The Geography of Higher Education and Spatial Inequalities*

Anaïs Fabre†

March 25, 2026

Abstract

This paper studies how the geography of the higher education market shapes spatial inequalities. Using data on the universe of college applicants and programs in France, I document that higher education options are unevenly distributed across space while students' demand is highly sensitive to geographic proximity. To quantify how these features shape regional gaps in educational attainment and spatial skill sorting, I build a dynamic model linking equilibrium sorting on the higher education market and location choices of entry-level workers. I show that students' and programs' preferences can be identified and estimated from data on choices and equilibrium outcomes. Thirteen percent of the regional variance in educational attainment is explained by the interaction of the uneven distribution of colleges and mobility frictions. The latter, however, insulate low-opportunity areas from an outflow of workers in the long-run, mitigating regional inequality. I show that low-opportunity areas could invest in the education of their local labor force while limiting out-migration by combining mobility scholarships with incentives to return.

*I am very grateful for the guidance and helpful comments from Matteo Bobba, Olivier De Groote, Thierry Magnac, and Arnaud Maurel. I also thank Francesco Agostinelli, Juan Pablo Atal, Emmanuelle Auriol, Steven Berry, Richard Blundell, Juan Camilo Castillo, Fabrice Collard, Pierre Dubois, Gilles Duranton, Tim Ederer, Hanming Fang, Ana Gazmuri, Eugenia Gonzalez-Aguado, YingHua He, Christian Hellwig, Clément Imbert, Vishal Kamat, Margaux Luflade, Nour Meddahi, Pepita Miquel-Florensa, Sophie Moinas, Kaivan Munshi, Augusto Ospital, François Poinas, Mounu Prem, Andrew Shephard, Augustin Tapsoba, and Petra Todd for their comments. I also thank workshop and seminar participants at the Toulouse School of Economics, the University of Pennsylvania, Northwestern University, the California Institute of Technology, UQAM, Tilburg University, University of Bonn, LMU Munich, the Institute for Fiscal Studies, Royal Holloway, Queen Mary University, University of Wisconsin-Madison, Arizona State University, University of Maryland, Washington University in Saint Louis, CMU Tepper School of Business, IMF, INSEAD, the NBER Summer Institute (Urban Economics), and the CESifo conference on the Economics of Education. Finally, I thank Lionel Bonneviale, Pierre Boulet and Guirane Ndao at the Ministry of Higher Education and Research for their support regarding data access. I acknowledge funding from the French National Research Agency (ANR) under the grant MATCHINEQ - ANR-22-CE26-0005-01.

†LMU Munich, anais.fabre@econ.lmu.de.

1 Introduction

While human capital is a fundamental economic input, individuals' educational attainment varies widely across birthplaces. The gap in college graduation rates between commuting zones in the United States reaches up to 50 percentage points (Chetty et al., 2026). These disparities may reveal the existence of local barriers to investment in human capital, which could in turn hinder growth (Hsieh et al., 2019). Spatial differences in educational attainment may also drive the observed concentration of skills across local labor markets, which impacts regional inequalities and efficiency (Diamond and Gaubert, 2022). Understanding the sources and welfare consequences of these spatial gaps is thus key.

Inequalities in economic opportunities across birthplaces and earnings differences across regions have been studied separately. While the causal effect of childhood location on individuals' lifetime outcomes has been established (Chetty and Hendren, 2018; Chyn and Katz, 2021), its underlying mechanisms remain unclear. Furthermore, its potential link to regional economic disparities has received little attention, as spatial skill sorting is typically understood as the result of workers' migration decisions (Diamond, 2016). This paper bridges these two strands of the literature by identifying a key driver of inequalities across both birthplaces and local labor markets: the geography of the higher education market. Using rich administrative data, quasi-experimental policy variation, and a dynamic structural model, it shows how mobility frictions interact with the within-country distribution of higher education programs to generate spatial gaps in students' welfare and educational attainment, which in turn contribute to spatial skill sorting and regional inequalities.

