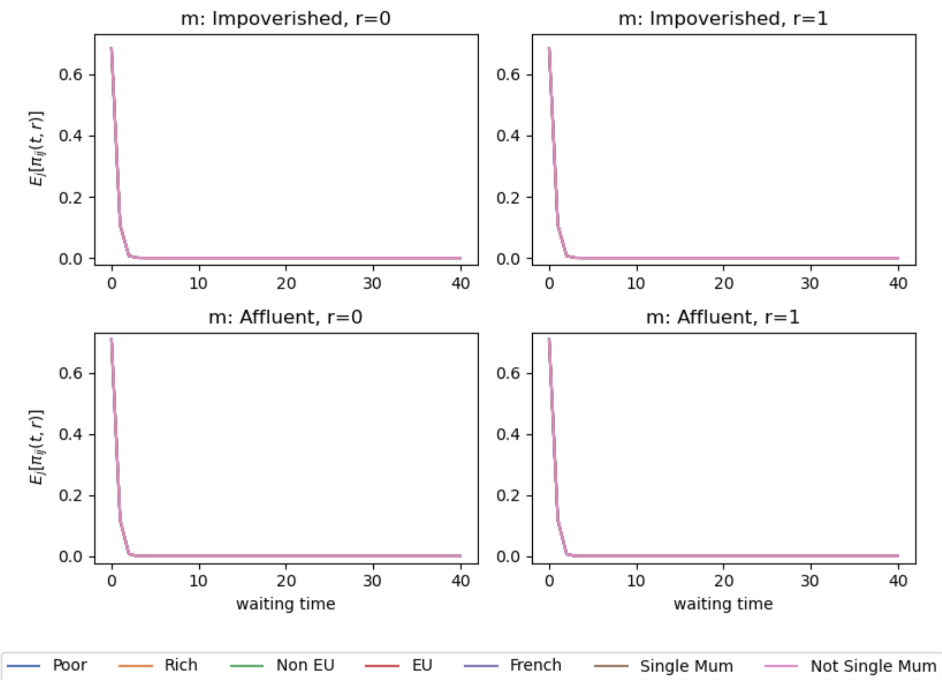


Figure B.10: Probability of receiving a given offer across applicant groups

Linear waiting time

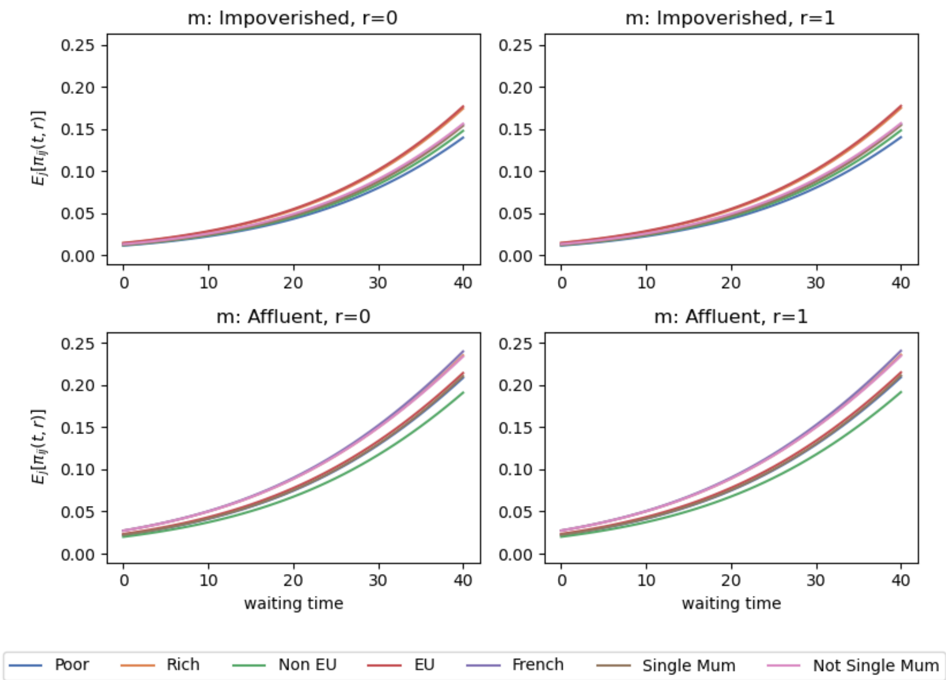


Figure B.11: Probability of receiving a given offer across applicant groups

Favor the Poor

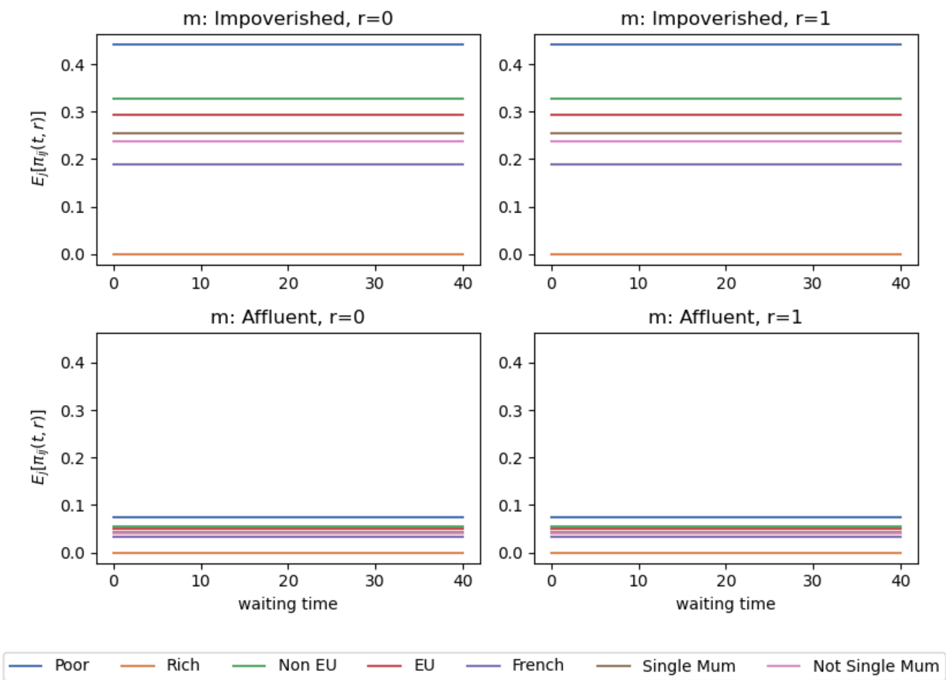


Figure B.12: Waitlist composition

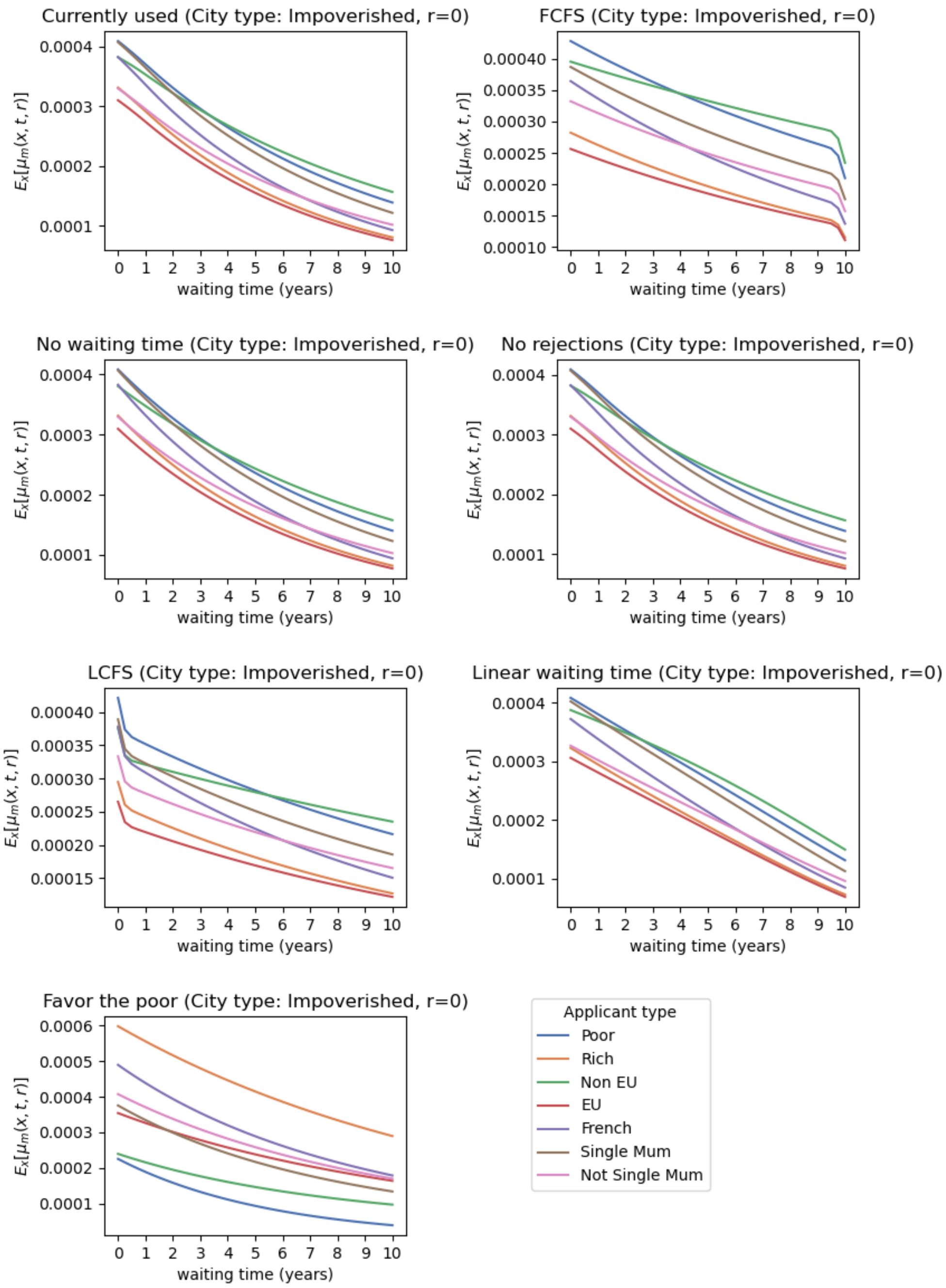


Figure B.13: Waitlist composition

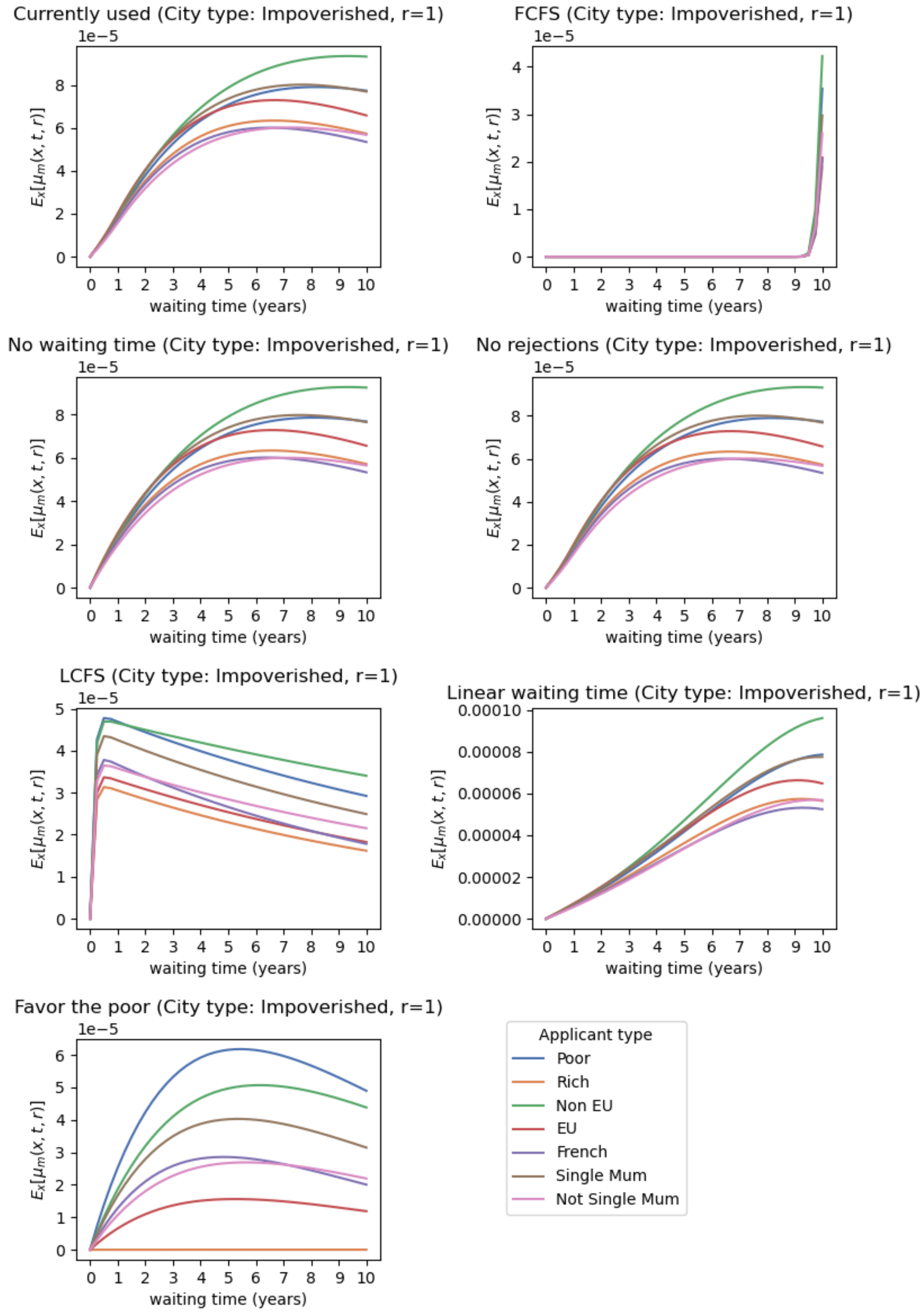


Figure B.14: Waitlist composition

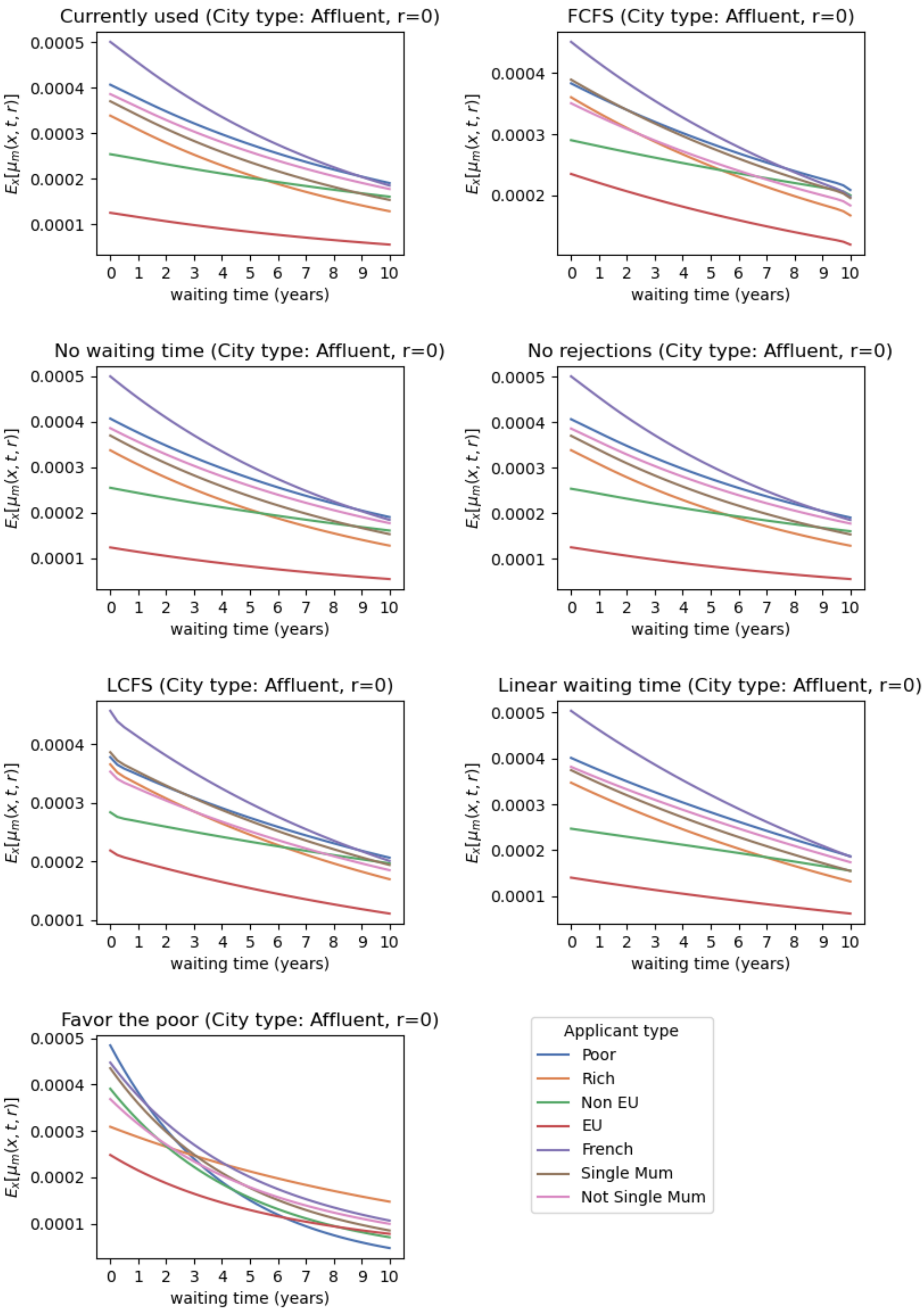


Figure B.15: Waitlist composition

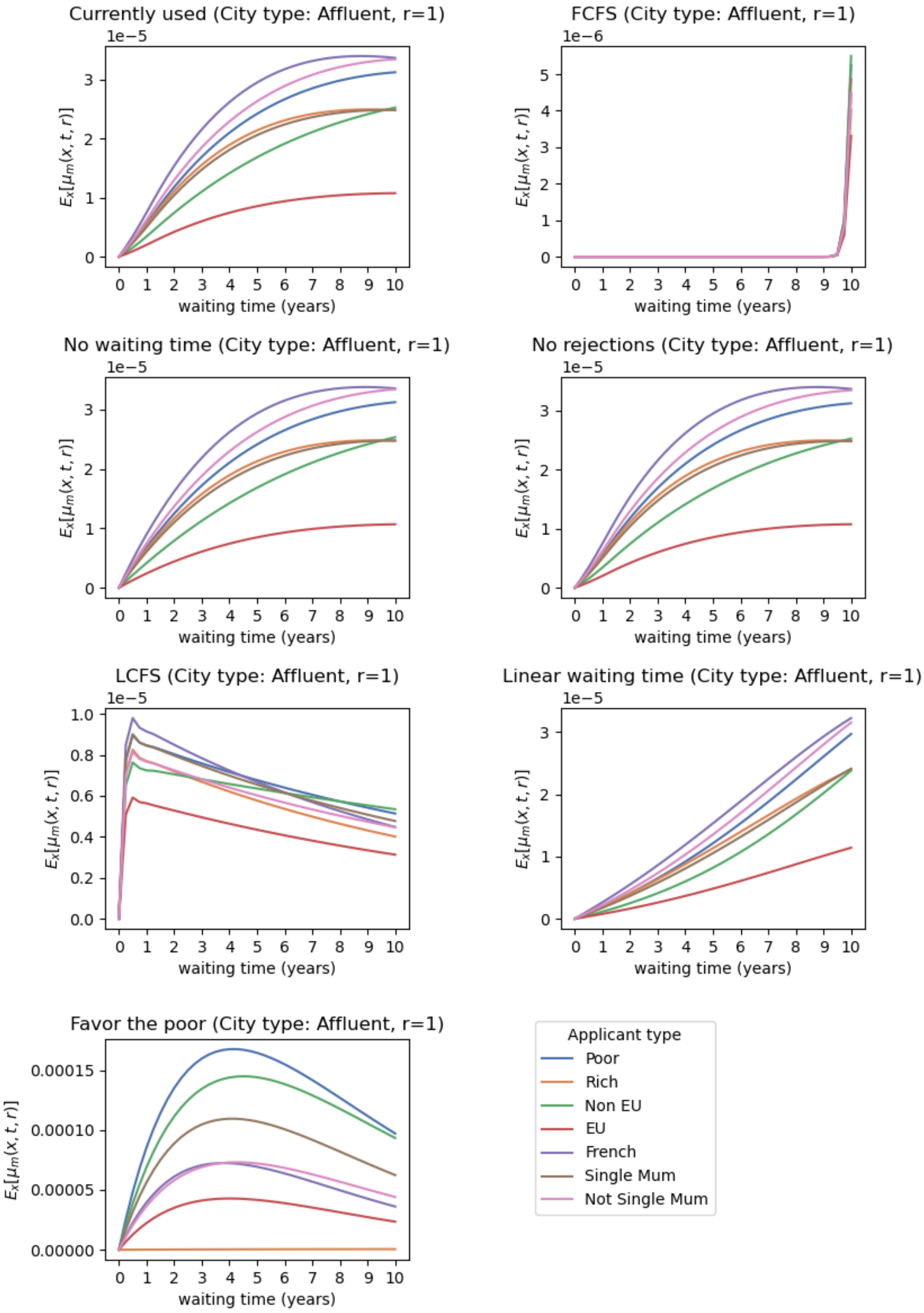


Figure B.16: Hazard function

Impoverished city