I investigate this link in four steps, providing several empirical and methodological contributions. First, I use data on the universe of college applicants and programs in France to document that (i) higher education options are highly unevenly distributed across space and (ii) students' demand for programs is elastic to geographic proximity. Second, I leverage local policy variations and survey data to study the impact of mitigating mobility frictions within the higher education market on students' sorting and on the spatial distribution of the labor supply of entry-level workers. Third, I build a dynamic model that combines a two-sided matching model of the higher education market with a model of location choice at labor market entry. I show how students' and programs' preferences can be identified from data on choices and equilibrium outcomes. Fourth, I take the model to the data to quan-

tify how mobility frictions—both pecuniary and non-pecuniary costs of studying away from home—interact with the spatial distribution of programs to generate spatial inequalities. I find that 13% of the regional variance in educational attainment is explained by the interaction between mobility frictions and the uneven distribution of colleges. This interaction also shapes the long-run spatial distribution of labor supply: mobility frictions within the higher education market reduce the share of high-skilled workers in low-opportunity areas but also insulate these locations from an outflow of workers, thereby mitigating regional divergence. Low-opportunity areas could invest in the education of their local labor force while limiting out-migration by combining mobility scholarships with incentives to return at labor market entry.

Specifically, I begin by using rich administrative data to document two facts linking the geography of the higher education market, individuals' educational choices, and spatial inequalities. First, I show that higher education options are unevenly distributed across space. Of the ninety-six provinces (*départements*) in metropolitan France, fifty account for less than 5% of the total supply of seats in academic programs, typically leading to a bachelor's degree. This does not reflect the spatial distribution of the population, as these areas concentrate 25% of students in secondary education. Second, I provide evidence suggesting that students' demand for higher education programs is highly sensitive to distance. More than 50% of students enroll in a program located less than 25 kilometers from their home location. This reflects students' choices rather than colleges' admission decisions, as the median offer they receive is located twenty kilometers farther away than the one they accept. By estimating a gravity model, I find that the elasticity of students' enrollment flows to distance is -1.9.

Overall, consistent with these two facts, I find that students' probability of enrolling in an academic program increases with their proximity to such institutions. This may in turn affect local labor markets, as I document a positive relationship between the share of a region's labor force with at least a three-year degree and the local availability of programs delivering such degrees.

I then leverage local policy changes to explore the direct and distributional effects of mitigating mobility frictions. I first evaluate a policy offering moving subsidies to low-income students conditional on enrolling out of their home region. I find that it led students to apply to programs farther away from home. To document the distributional effects of removing barriers to mobility, I then zoom in on the Greater Paris area and exploit a change

in geographic priorities determining students' university admission. Removing the priority granted to students from Paris to enroll in local universities led to a 50% increase in the share of high-ability students from peripheral locations enrolled in programs in Paris. I find that these students crowded out lower-ability applicants from Paris, suggesting that spatial mobility within the higher education market may have substantial effects on equilibrium sorting. Survey data on education and work migration flows further suggest that, if such reallocation occurred at the scale of the country, it could affect the spatial distribution of labor supply across regions. Indeed, more than half of students who graduate outside their home location do not return there at labor market entry, while 17% enter the labor market in their graduation location.

To quantify the impact of mobility frictions within the higher education market on spatial inequalities, I then build a dynamic model encompassing these mechanisms. It flexibly accounts for differences in preferences and admission chances across students and captures the long-run distributional effects of students' spatial mobility. The framework combines a two-sided matching model of the higher education market, in which students are forward-looking, with a model of optimal location choices at labor market entry. In the first period, students and programs meet on a centralized platform and express their preferences over one another. Each student's payoff from enrolling in a given program depends on its associated flow and continuation values, as well as on mobility frictions, which capture both pecuniary and non-pecuniary costs of studying away from home. I assume that the resulting equilibrium is stable with respect to students' lifetime utility and programs' priorities, and that programs rank applicants truthfully. Graduation outcomes are flexibly modeled as a function of student and program characteristics, allowing students to switch degrees and location during their studies. The second stage takes place upon graduation, when individuals choose where to enter the labor market, taking into account migration costs, home bias, and location-specific wages and amenities.