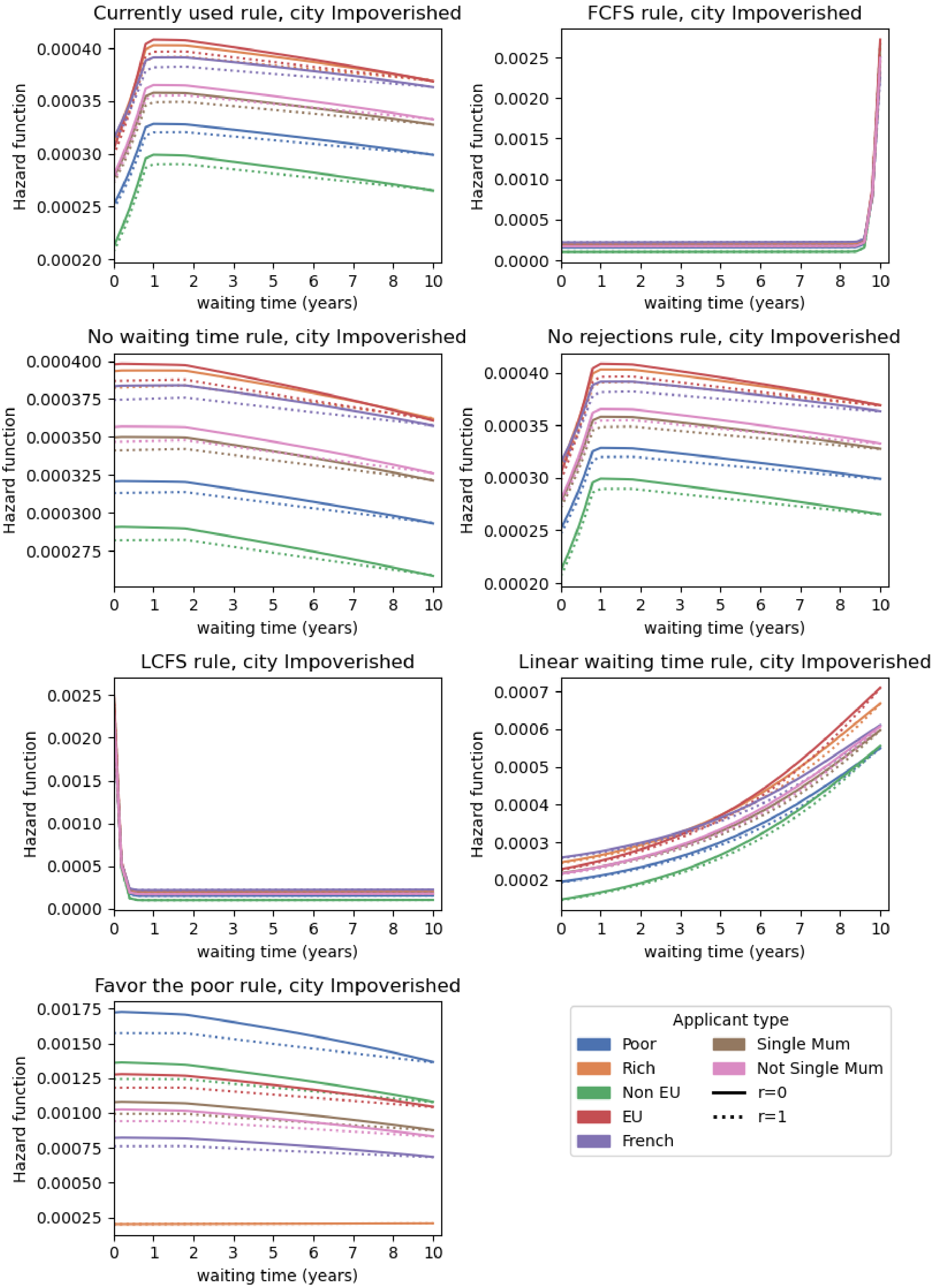


Figure B.17: Hazard function

Affluent city

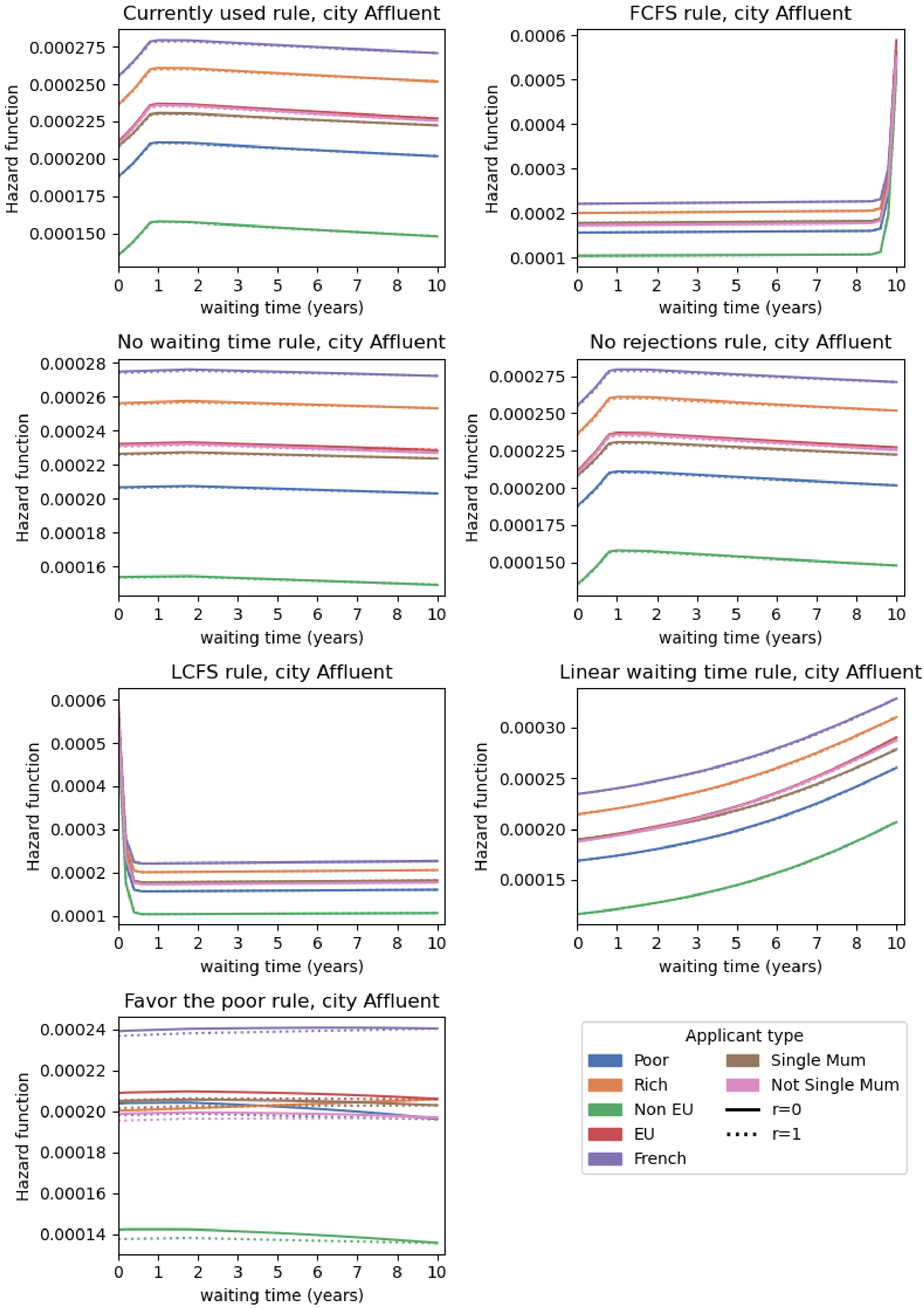


Figure B.18: Match value

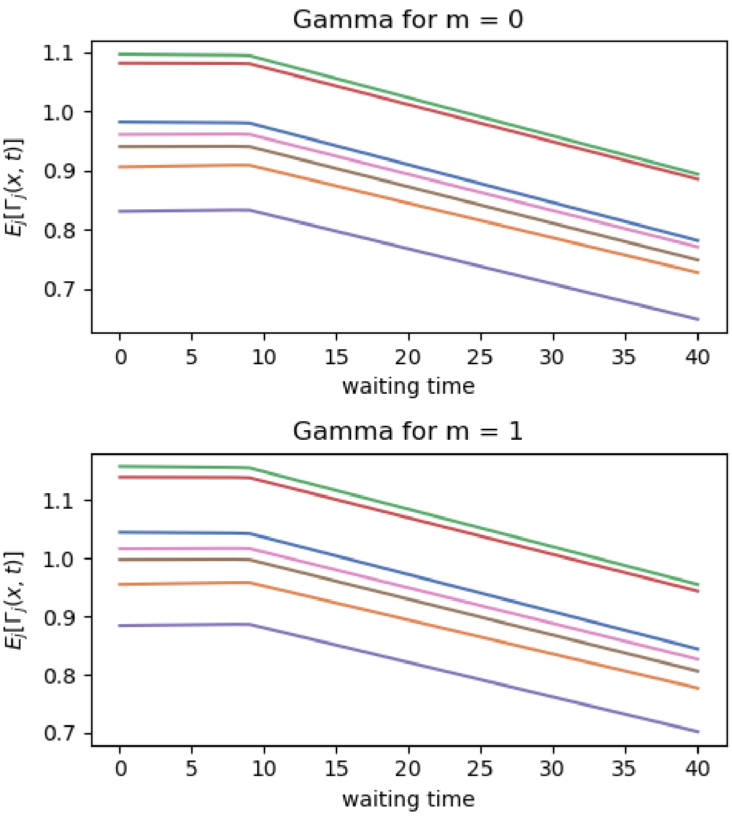


Figure B.19: Probability to leave the waitlist voluntarily

By applicant group

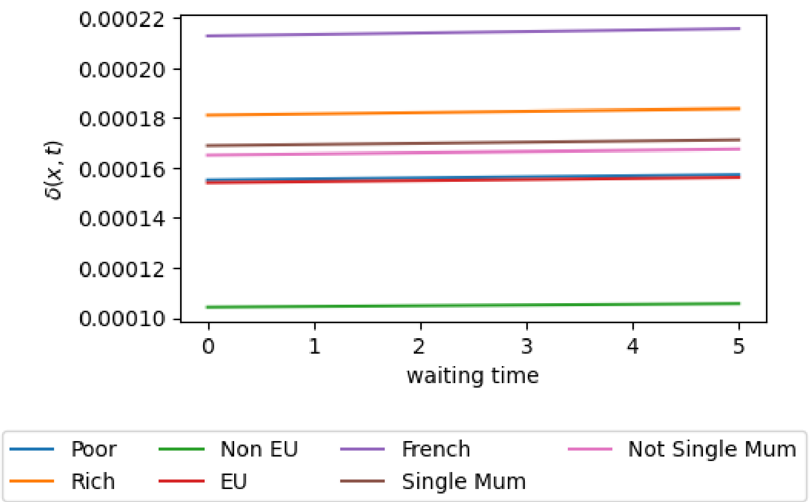


Figure B.20: Average value of a match weighted by probability to receive the offer

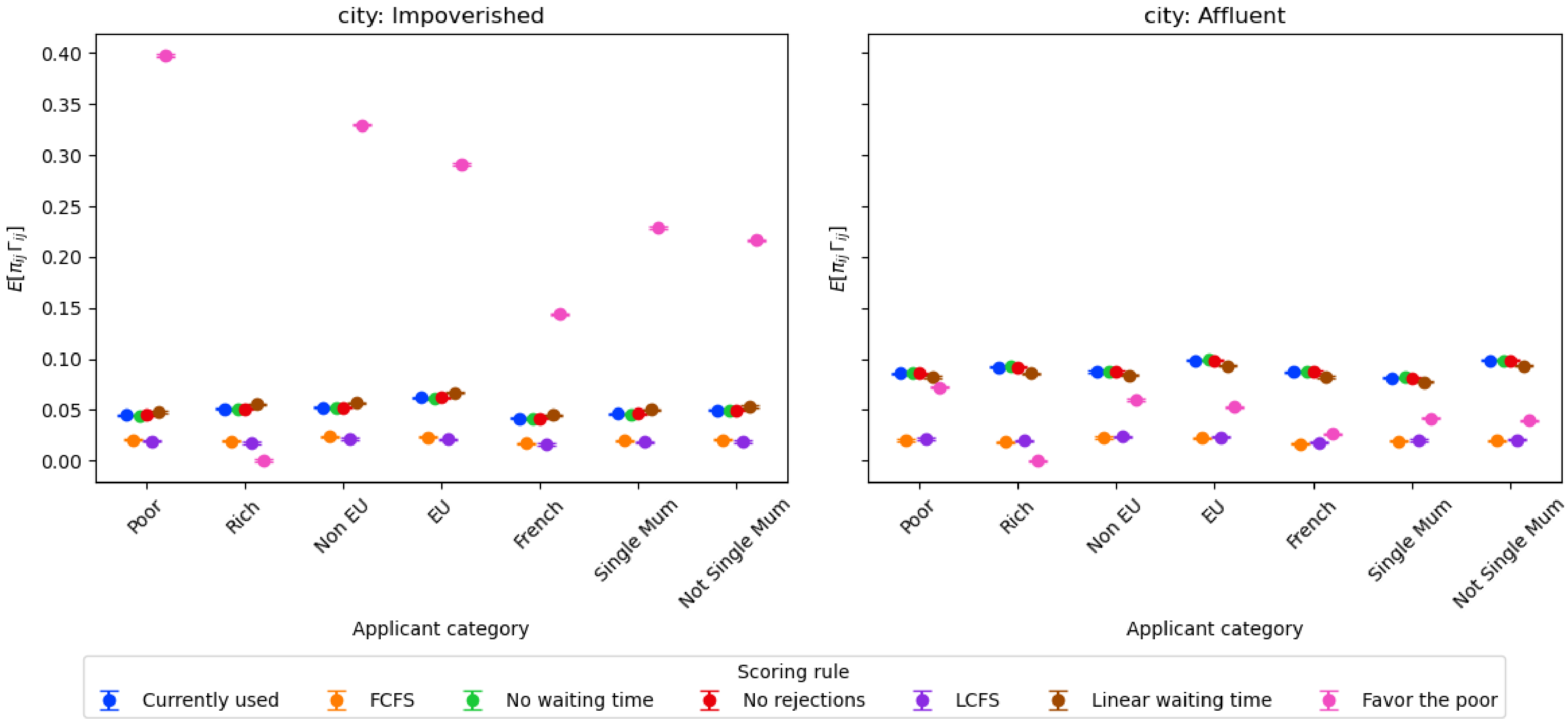


Figure B.21: Probability to apply to Affluent city vs Impoverished

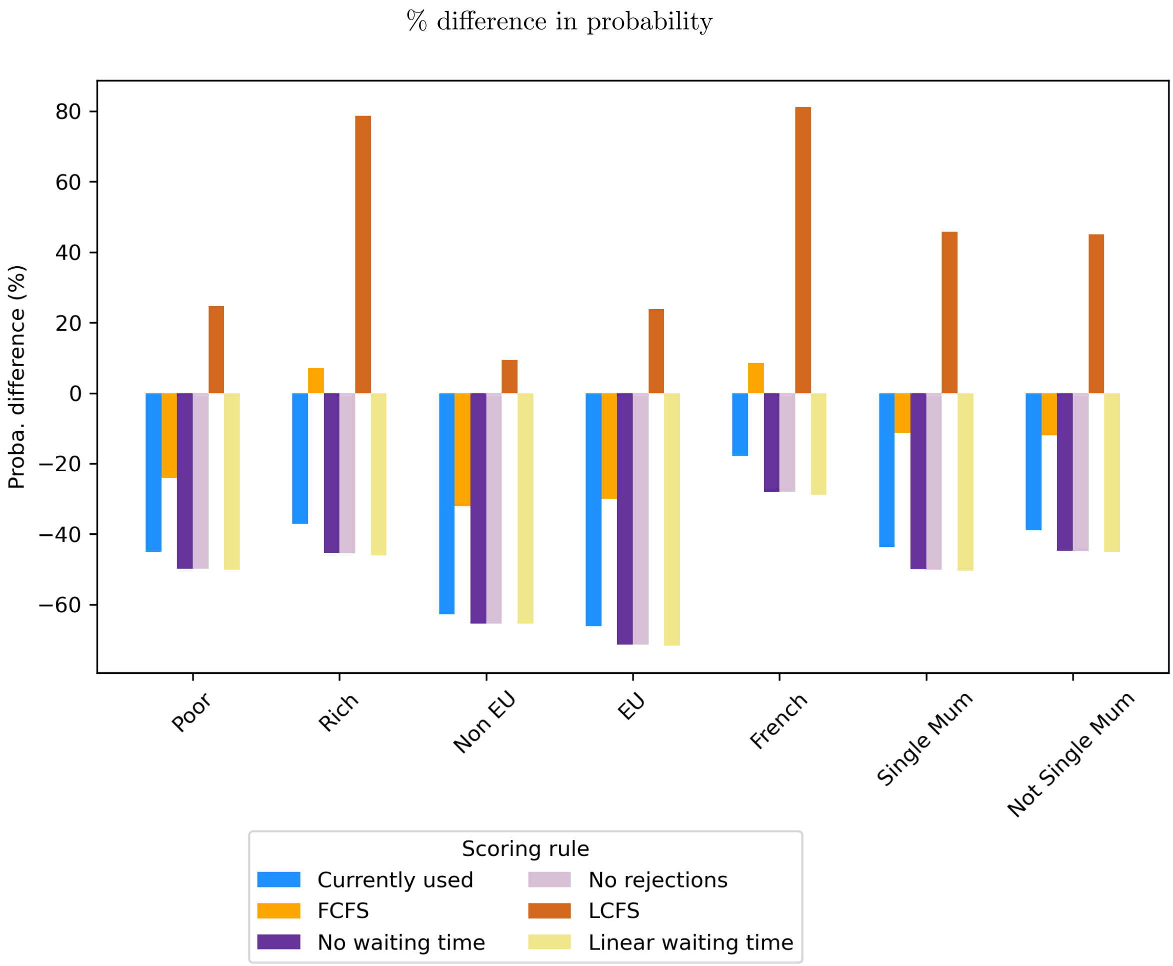


Figure B.22: Probability not to apply to social housing

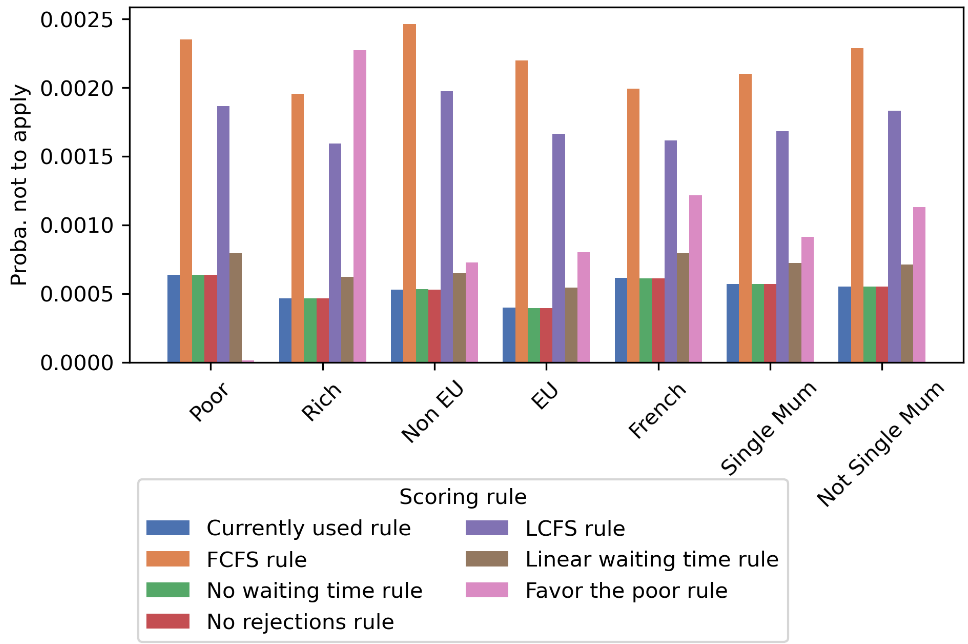
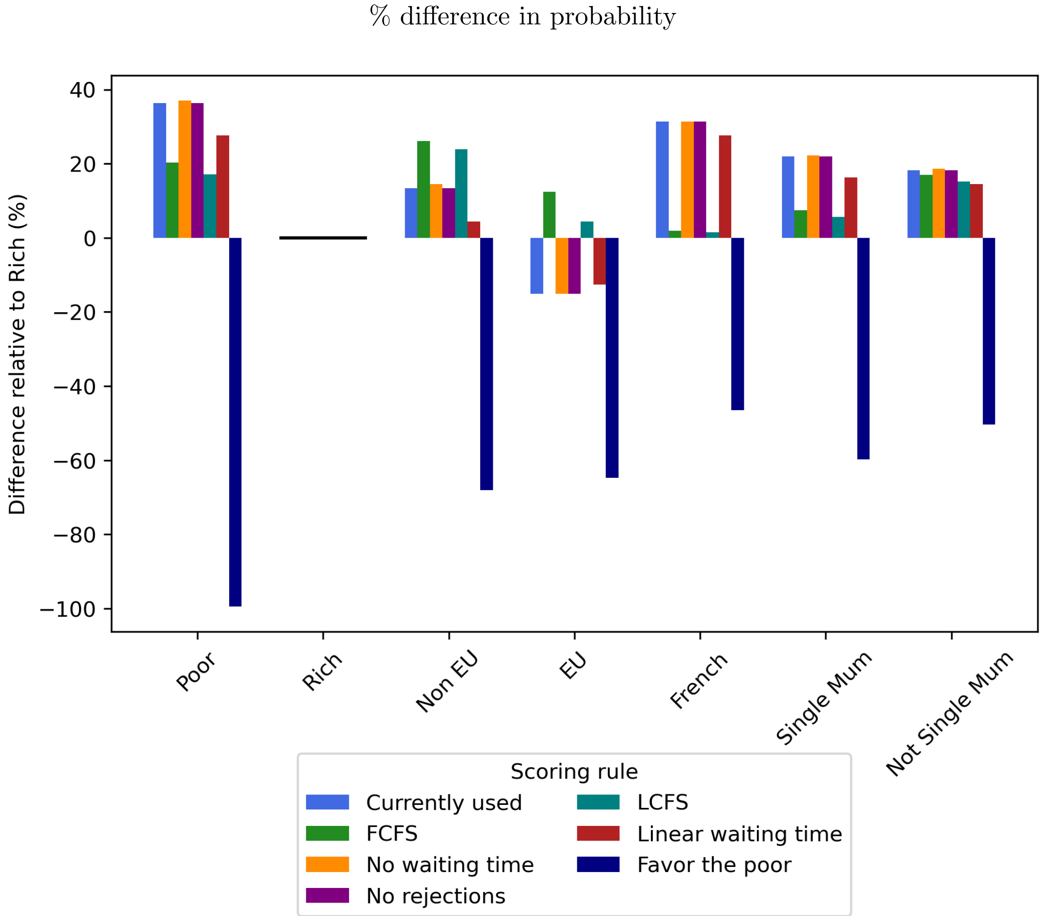


Figure B.23: Probability not to apply, difference across applicants



B.4 Robustness checks

Table B.2: CCP Results for applicants with only one listed city

monthly income/uc spline (thousands of €)	
0 - 1.458	0.0118
	(0.224)
1.458 - 2.09 0.405	
	(0.288)
2.09+	0.337
	(0.190)
job contract type - ref: other	
permanent	0.164*
	(0.0836)
unemployed	0.149
	(0.123)
citizenship - ref: French	
EU	0.195
	(0.140)
Non EU	0.192*
	(0.0797)
waiting time (months)	
0 - 26	-0.00893
	(0.00538)
26+	-0.00634
	(0.00545)
nb rejections	0.0540
	(0.0741)
rent spline (hundreds of €)	
0 - 541	-0.00337
	(0.0190)
541+	-0.0208
	(0.0707)
floor - ref: medium (2-4)	
low (0-1)	0.0329
	(0.0646)
high (4+)	0.157
	(0.140)
Controls: rent, number of rooms per person, building age, building maintenance state, unit in urban priority area, woman,large city, social housing increase in the city, social housing prevalence in city, and interactions.	
N	1,838

Note: "Quality of nearest secondary school" refers to the % of students who got a cum laude distinction or above at the "Brevet des Collèges" (Year 9 examination) in the school closest in distance to the social housing unit. "Urban priority areas" are a translation of "Quartier Prioritaire de la Ville" (QPV). They are neighborhoods identified as poorer and less developed. "Social housing increase in the city" refers to the amount of social housing built in the city between 2012 and 2019 as a percentage increase. "Social housing prevalence in the city" refers to the percentage of housing units that are social housing in the city in 2017.

See Online Appendix for a more detailed table.