I show how to identify the primitives of this model without imposing parametric assumptions on the deterministic part of agents' payoffs. I first identify preferences over local labor markets using data on individuals' home, study, and work locations. I then show how data on programs' rankings and realized matches can be combined to identify programs' priorities and students' preferences in a two-sided matching model. I propose a tractable estimation procedure.

Parameter estimates show that students' demand is highly elastic to distance. The average student would be willing to forgo 741 euros in future monthly entry-level wages, corresponding to 45% of the average entry-level wage, to enroll in a program located 100 kilometers closer to her home location. Mobility frictions, rather than programs' characteristics, are the main component of the value associated with a program for students from higher education deserts. Regarding local labor market choices, I find that individuals display a strong preference for their home location. Yet, for individuals studying away from home, the costs of moving from their graduation location outweigh the preference to return, on average. Overall, model predictions closely match both enrollment and work flows between education hubs and deserts.

I use the estimated model to quantify how mobility frictions interact with the geography of the higher education market to generate spatial inequalities. Holding fixed the distribution of programs, I simulate individuals' education and labor market choices in a counterfactual scenario in which mobility frictions are eliminated. This breaks the link between students' demand and the local supply they face. On the higher education market, predicted sorting depends both on how students shift their demand and on how programs select whom to admit given the new pool of applicants. The model also enables me to simulate counterfactual location choices at entry on the labor market. Overall, I use the model to quantify changes in educational attainment and welfare between students who grew up in different locations, as well as in skills and economic activity across regions.

Mobility frictions within the higher education market account for 50% of the gap in bachelor's degree attainment between higher education hubs and deserts, and for about 13% of the variance in educational attainment across provinces. Eliminating these frictions would also reduce welfare inequalities across students from hubs and deserts. At the same time, mobility frictions shape the long-run spatial distribution of labor supply. They reduce the share of high-skilled workers in low-opportunity areas, thereby contributing to spatial skill sorting, but they also insulate these areas from an outflow of workers. When they are eliminated, the number of individuals from deserts entering the labor market in hubs rises by 9%, increasing inequalities across local labor markets.

I perform two additional quantification exercises. I show that modeling programs' admission decisions is key to capturing these effects. I also eliminate both mobility frictions within the higher education market and moving costs between study and home locations. When

individuals can return home at no cost, inequalities across both students and local labor markets decline. Yet, students from low-opportunity areas accumulate less human capital, because they anticipate entering home labor markets where returns to education are lower.

Finally, I explore whether mobility scholarships tied to incentives to return home after graduation could help low-opportunity areas invest in their local human capital while limiting the long-run outflow of workers. For a cost below 3% of entry-level workers' total wages, this policy would increase the share of students from low-opportunity areas graduating with at least a bachelor's degree by 3.29% and the number of high-skilled workers entering the labor market there by 16%.

Overall, this paper contributes to several strands of the literature. First, it relates to work on the causal effects of places on individuals' lifetime outcomes (Chetty et al., 2014; Chetty et al., 2026) by quantifying the impact of differences in the local supply of higher education options. The existence of educational deserts has been documented in the United States (Hillman, 2016). Studies in Belgium (Kelchtermans and Verboven, 2010), England (Gibbons and Vignoles, 2012), and the United States (Fu et al., 2022; Ishimaru, 2025) show that students' demand for higher education programs is sensitive to proximity, affecting their welfare and economic outcomes. I contribute to this literature by studying these mechanisms in a two-sided equilibrium model of the higher education market, which allows me to account for how eliminating mobility frictions reshapes sorting under capacity constraints and programs' admission decisions, and by linking this market to graduation outcomes and labor market location choices to quantify the long-run distributional effects of students' imperfect spatial mobility.