Appendix C

Appendix for Chapter 3

C.1 Parallel trends assumptions

C.1.1 Event study

Figure C.1: Applicant characteristics trends

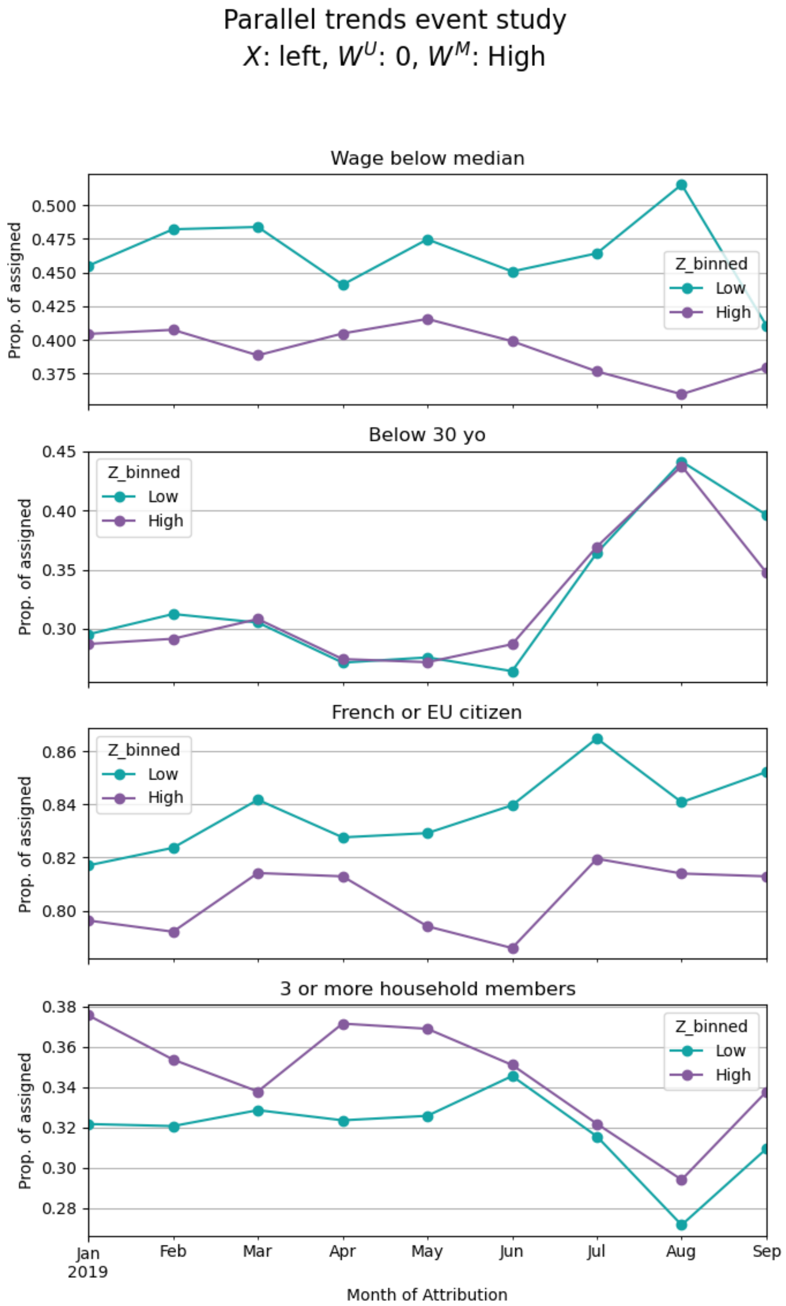


Figure C.2: Applicant characteristics trends

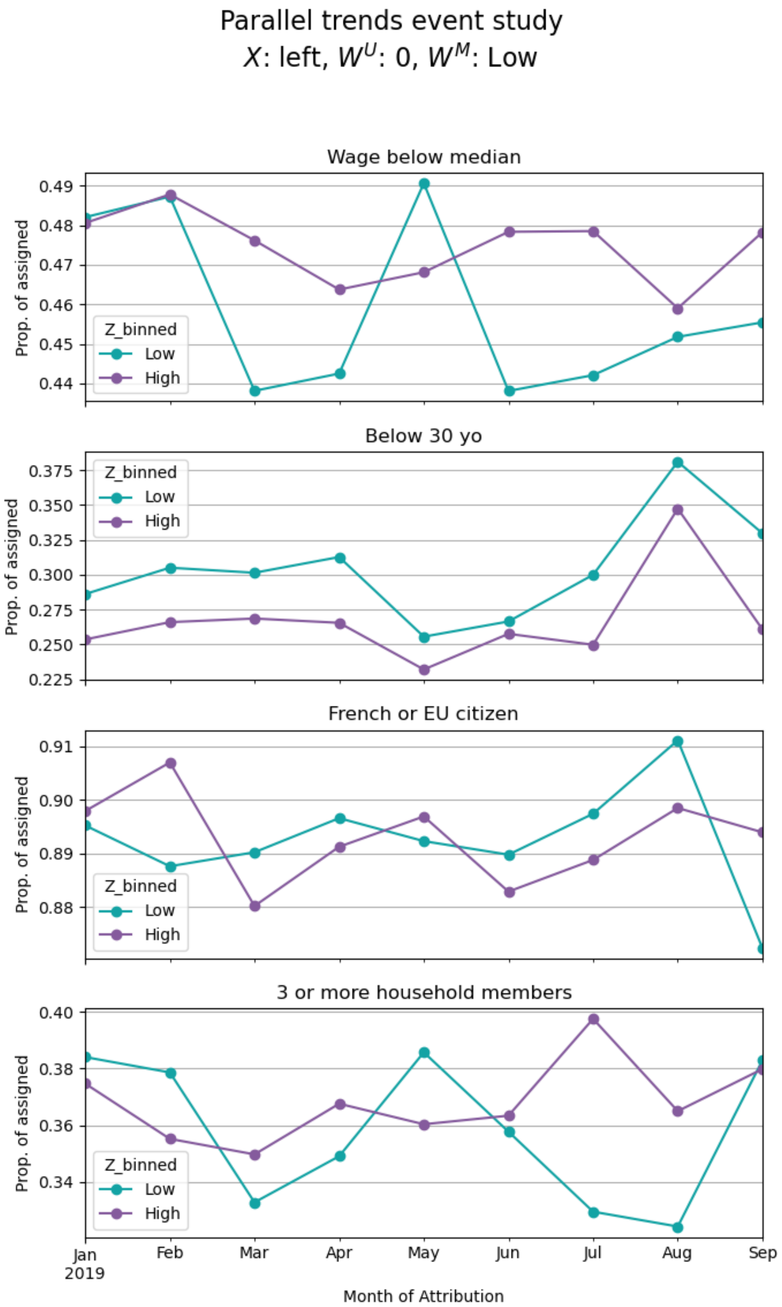


Figure C.3: Applicant characteristics trends

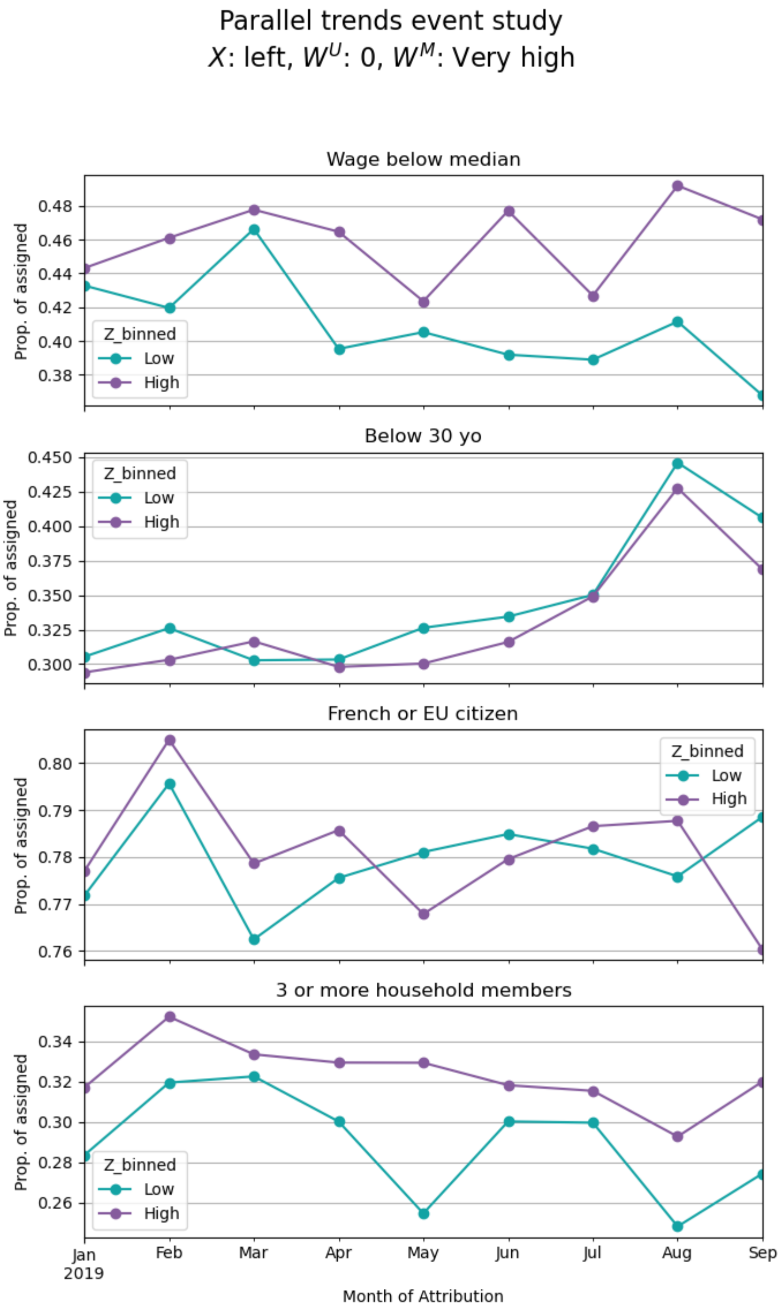


Figure C.4: Applicant characteristics trends

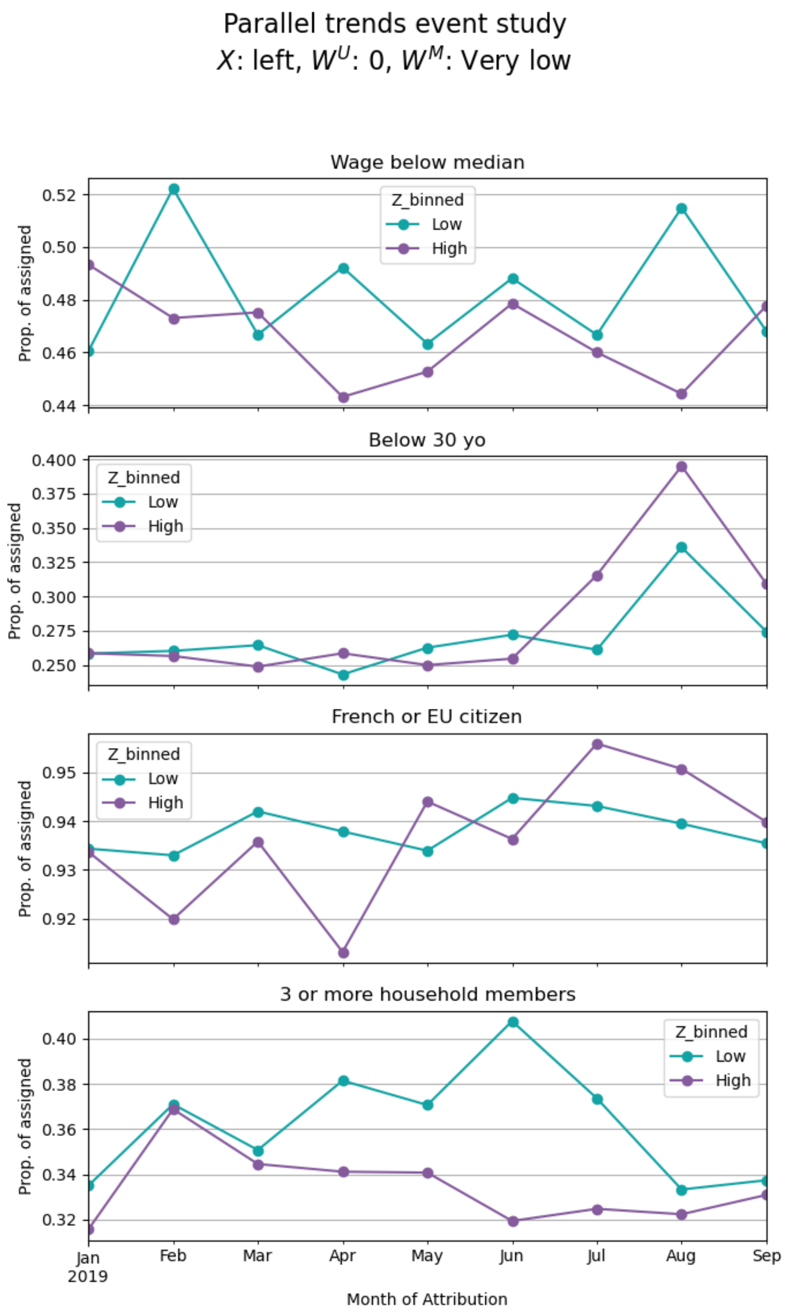


Figure C.5: Applicant characteristics trends

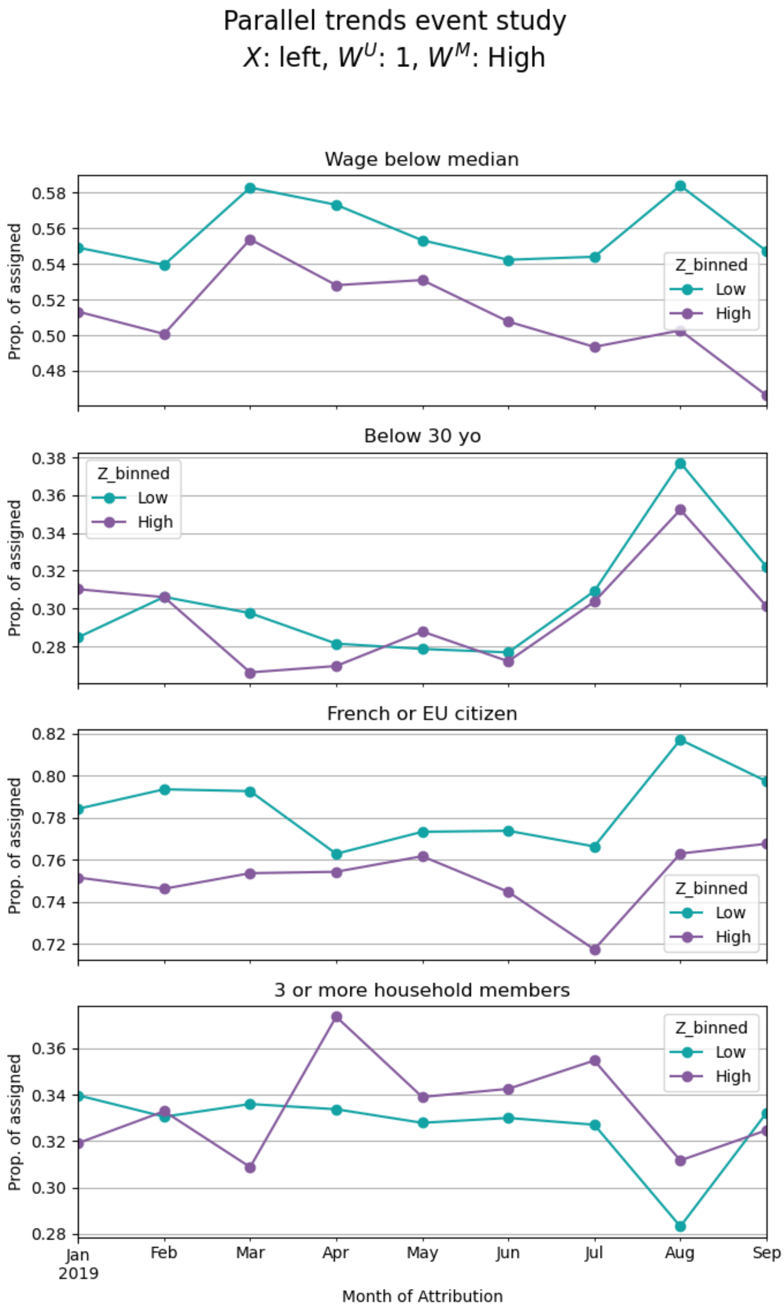


Figure C.6: Applicant characteristics trends

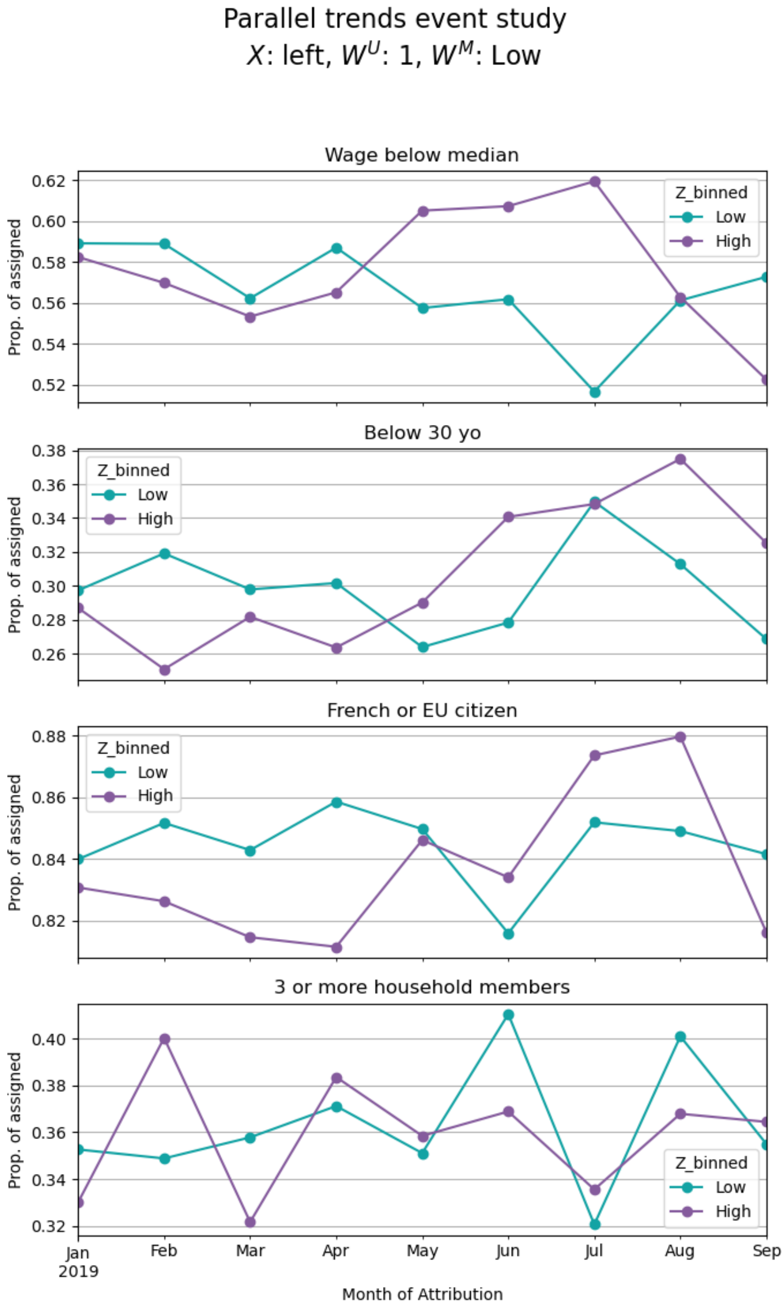


Figure C.7: Applicant characteristics trends

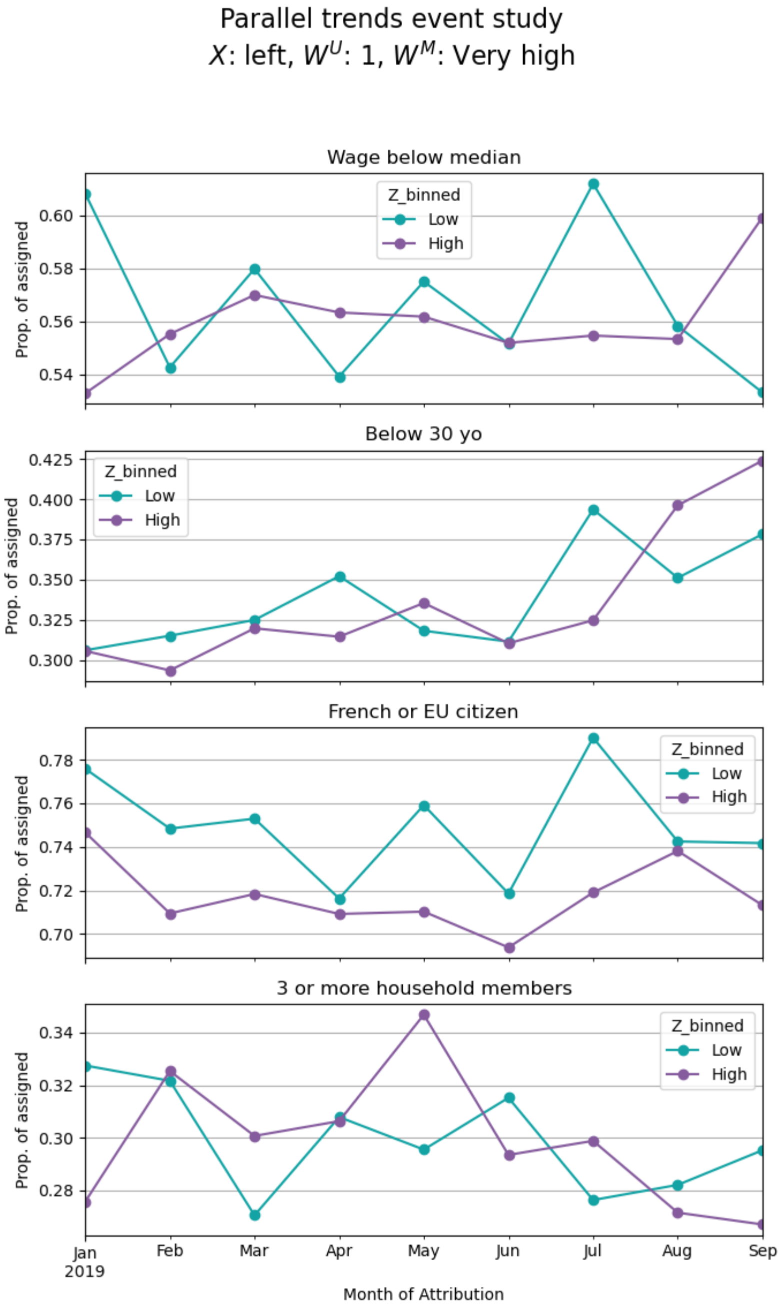


Figure C.8: Applicant characteristics trends

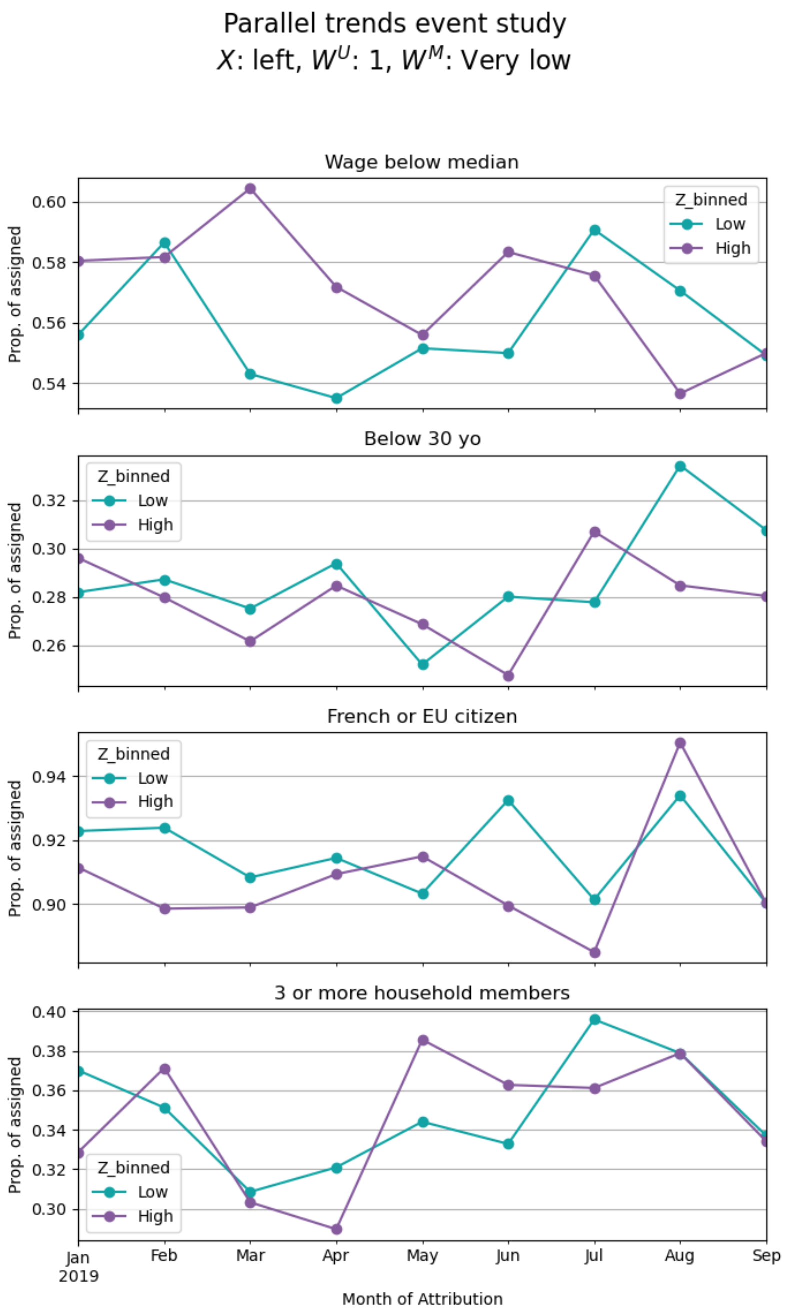


Figure C.9: Applicant characteristics trends

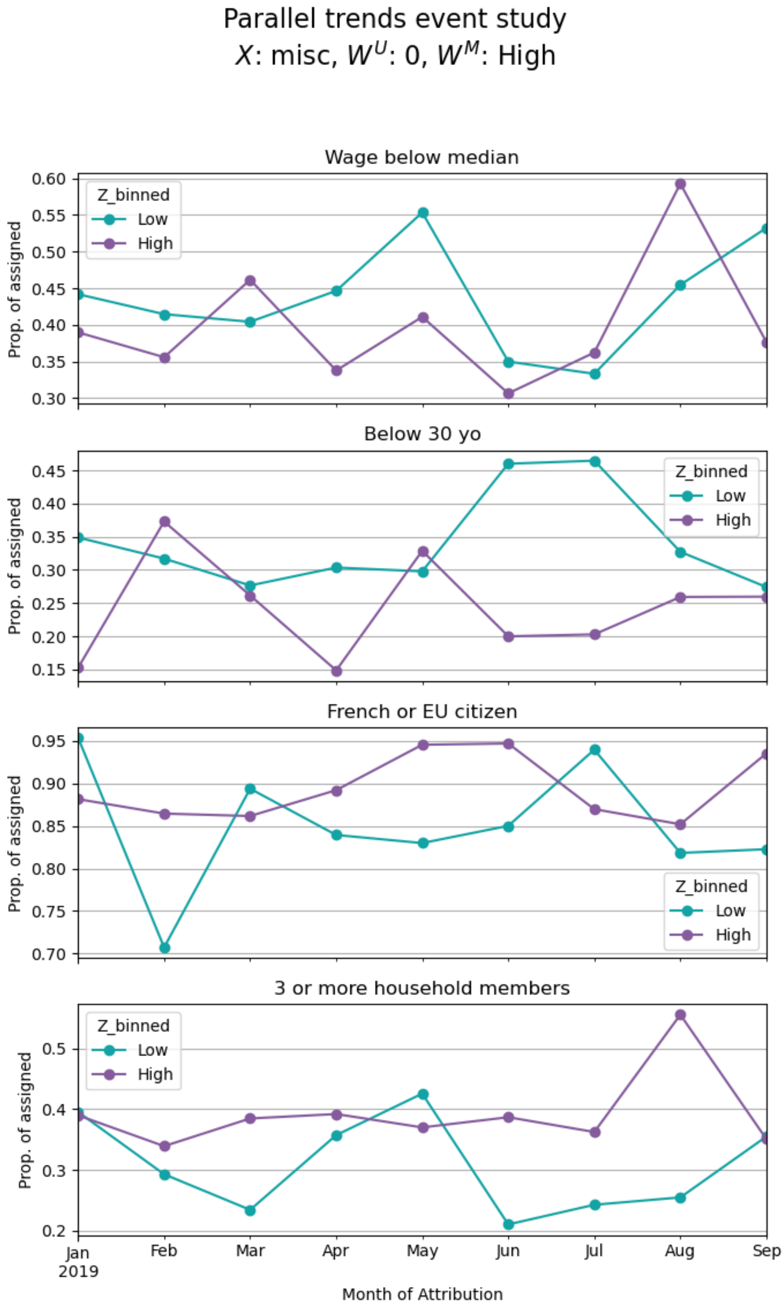


Figure C.10: Applicant characteristics trends

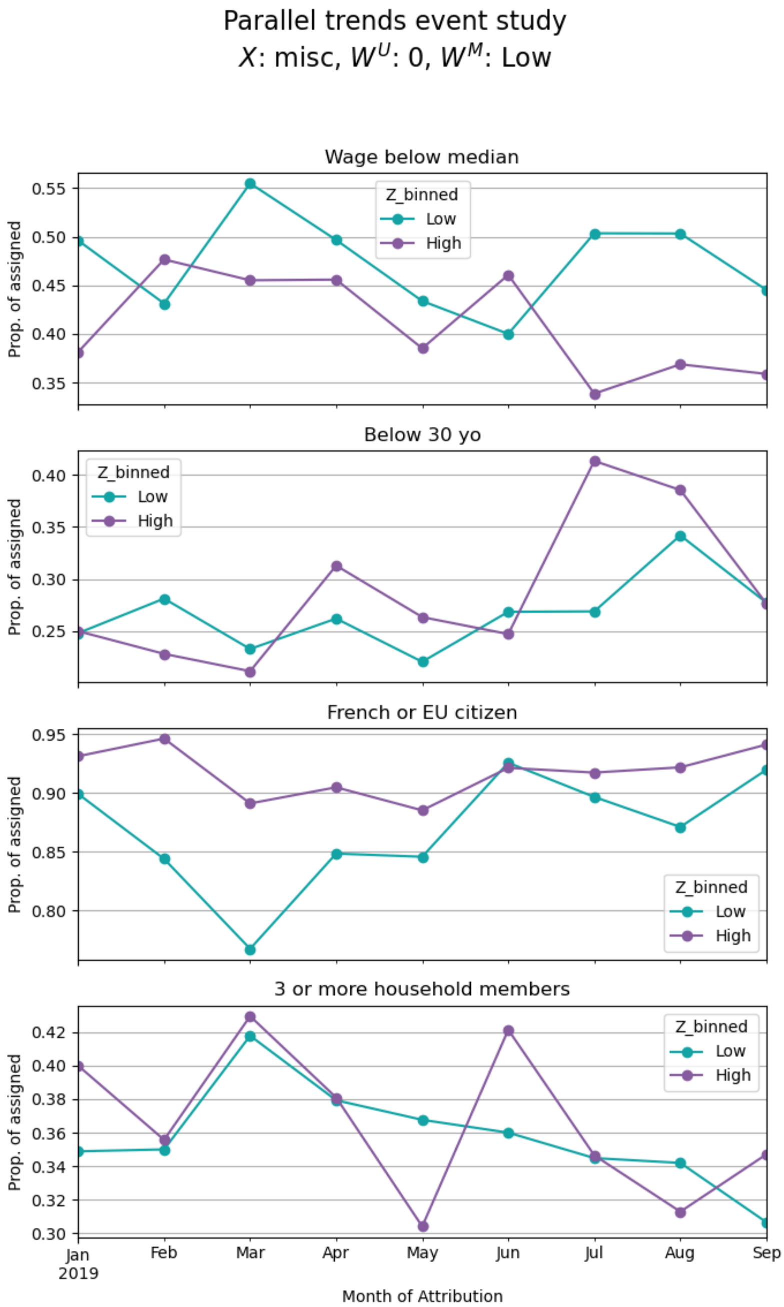


Figure C.11: Applicant characteristics trends

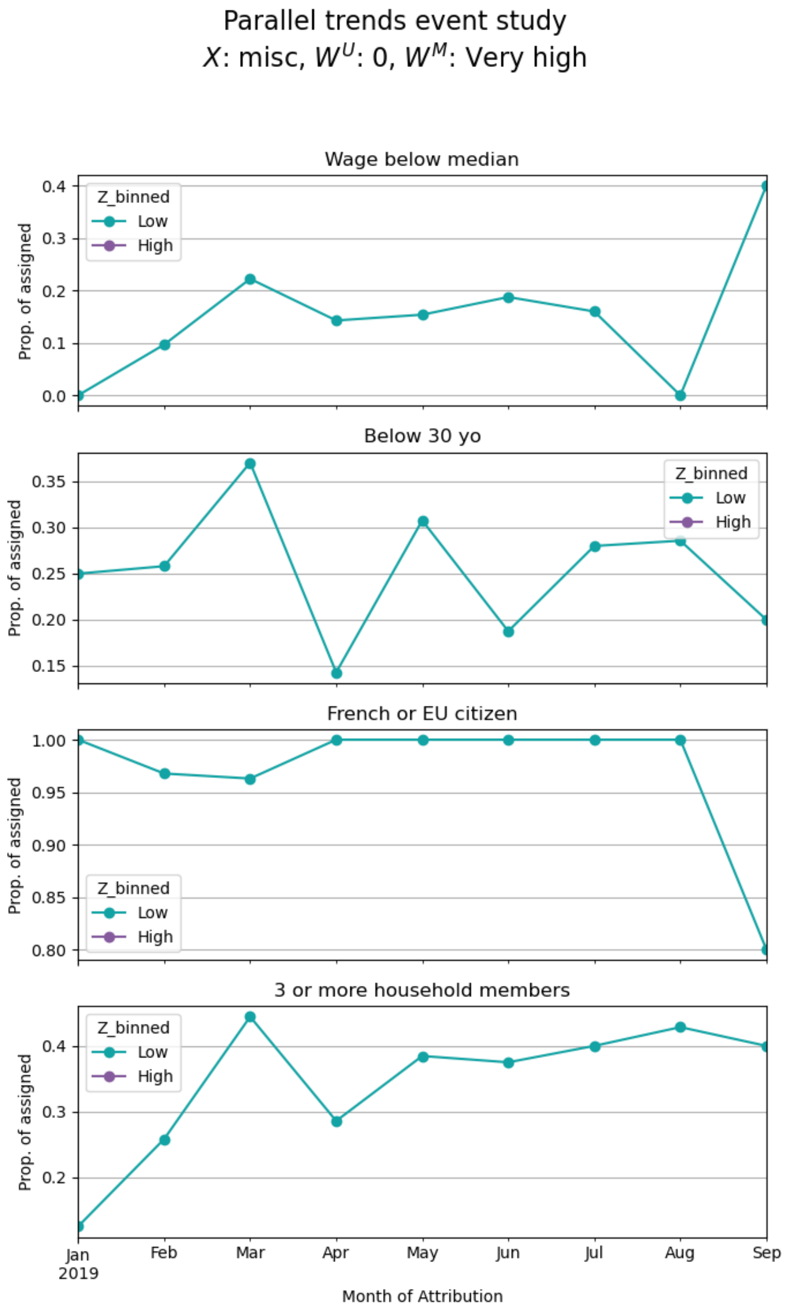


Figure C.12: Applicant characteristics trends

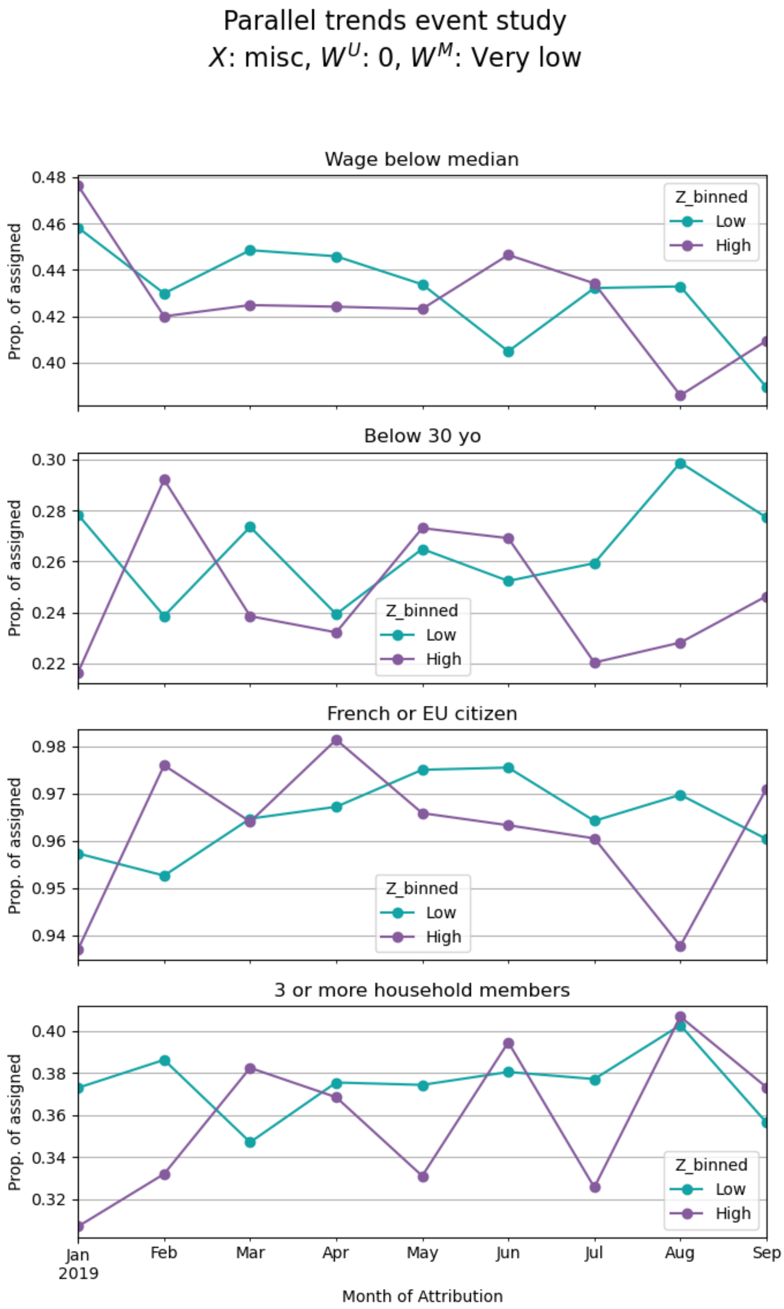


Figure C.13: Applicant characteristics trends

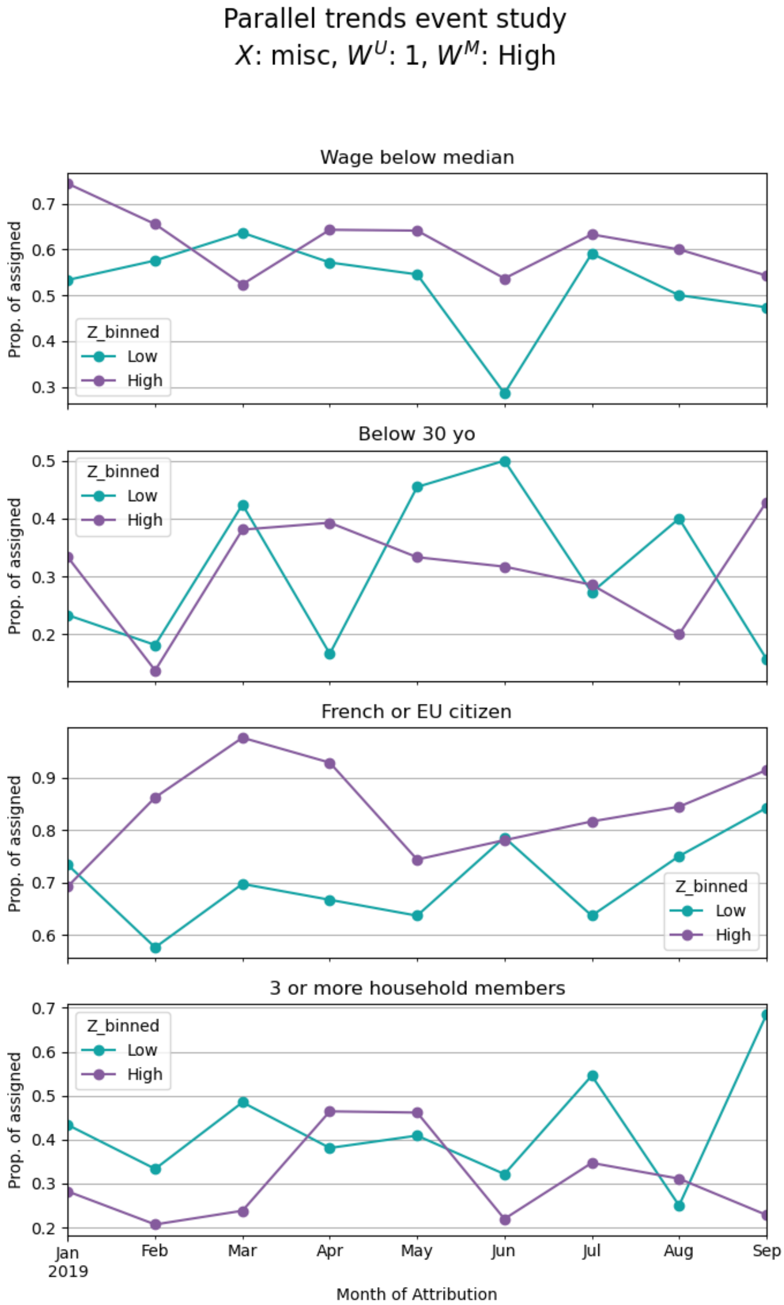


Figure C.14: Applicant characteristics trends

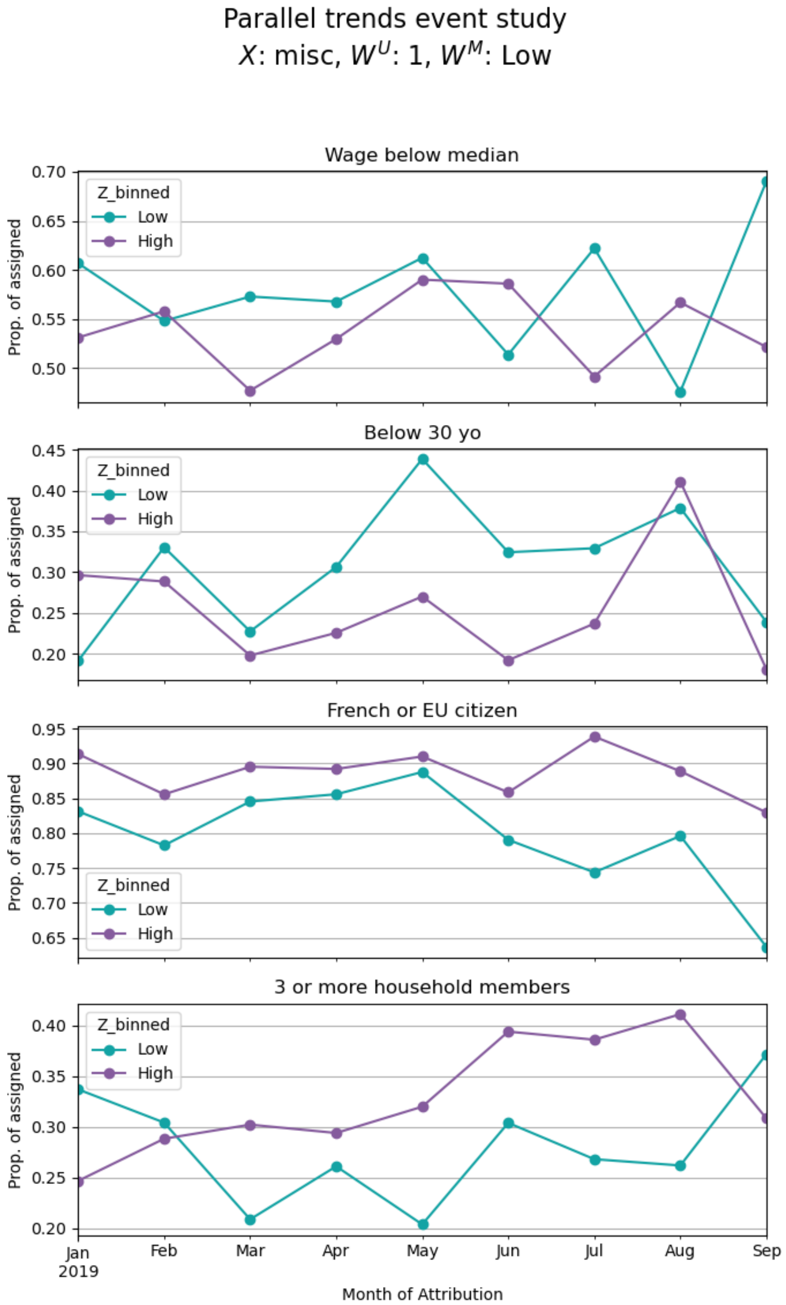


Figure C.15: Applicant characteristics trends

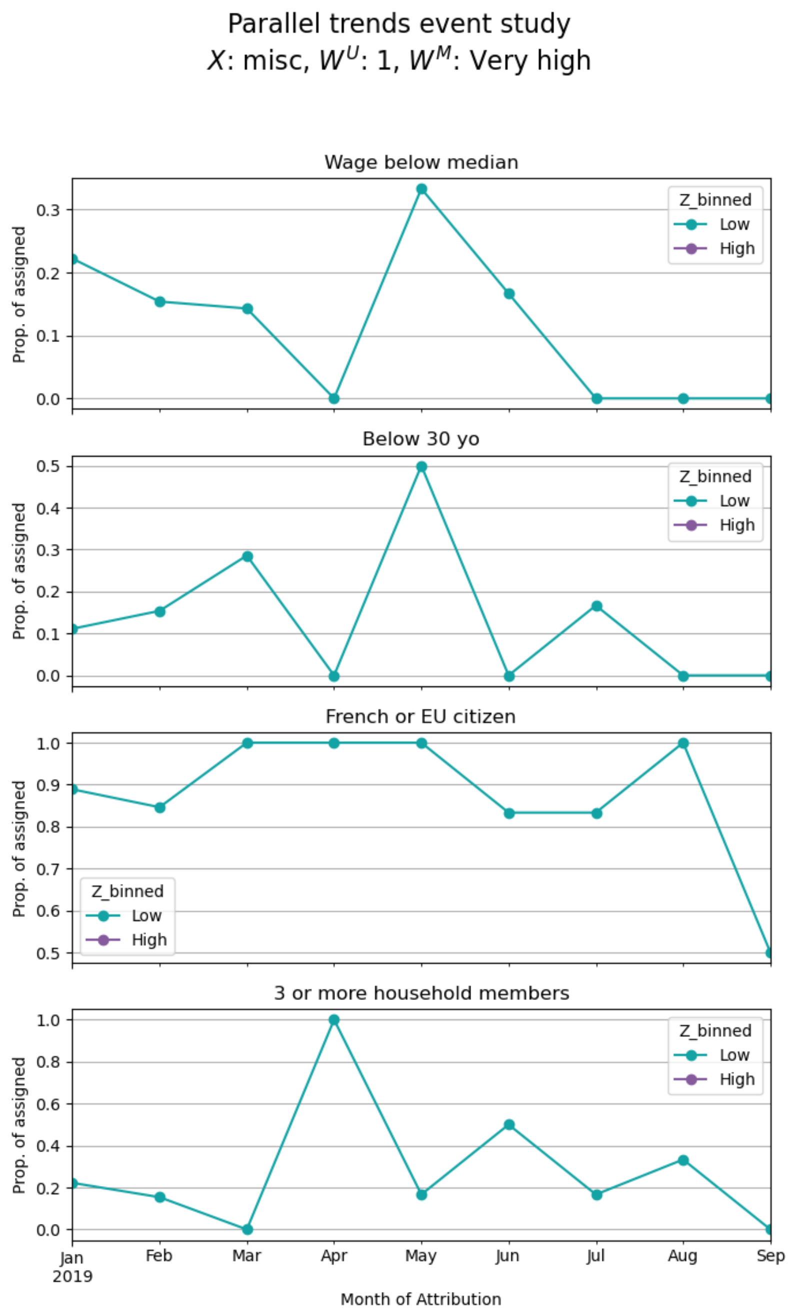


Figure C.16: Applicant characteristics trends