This framework relates to the literature studying equilibrium sorting in the higher education market (Epple et al., 2006; Fu, 2014; Kapor, 2025). Relying on the centralized nature of the French higher education market, I build on work estimating students' preferences in centralized school choice and college admissions when schools' or programs' priorities are observed (Agarwal and Somaini, 2018; Fack et al., 2019; Calsamiglia et al., 2020; Kapor et al., 2024). Using data on programs' rankings of applicants, I show how preferences on both sides of the market can be identified without imposing exclusion restrictions (Agarwal and Somaini, 2025; He et al., 2024) or taking a stand on students' application strategies. I also contribute to this literature by linking the equilibrium model of the higher education market to a location choice model of entry-level workers (Kennan and Walker, 2011). This explicitly

accounts for the fact that the human capital investment decisions of forward-looking students may depend on local returns to education in the presence of home bias. It also allows me to quantify how mobility frictions affect not only students' outcomes, but also the spatial distribution of skills and inequalities across local labor markets.

By accounting for these dynamics, this paper also contributes to the literature on regional inequality and spatial skill sorting. Spatial sorting of skills is a key driver of regional disparities (Combes et al., 2008; Dauth et al., 2022; Card et al., 2025), and this literature has primarily emphasized workers' migration decisions as a source of divergence across places (Diamond, 2016). I show that part of this sorting is shaped upstream, through "sorting at birth" (Combes et al., 2012): unequal access to higher education contributes to persistent differences in local skill composition and shapes inequalities across local labor markets. The framework also allows me to evaluate mobility scholarships for local students, tied to return incentives at labor market entry, as a policy tool through which low-opportunity areas can invest in local human capital.

2 Context and Data

This section describes the French higher education market and the college admission procedure and introduces the data sources used throughout the paper.

2.1 The French higher education market

Higher education in France is centrally regulated by the Ministry of Higher Education and Research and organized into thirty catchment areas, which implement national policy locally.

Students can access higher education provided they receive a passing grade on the national end-of-high-school exam.[1] There are six main types of programs available. First, universities offer three-year bachelor's programs, after which students can continue for two additional years to obtain a master's degree. Second, students can apply to two-year elite preparatory programs, which are highly selective and prepare students for competitive exams for admission to prestigious higher education institutions. Third, engineering and business schools offer selective programs that also typically lead to a master's degree.

[1] Appendix B provides more details about the high school system in France.

Three additional program types provide shorter, more vocational training. First, two-year technology-oriented programs train mid-level technical professionals. Second, two-year vocational programs place greater emphasis on apprenticeships and internships. Third, social work and nursing schools train workers in two to three years.

2.2 College admission procedure

Since 2009, higher education admissions have been centralized on an online platform. Each year, students in their final year of high school register on the platform. An online search engine is available: students can search for programs by major and location, and a map tool displays institutions offering a given major. For each program, the platform provides a general description, broad admission criteria, the number of seats, and the number of applicants in the previous year.

Students then submit applications on the platform. Programs are defined by an institution-type-major combination. For example, applying to economics and mathematics at the Sorbonne University requires two separate applications. In principle, the number of applications is limited to ten. However, the effective number of applications allowed is substantially larger in practice as, in many cases, several applications are processed as a single one. For instance, listing up to ten institutions offering a two-year elite program in physics counts as one application. In the selected sample, which I define below, the average number of applications is 10, and 25% of students submit 13 applications or more (Table A.3).

After applications are submitted, each program ranks its applicants based on information submitted by students and high schools. In particular, programs observe students' high school grades and may require additional materials such as a resume or cover letter. Programs do not have to disclose how the different characteristics they observe are weighted to generate their rankings: programs' priorities are therefore unobserved. Interviews of members of admission committees, however, reveal that programs typically first choose the weights to assign to each observed characteristic.[2] These weights are then fed into an algorithm generating the ranking of applicants. The use of cover letters and resumes seems less common.

Program rankings are then adjusted to take into account some equity-based admission criteria chosen by policymakers ex ante. In particular, each program faces a target minimum

[2]See 'Parcoursup : enquête dans les coulisses des commissions d'examen des vœux', Le Monde, Séverin Graveleau, Soazig Le Nevé et Eric Nunès. Published on May 23^{rd}, 2023. Last accessed: August 18^{th}, 2023.

share of scholarship-eligible (low-income) students. University programs are also subject to place-based quotas that set a maximum share for admitted students from outside the catchment area.