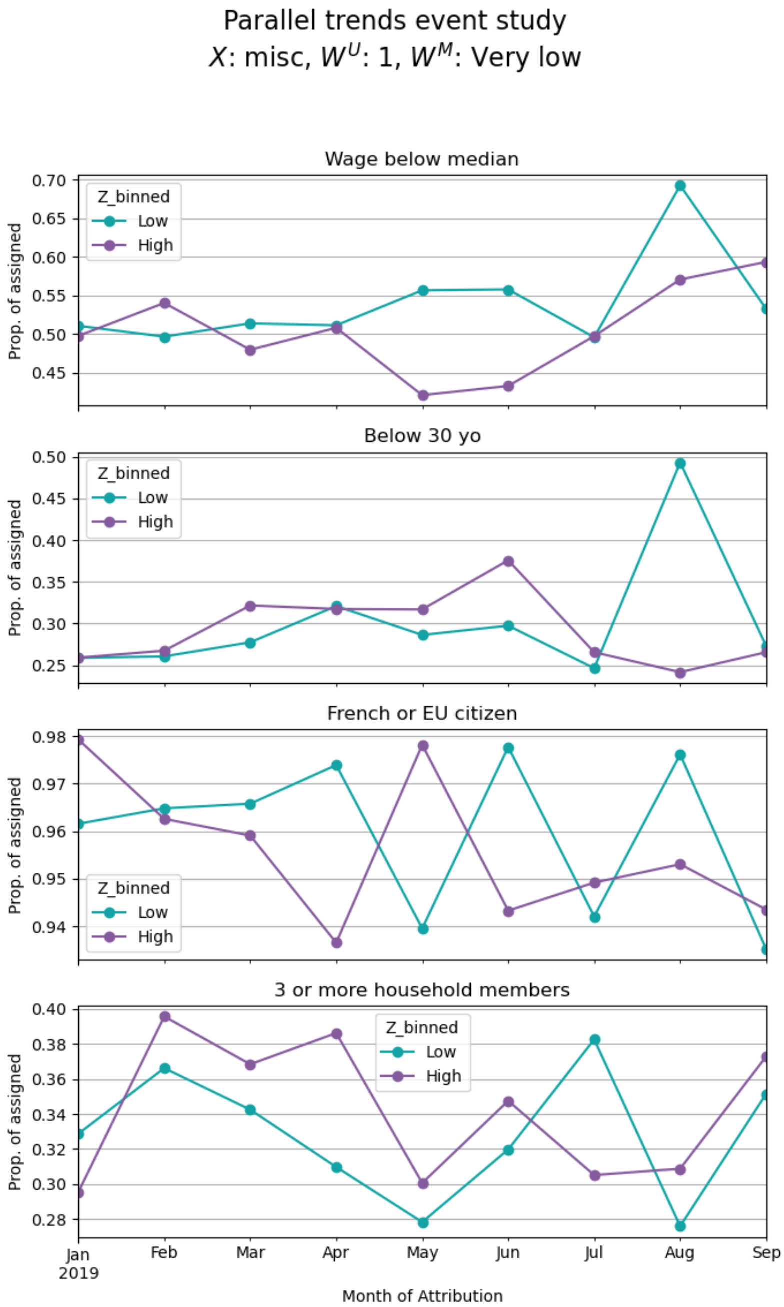


Figure C.17: Applicant characteristics trends

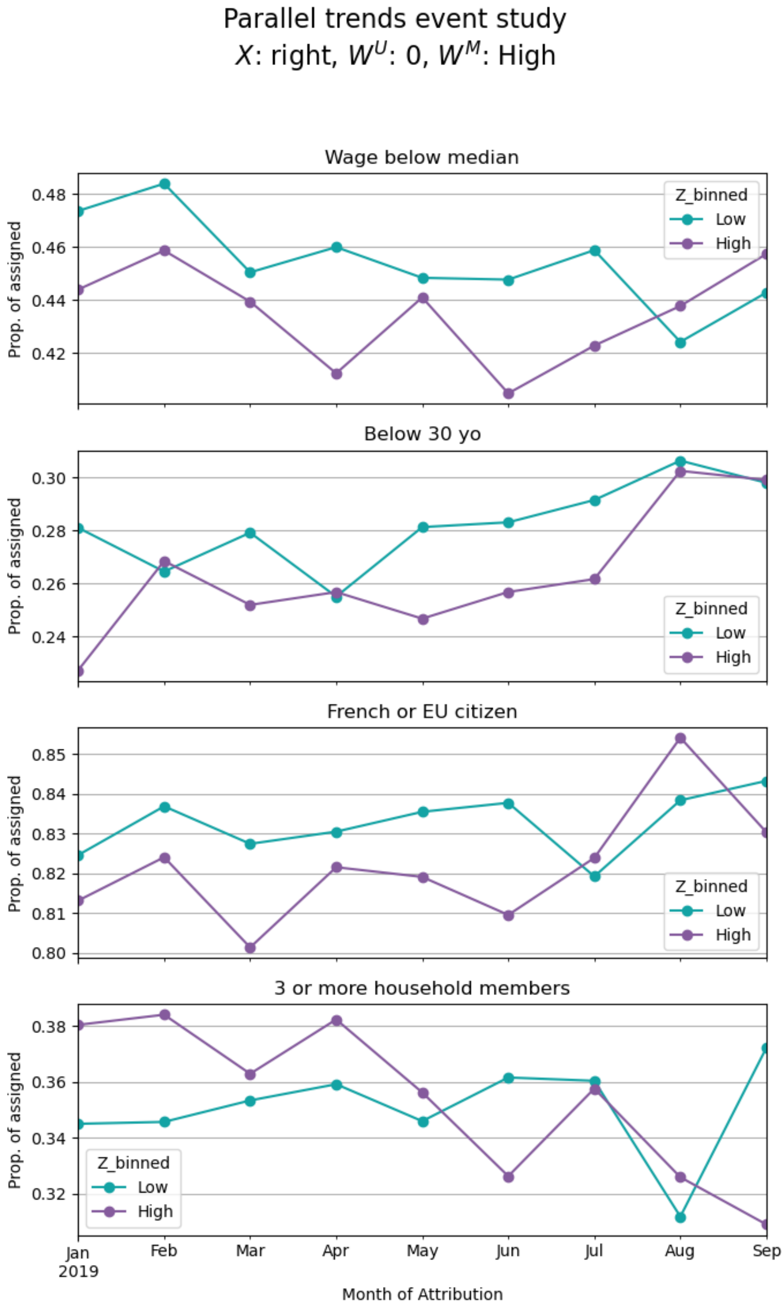


Figure C.18: Applicant characteristics trends

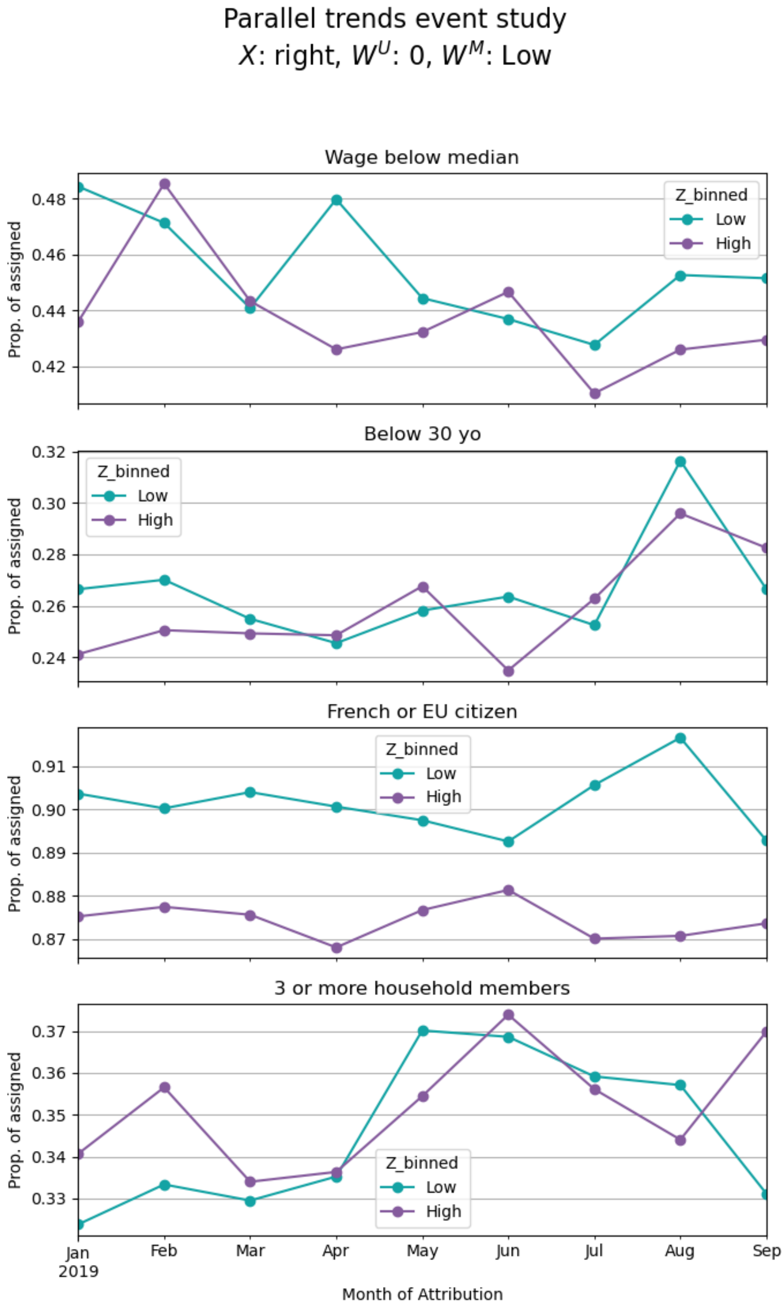


Figure C.19: Applicant characteristics trends

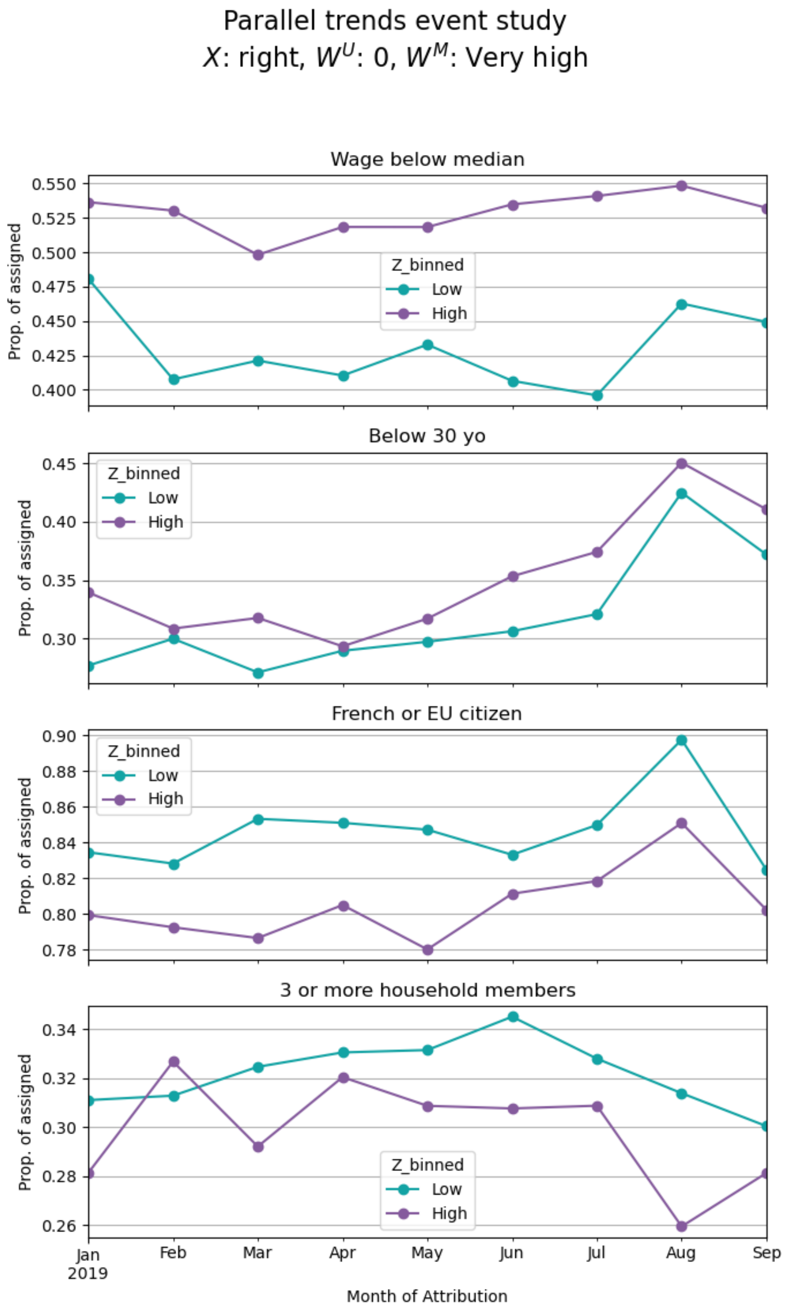


Figure C.20: Applicant characteristics trends

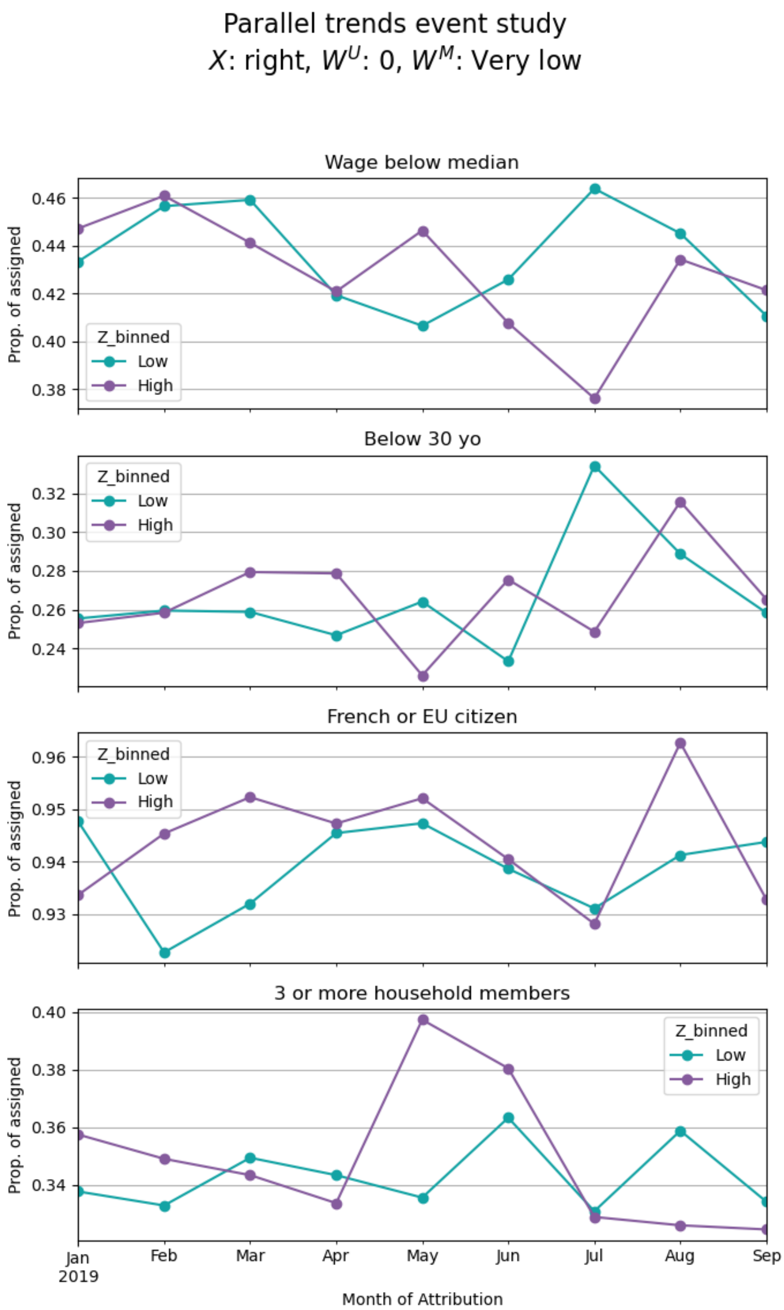


Figure C.21: Applicant characteristics trends

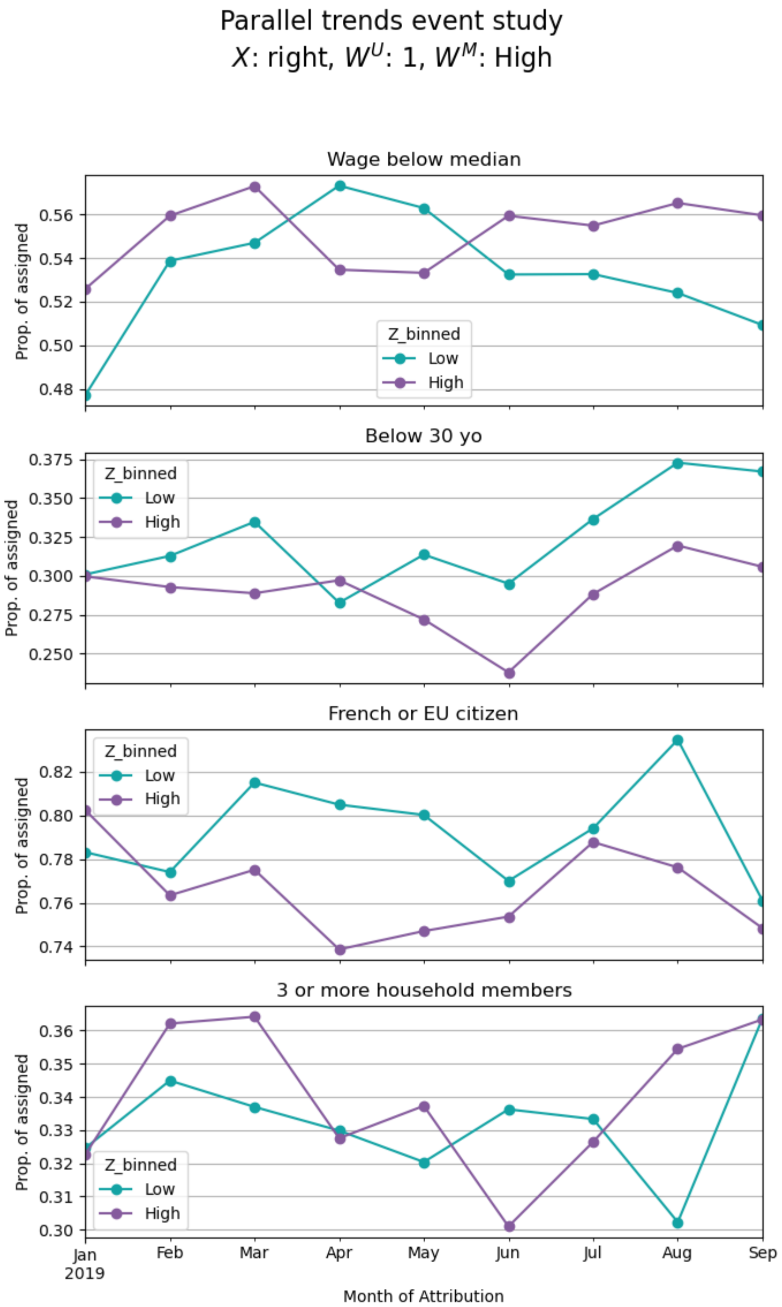


Figure C.22: Applicant characteristics trends

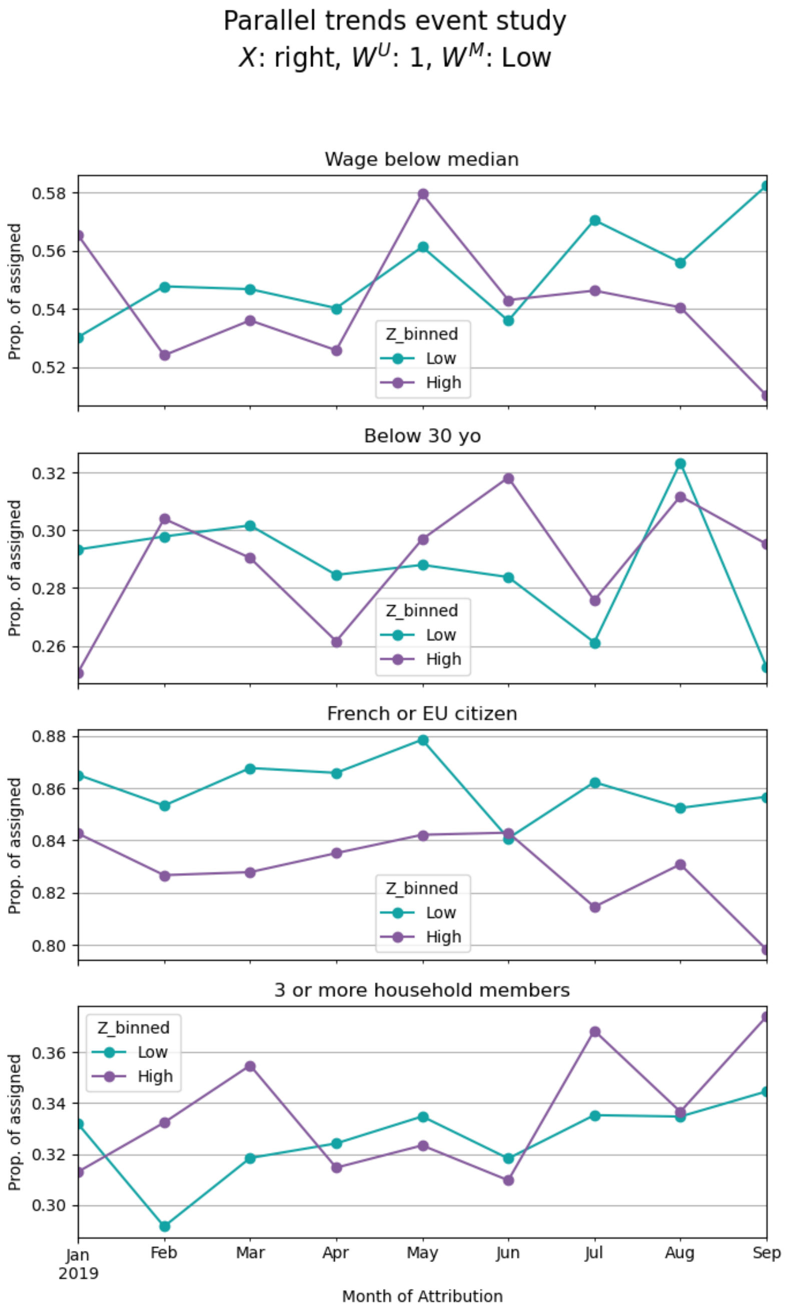


Figure C.23: Applicant characteristics trends

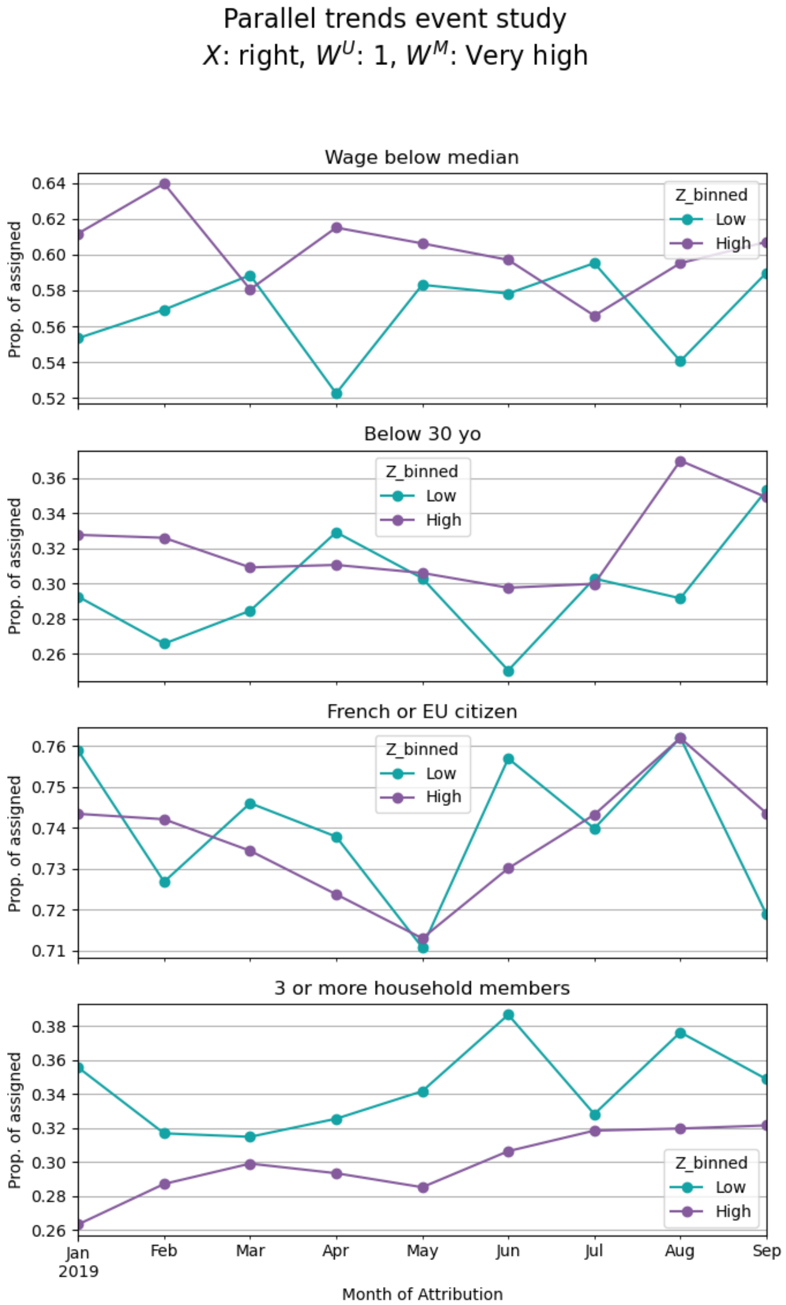
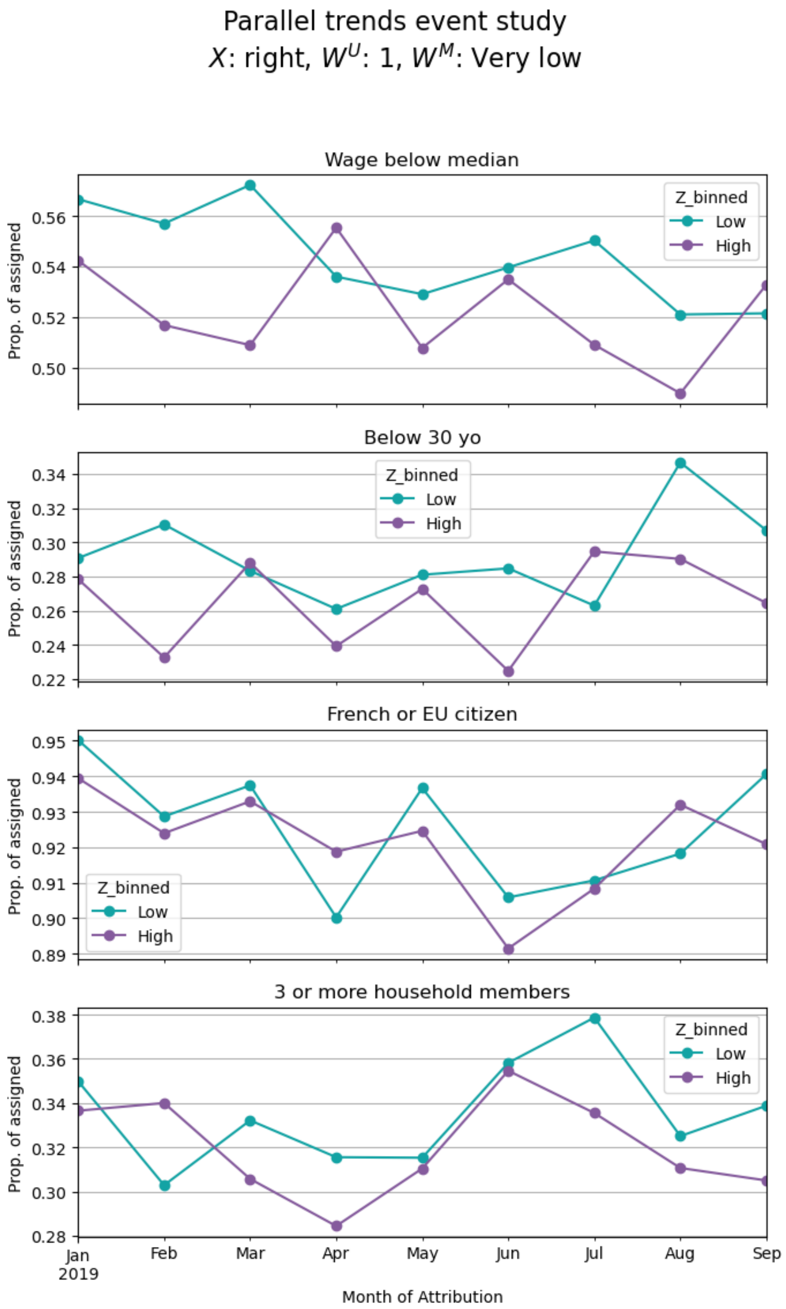