Once rankings are finalized, the admission stage begins. Programs issue offers starting from the top of their ranked lists, up to capacity.[3] Students may receive multiple offers: they can accept one and decline the others, in which case offers are extended to the next applicants in the ranking. Students can also hold an offer while waiting for a preferred program. Throughout the process, students observe their waitlist rank and the rank of the last admitted student in the previous year, which helps them predict whether they will receive an offer.

2.3 Data

I assemble a dataset covering the universe of applicants and programs on the French college admission platform and tracking students until they exit higher education. I complement these administrative records with a survey following individuals at entry into the labor market. This section describes these different data sources.

College Admission System (Parcoursup): First, I use administrative data from the college admission platform in France, which cover the universe of applicants each year. The data include a rich set of student characteristics, such as gender, parents' socio-economic status, scholarship status, and test scores. Importantly, I also observe each student's residential location at the ZIP code level—there are 37,901 ZIP codes in France. Table A.1 reports summary statistics for the main sample used throughout the paper. I focus on high school applicants living in metropolitan France, yielding a sample of 507,605 students in 2019.

The data also cover the universe of programs available on the platform. For each program, I observe its type, major, number of seats, and GPS coordinates, allowing me to compute the distance between each student and each program. I restrict attention to the 10,936 in-person programs located in metropolitan France. Table A.2 reports summary statistics for this sample.

Finally, this dataset records all activity on the platform. It includes all submitted applica-

[3] To make the process run more smoothly, the platform allows programs to send more offers than available seats.

tions and each program's ranking of applicants, and it tracks offers, students' acceptance and rejection decisions, and final enrollment outcomes. Tables A.3 and A.4 describe application, offer, and enrollment patterns.

Higher Education Census (SISE): Second, I use administrative records that follow each student enrolled in the French higher education system until they exit. I merge these records with the platform data, which allows me to track students from entry to graduation or dropout. I observe whether students switch programs, the highest degree they obtain, and their province of graduation.

Location Choices and Earnings Survey (Enquêtes Génération): I complement these datasets with a survey that follows individuals after they leave the education system. I restrict attention to high school graduates. The survey records individuals' home, graduation, and work locations—allowing me to observe migration decisions—and reports wages within three years after graduation. Table A.5 summarizes work and study flows and earnings.

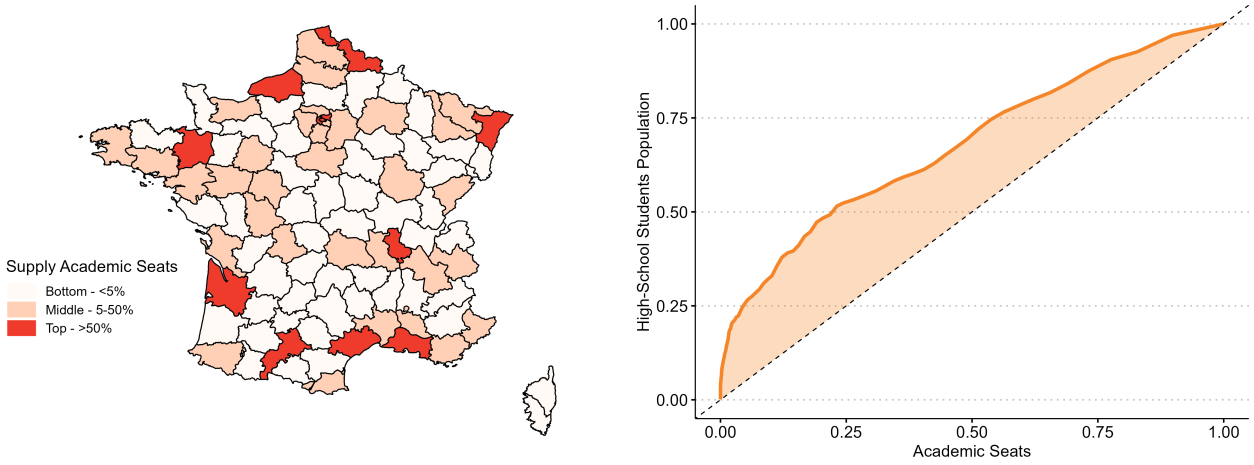
3 Descriptive Evidence and Policy Evaluation

In this section, I first present descriptive evidence linking the geography of higher education to regional inequalities in educational attainment and spatial skill sorting. I then use local policies that mitigate mobility frictions to study their effects on students' demand and document their distributional consequences.

3.1 Spatial distribution of colleges and inequalities

I first document that the spatial distribution of academic programs is uneven and does not mirror that of the student population. I then show that students' demand is highly sensitive to geographic proximity, consistent with the presence of monetary and non-monetary mobility costs. Together, these patterns coincide with observed spatial gaps in enrollment choices and skill sorting.

Figure 1: Geographic distribution of academic seat supply and high school student population



a) Spatial distribution of academic seats **b) Student-seat mismatch**

NOTES. Panel a) maps the distribution of academic seats in metropolitan France. Provinces are ranked according to their position in the distribution of academic seats. Provinces in white lie at the bottom of the distribution and jointly account for less than 5% of total academic seats. Provinces in red lie at the top of the distribution; together they account for half of total academic seats. Panel b) relates the geographic distributions of academic seats and high school students. The x-axis reports the cumulative share of academic seats across provinces and the y-axis the corresponding cumulative share of the high school student population, where provinces are ranked by their share of academic seats. For example, the provinces that account for 25% of academic seats are home to 50% of high school students.

3.1.1 Fact 1: Higher education programs are unevenly distributed across space

The spatial distribution of higher education programs is highly uneven. I focus on academic programs—universities and two-year elite programs—which are the main pathways to a bachelor's degree. Panel a) of Figure 1 maps the spatial distribution of first-year academic seats at the province level. Of the 96 provinces in metropolitan France, ten do not offer a single academic seat.[4] Overall, half of the provinces jointly account for less than 5% of the total supply of academic seats. I refer to these areas as higher education deserts. In contrast, 50% of the total supply is concentrated in only 12 provinces, pointing to a highly uneven spatial distribution of academic seats.

This geography could, however, merely reflect the spatial distribution of students. Panel b) of Figure 1 shows that it does not: provinces accounting for less than 5% of the total supply of academic seats are home to 25% of high school students. There is thus a substantial mismatch between the spatial distributions of academic seats and students.[5]

[4]In total, 17 provinces offer no university programs, while 18 provinces do not have any two-year elite programs.

[5]This mismatch contrasts with the number of seats in two-year technological and vocational programs,

Table 1: Number of programs near students' home location

	Quantile			
	10^{th}	25^{th}	50^{th}	75^{th}
	(1)	(2)	(3)	(4)
Panel A: 10 kilometer-radius				
Three-Year University Program	0	0	0	29
Two-Year Elite Program	0	0	0	9
Engineering & Business Schools	0	0	0	5
Two-Year Technical Program	0	0	1	10
Two-Year Vocational Program	0	2	16	63
Nursing Schools	0	0	1	8
Panel B: 40 kilometer-radius				
Three-Year University Program	0	6	35	115
Two-Year Elite Program	0	3	11	34
Engineering & Business Schools	0	2	7	25
Two-Year Technical Program	2	6	13	29
Two-Year Vocational Program	30	53	92	267
Nursing Schools	1	4	9	24

NOTES. Number of programs, by program type, located within a given radius of students' home location. Panel A considers a ten-kilometer radius, while Panel B presents results for a forty-kilometer radius. I consider 40 kilometers as the threshold defining students' commuting distance, to the extent that 96% of daily commuting trips in France are 40 kilometers or less. This table shows that 10% of the students have no three-year university programs located within a forty-kilometer radius around their home location.