Figure C.24: Applicant characteristics trends



C.1.2 Difference-in-differences

Figure C.25: Applicant characteristics trends

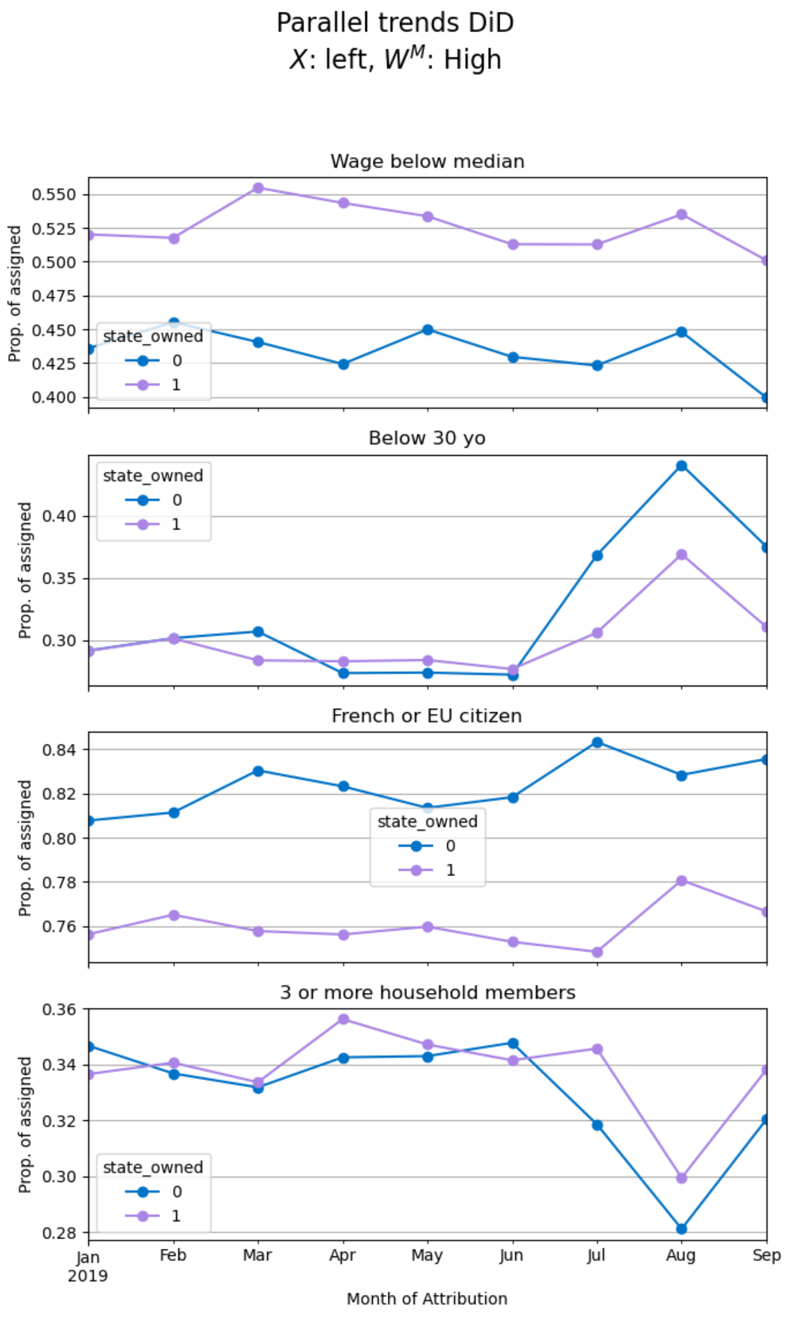


Figure C.26: Applicant characteristics trends

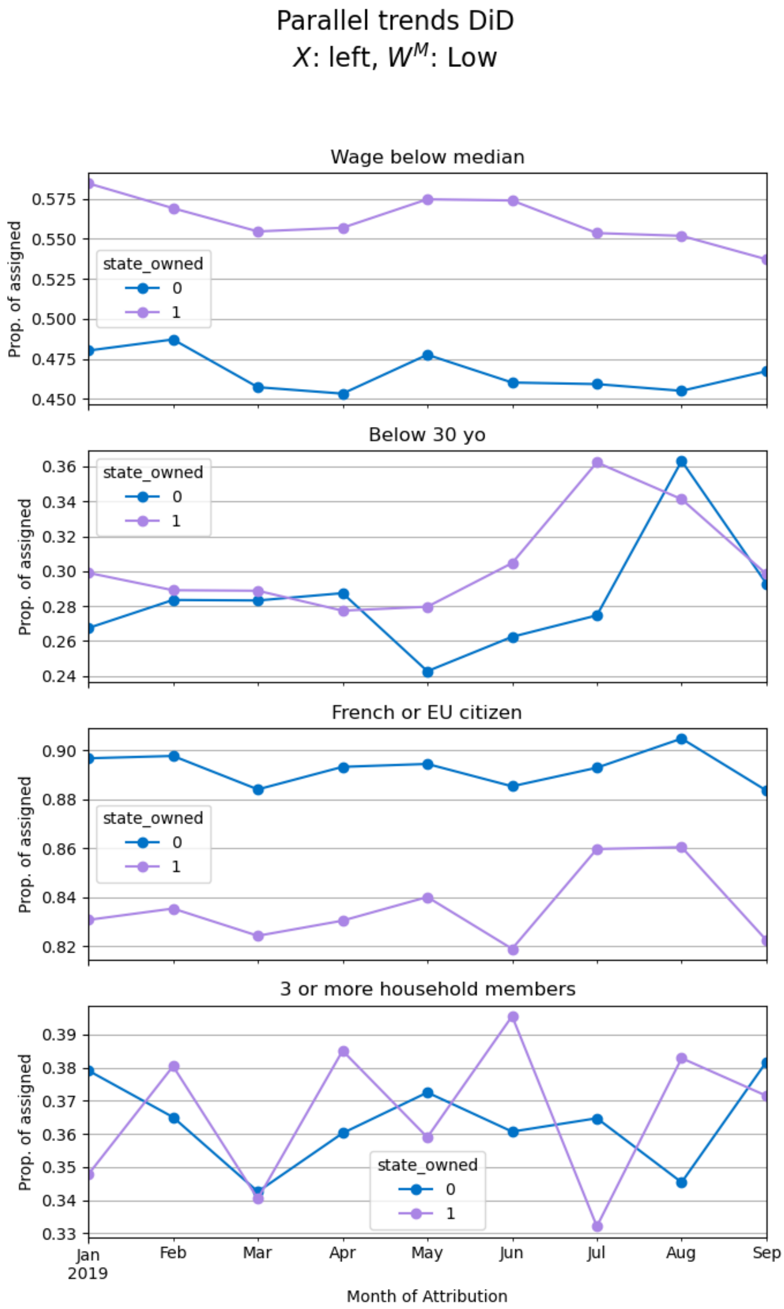


Figure C.27: Applicant characteristics trends

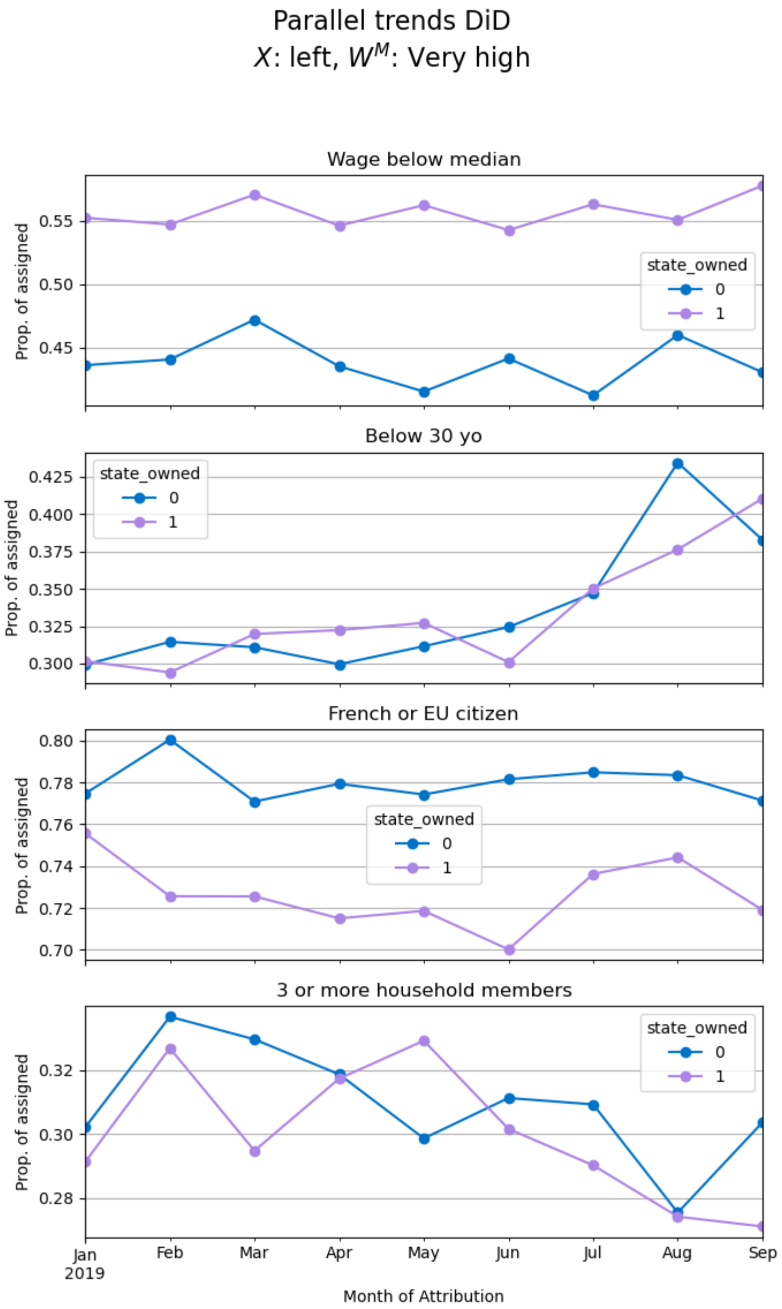


Figure C.28: Applicant characteristics trends

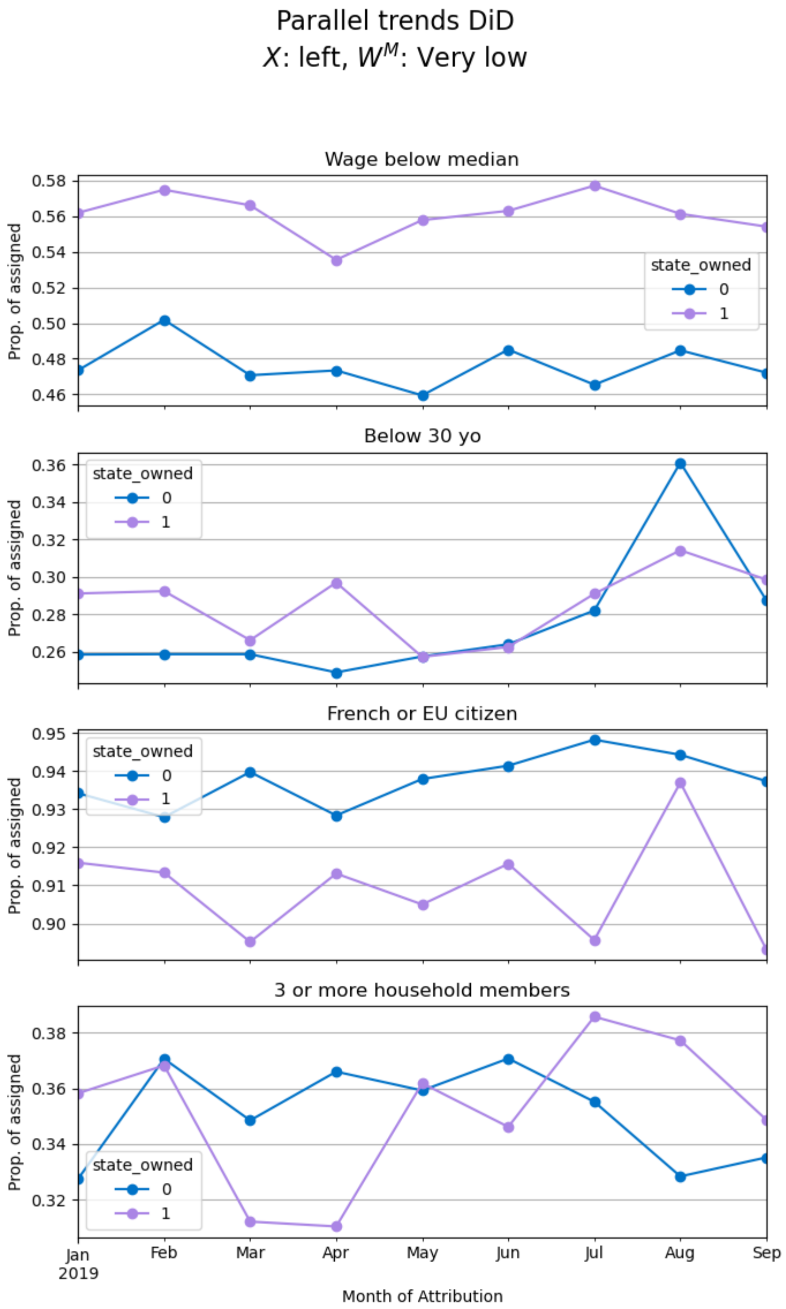


Figure C.29: Applicant characteristics trends

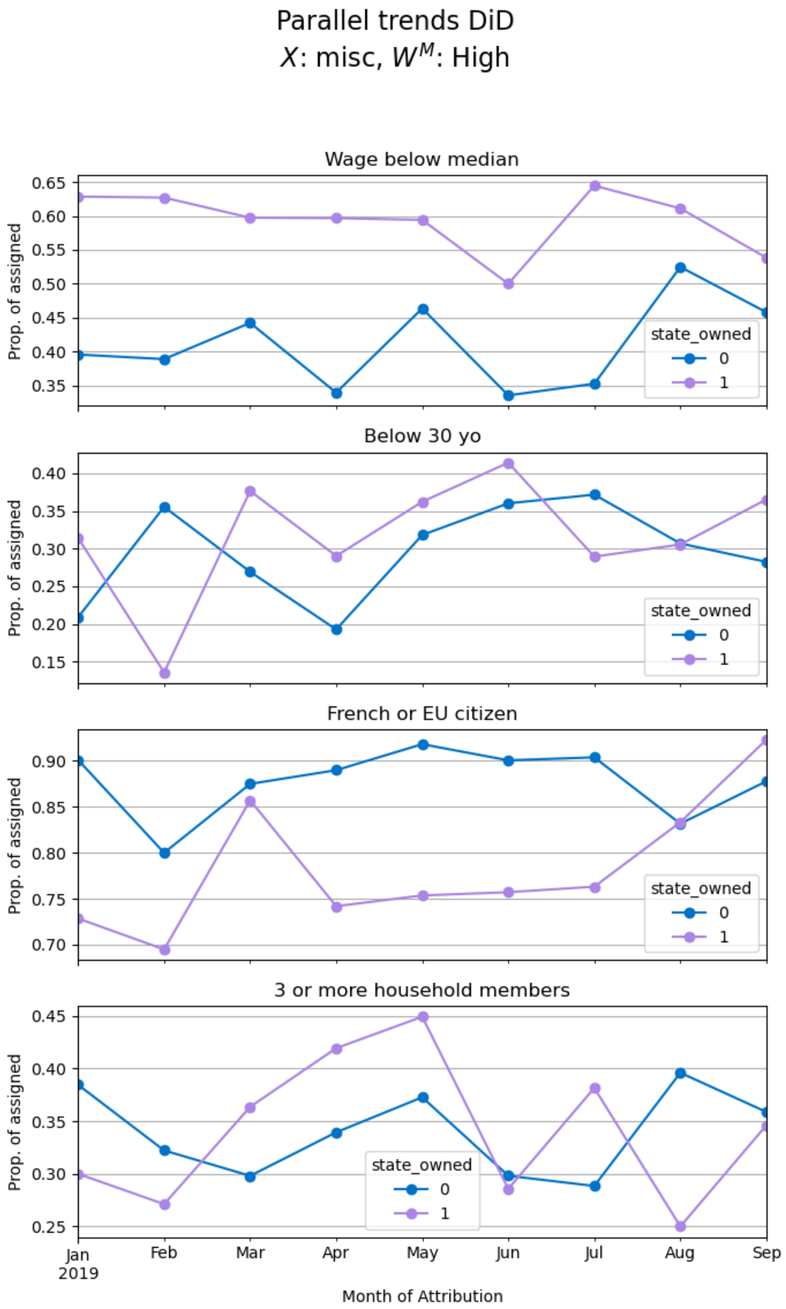


Figure C.30: Applicant characteristics trends

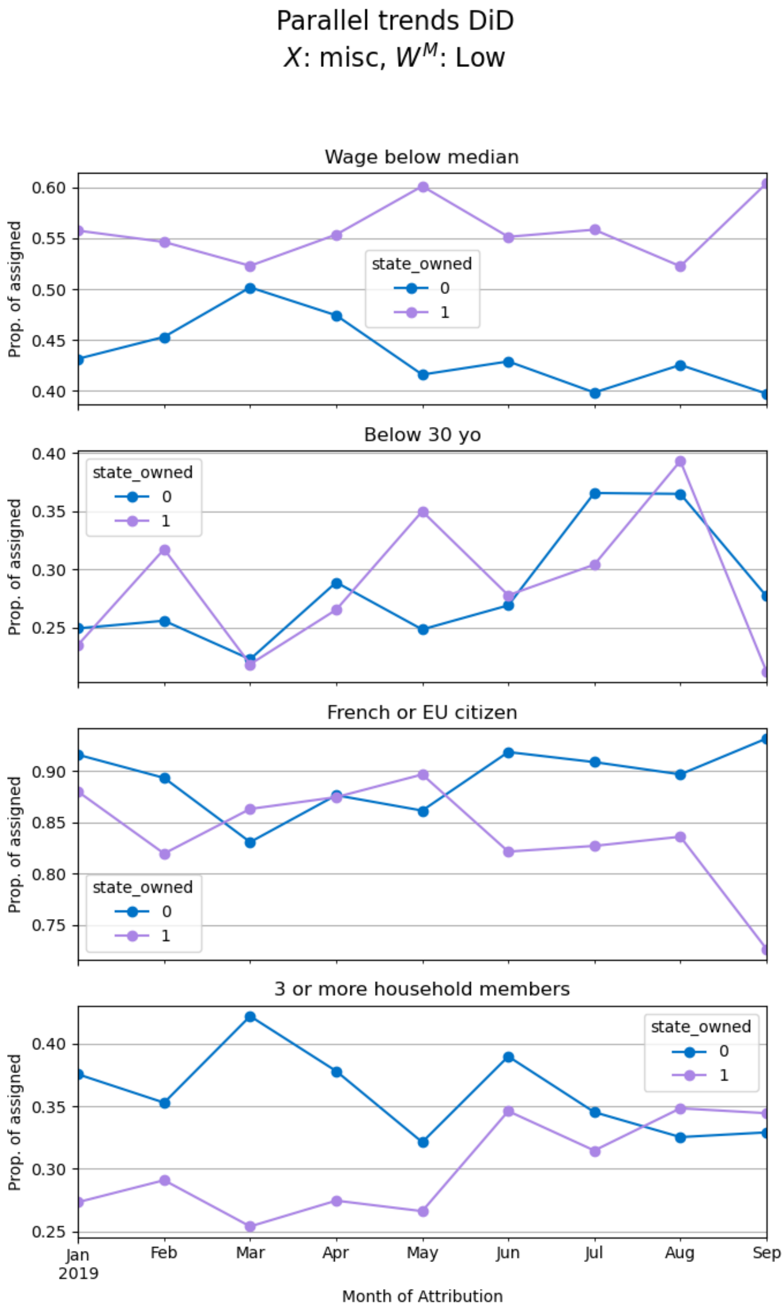


Figure C.31: Applicant characteristics trends

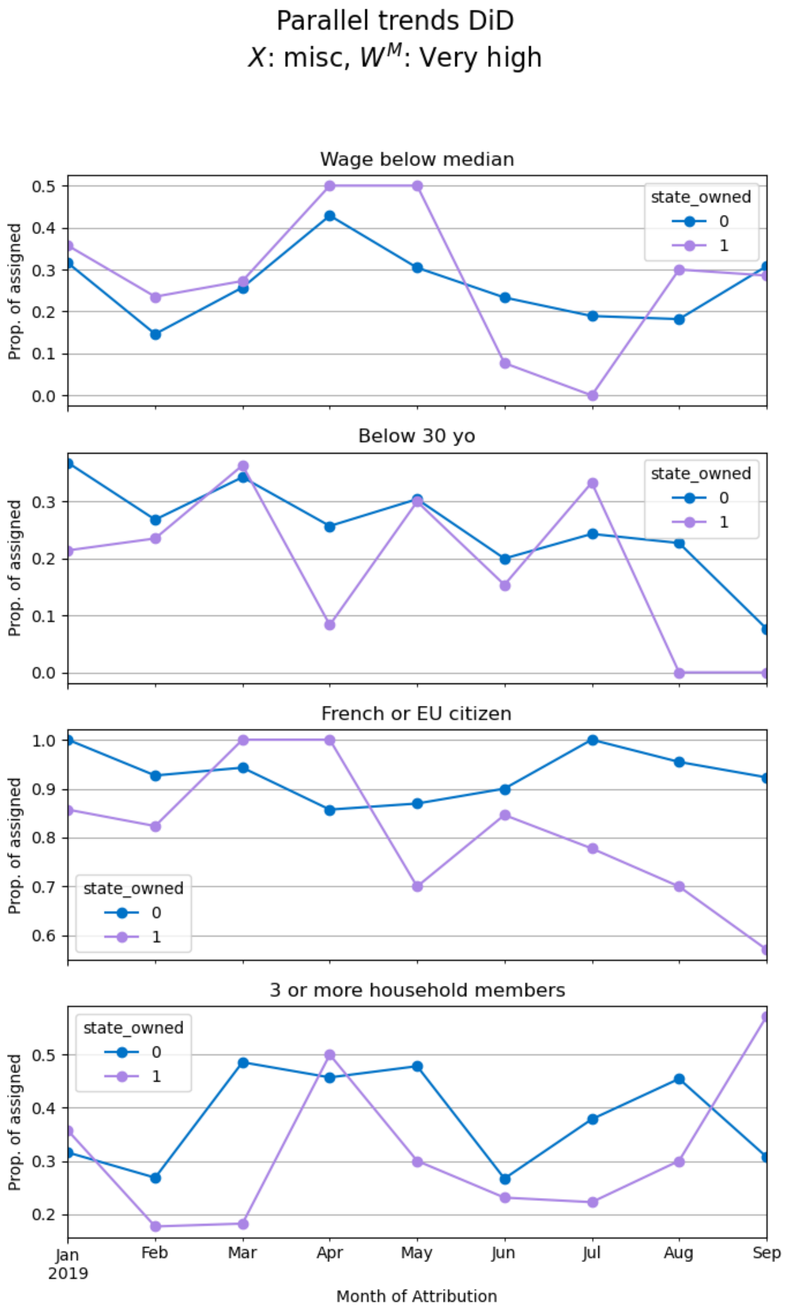


Figure C.32: Applicant characteristics trends

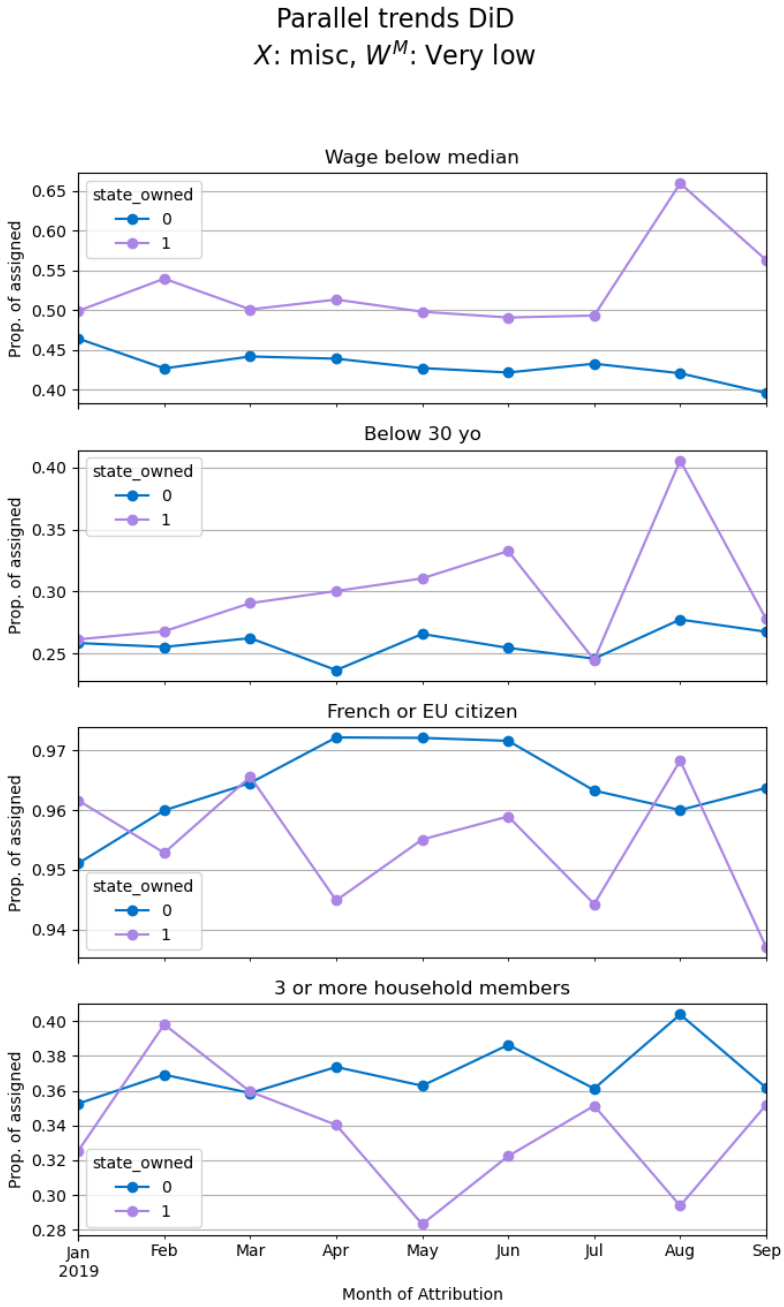


Figure C.33: Applicant characteristics trends

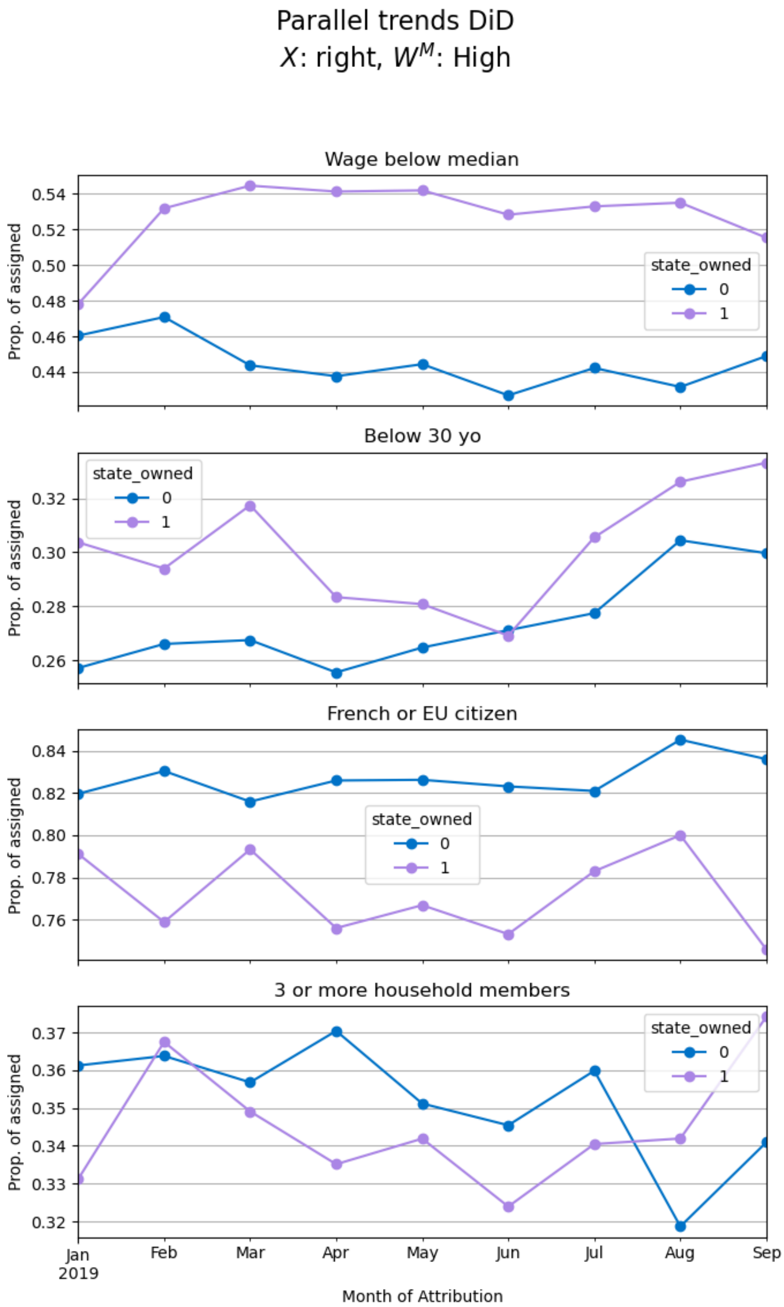


Figure C.34: Applicant characteristics trends

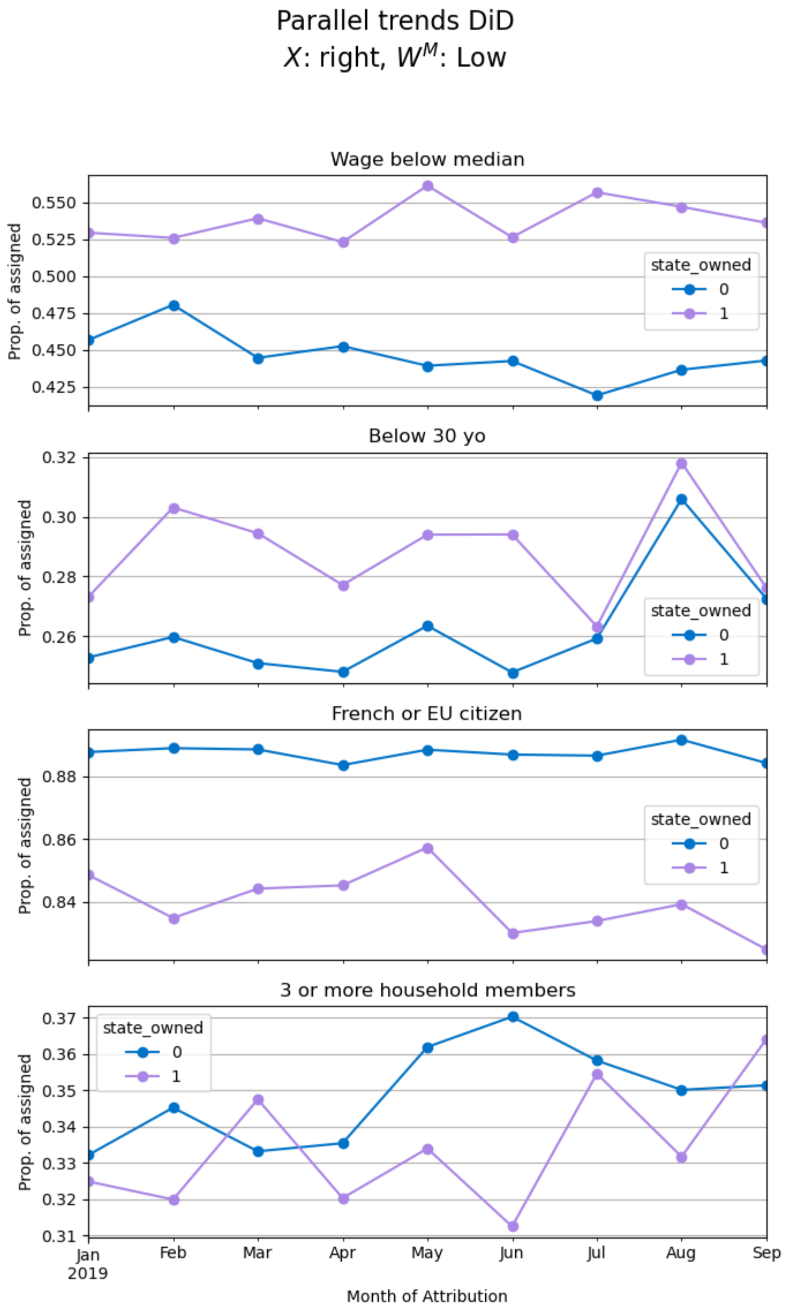


Figure C.35: Applicant characteristics trends

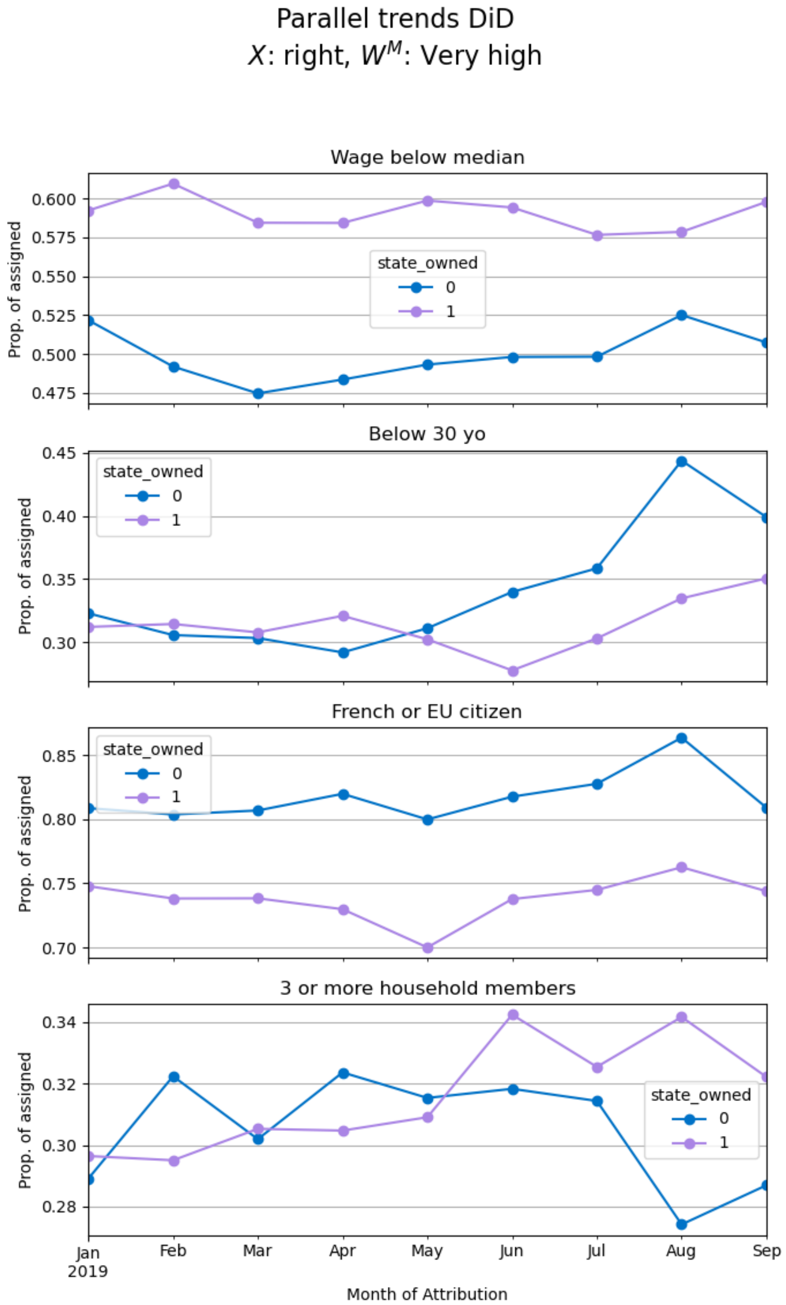
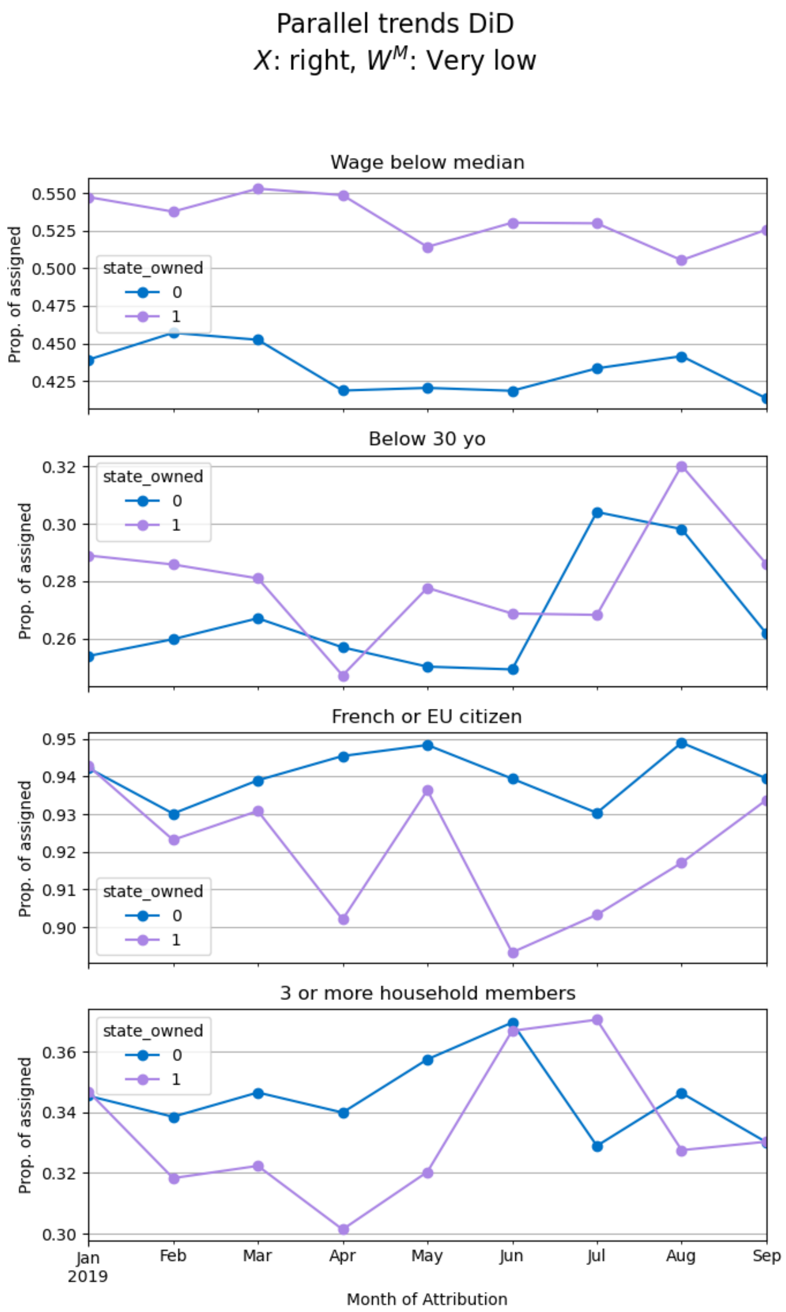


Figure C.36: Applicant characteristics trends



C.2 Results

C.2.1 Event study estimates

Table C.1: Event study estimates

	poor	young	can vote	family
D × Z	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Center × D × Z		reference category		
X = Far Right × D × Z	-0.004*	0.002	0.000	-0.001
	(0.002)	(0.002)	(0.001)	(0.002)
X = Left × D × Z	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Misc × D × Z	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Right × D × Z	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Unstated × D × Z		ommitted		
X = Center × D	0.060	-0.071*	0.016	-0.056
	(0.032)	(0.030)	(0.023)	(0.031)
X = Far Right × D	0.095*	-0.100**	0.020	-0.040

* p<0.05, ** p<0.01, *** p<0.001

Table C.1: Event study estimates (continued)

	poor	young	can vote	family
	(0.040)	(0.038)	(0.029)	(0.039)
X = Left × D	0.068*	-0.080**	0.017	-0.050
	(0.031)	(0.030)	(0.023)	(0.031)
X = Misc × D	0.058	-0.063*	0.015	-0.056
	(0.032)	(0.030)	(0.023)	(0.031)
X = Right × D	0.059	-0.075*	0.020	-0.052
	(0.031)	(0.030)	(0.023)	(0.031)
X = Unstated × D	0.056	-0.068*	0.019	-0.044
	(0.032)	(0.030)	(0.023)	(0.031)
state_owned	0.108***	0.012***	-0.054***	-0.013***
	(0.001)	(0.001)	(0.001)	(0.001)
city_size = High × D	0.003	-0.008	-0.013***	0.012*
	(0.005)	(0.005)	(0.004)	(0.005)
city_size = Low × D	0.003	0.000	-0.011**	0.006
	(0.005)	(0.005)	(0.004)	(0.005)
city_size = Very High × D	-0.012*	-0.003	-0.014***	0.008
	(0.005)	(0.005)	(0.004)	(0.005)
city_size = Very Low × D	reference category			

* p<0.05, ** p<0.01, *** p<0.001

Table C.1: Event study estimates (continued)

	poor	young	can vote	family
month_attrib = 01/19		reference category		
month_attrib = 02/19	0.007	0.003	-0.003	0.012***
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 03/19	-0.002	0.004	-0.002	0.001
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 04/19	-0.008*	-0.005	-0.005	0.007
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 05/19	-0.001	-0.003	-0.006*	0.007*
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 06/19	-0.008*	-0.001	-0.006*	0.013***
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 07/19	-0.008*	0.018***	0.000	0.011**
	(0.004)	(0.003)	(0.003)	(0.003)
month_attrib = 08/19	-0.010**	0.058***	0.011***	0.001
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 09/19	-0.005	0.034***	-0.005*	0.005
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 10/19	-0.075*	0.093**	-0.015	0.051

* p<0.05, ** p<0.01, *** p<0.001

Table C.1: Event study estimates (continued)

	poor	young	can vote	family
	(0.031)	(0.030)	(0.023)	(0.031)
month_attrib = 11/19	-0.067*	0.091**	-0.021	0.050
	(0.031)	(0.030)	(0.023)	(0.031)
month_attrib = 12/19	-0.073*	0.091**	-0.023	0.052
	(0.031)	(0.030)	(0.023)	(0.031)
month_attrib = 01/20	-0.081**	0.094**	-0.020	0.053
	(0.031)	(0.030)	(0.023)	(0.031)
month_attrib = 02/20	-0.081*	0.096**	-0.024	0.053
	(0.031)	(0.030)	(0.023)	(0.031)
month_attrib = 03/20	-0.025***	0.018***	-0.016***	0.011**
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 04/20	-0.041***	0.023**	0.004	0.017*
	(0.008)	(0.008)	(0.006)	(0.008)
_cons	0.451***	0.279***	0.868***	0.336***
	(0.003)	(0.003)	(0.002)	(0.003)
N	525,147	525,147	525,147	525,147
adj. R^2	0.082	0.034	0.088	0.032

* p<0.05, ** p<0.01, *** p<0.001

C.2.2 Difference-in-differences estimates

Table C.2: DiD estimates

	poor	young	can vote	family
D	-0.008	-0.000	0.001	-0.005
	(0.007)	(0.006)	(0.005)	(0.007)
D × Z	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Center × D	reference category			
X = Far Right × D	0.046	-0.047	-0.008	-0.019
	(0.030)	(0.028)	(0.022)	(0.029)
X = Left × D	0.009	-0.011*	-0.003	0.000
	(0.006)	(0.005)	(0.004)	(0.006)
X = Misc × D	0.008	0.011	-0.001	0.014
	(0.009)	(0.008)	(0.006)	(0.009)
X = Right × D	-0.001	-0.003	-0.003	-0.004
	(0.006)	(0.005)	(0.004)	(0.006)
X = Unstated × D	0.005	-0.002	-0.005	0.002
	(0.008)	(0.008)	(0.006)	(0.008)
X = Center × D × Z	reference category			
X = Far Right × D × Z	-0.002	0.002	0.000	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
X = Left × D × Z	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Misc × D × Z	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
X = Right × D × Z	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)

* p<0.05, ** p<0.01, *** p<0.001

Table C.2: DiD estimates (continued)

	poor	young	can vote	family
X = Unstated × D × Z		ommitted		
state_owned	0.108***	0.010***	-0.051***	-0.012***
	(0.002)	(0.002)	(0.001)	(0.002)
city_size = High × D	0.010	-0.006	0.015***	0.013*
	(0.006)	(0.005)	(0.004)	(0.006)
city_size = Low × D	0.009	-0.004	0.002	0.004
	(0.006)	(0.005)	(0.004)	(0.006)
city_size = Very High × D	-0.001	0.002	0.016***	0.008
	(0.006)	(0.005)	(0.004)	(0.006)
city_size = Very Low × D		reference category		
month_attrib = 01/19		reference category		
month_attrib = 02/19	0.007	0.003	-0.003	0.012***
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 03/19	-0.002	0.004	-0.002	0.001
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 04/19	-0.008*	-0.005	-0.005	0.007
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 05/19	-0.001	-0.003	-0.006*	0.007*
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 06/19	-0.009*	-0.001	-0.006*	0.013***
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 07/19	-0.009*	0.018***	0.000	0.011**
	(0.004)	(0.003)	(0.003)	(0.003)

* p<0.05, ** p<0.01, *** p<0.001

Table C.2: DiD estimates (continued)

	poor	young	can vote	family
month_attrib = 08/19	-0.010*	0.058***	0.011***	0.001
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 09/19	-0.005	0.035***	-0.005*	0.005
	(0.004)	(0.003)	(0.003)	(0.004)
month_attrib = 10/19	-0.015***	0.020***	-0.011***	0.006
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 11/19	-0.007	0.018***	-0.017***	0.005
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 12/19	-0.013**	0.017***	-0.019***	0.007
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 01/20	-0.021***	0.021***	-0.016***	0.008
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 02/20	-0.021***	0.023***	-0.019***	0.008*
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 03/20	-0.025***	0.017***	-0.016***	0.010*
	(0.004)	(0.004)	(0.003)	(0.004)
month_attrib = 04/20	-0.0413***	0.023***	-0.003	0.017*
	(0.008)	(0.008)	(0.006)	(0.008)
_cons	0.451***	0.280***	0.867***	0.336***
	(0.003)	(0.003)	(0.002)	(0.003)
N	525,147	525,147	525,147	525,147
adj. R^2	0.082	0.034	0.088	0.032

* p<0.05, ** p<0.01, *** p<0.001