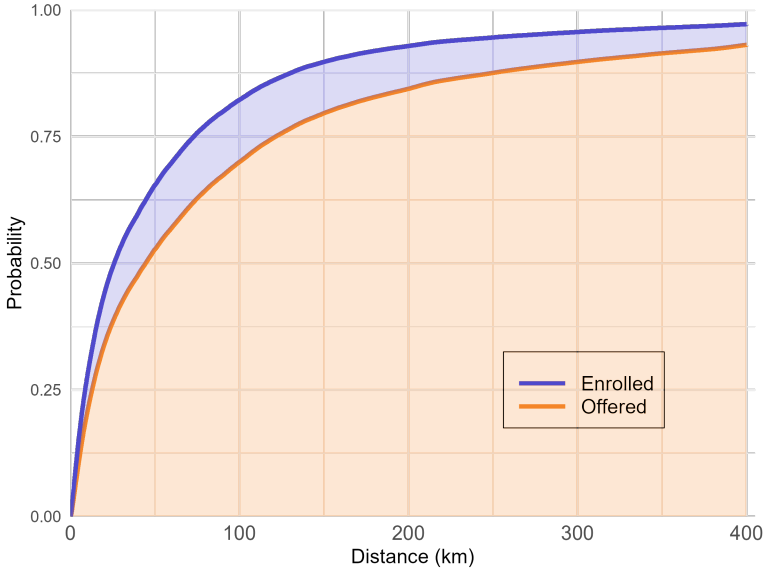
As a result, many students face a limited local supply of academic programs. Table 1 reports the distribution of the number of programs, by type, within 10 and 40 kilometers of students' residence. I interpret 40 kilometers as a commuting threshold, since 96% of daily commuting trips in France are 40 kilometers or less.[6] Half of students do not have a single university program within ten kilometers of home. Expanding the radius to daily commuting distance, more than 10% of applicants still have no university nearby, and a quarter have access to at most six university programs within 40 kilometers. The spatial distribution of two-year elite programs and engineering and business schools is similar to that of university programs.[7] Because each program corresponds to a specific major, these patterns imply that

which is considerably more evenly distributed across space (Figure A.2) and thus better matches the spatial distribution of high school students (Figure A.3).

[6]According to the same study, in France, the average individual only makes 6.3 trips of more than 80 kilometers from their home each year. Source: SDES, Insee – Enquête Mobilité des Personnes 2018-2019. Data: 'La mobilité locale et longue distance des Français - Enquête nationale sur la mobilité des personnes en 2019', accessed at statistiques.developpement-durable.gouv.fr. Last accessed: March 3^{rd}, 2026.

[7]Figure A.1 provides a map of the mean number of academic programs located within commuting distance of students' home location at the province level. The mean number of academic programs located within a

Figure 2: Cumulative Distribution - Distance of Enrollment & Offers



NOTES. This figure shows the cumulative distribution function of the distance (in kilometers) between students and the program in which they enroll (purple) and between students and the programs for which they receive offers (orange).

such students have a very limited set of majors available locally, as illustrated by Table A.6. For example, 10% of students must travel at least 93 kilometers to reach a medicine program.

3.1.2 Fact 2: Students' demand is elastic to geographic proximity

The concentration of academic programs across space would have no consequences for students' college choices absent monetary or non-monetary mobility costs. Such costs could include travel expenses, housing costs, and the emotional toll of being far from family and friends. Enrollment patterns suggest that such costs are present and play an important role in students' higher education choices. Figure 2 shows that 25% of students enroll less than 8.5 kilometers from home, and the median enrollment distance is 25 kilometers. This pattern is not driven by students receiving offers only from programs located very close to home. On the contrary, the distribution of the distance to received offers (orange line) first-order stochastically dominates that of the distance to accepted offers (purple line). The median offer is located 20 kilometers farther away than the accepted offer. This gap increases to 50 kilometers at the third quartile of the distribution.

This evidence suggests that students favor programs closer to home, consistent with mo-

40-kilometer radius around students' home location is less than ten in 24 provinces.

bility frictions. To measure the elasticity of enrollment flows with respect to distance, I borrow from the migration and trade literature and estimate a gravity model. I regress the flow of students from each origin province enrolling in programs located in each destination province on the distance between the two provinces' centroids, including origin and destination fixed effects.[8] I estimate an elasticity of enrollment with respect to distance of -1.9 (Table A.7), indicating that college choices are highly sensitive to distance from home.

3.1.3 Implications: Spatial inequalities

Overall, the interaction of mobility frictions with the uneven distribution of higher education programs is likely to generate important inequalities between students from higher education hubs and deserts. Panel a) of Figure 3 shows a positive relationship between the share of students from a province enrolled in an academic program and the local availability of such programs. In addition, because individuals often work in their home region—given moving costs and home bias (Kennan and Walker, 2011)—this relationship can contribute to skill sorting across local labor markets. Panel b) of Figure 3 shows a similar positive relationship between the share of a province's labor force with at least a bachelor's degree and the local availability of academic programs.

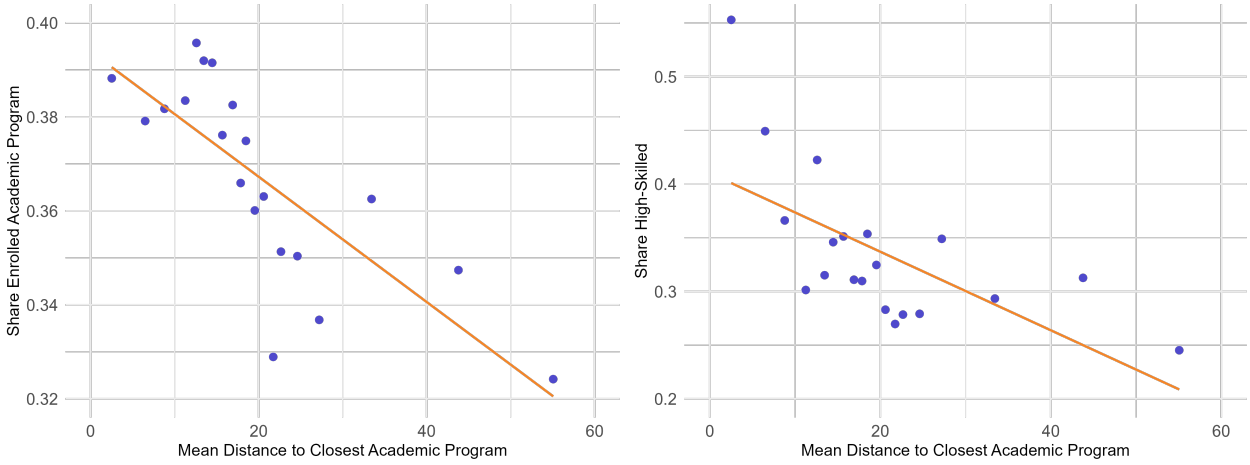
These correlations could reflect factors other than mobility frictions, such as differences in students' pre-college human capital, preferences, or returns to education across provinces. In Section 4, I build a model that accounts for these factors, allowing me to quantify the contribution of mobility frictions, interacting with the geography of higher education, to spatial inequalities in educational attainment and skill sorting.

3.2 Mitigating mobility frictions: demand responses and distributional consequences

The patterns above suggest that mobility frictions within higher education contribute to spatial inequalities. This section uses two local policy changes, along with survey data, to document how relaxing mobility constraints reshapes students' demand and affects distributional outcomes, both within higher education and in subsequent labor market choices—key margins that the model in Section 4 is designed to capture.

[8] I estimate this model using a Pseudo-Poisson Maximum-Likelihood estimator. See Silva and Tenreyro (2006) for a discussion of this approach in gravity models.

Figure 3: Higher Education Market Geography and Spatial Inequalities



a) Spatial gaps in academic enrollment **b) Spatial skill sorting**

NOTES. Panel a) is a binned scatter plot relating the share of students from a province enrolled in an academic program to the mean distance (in kilometers) to the closest academic program for the average student in the province. Panel a) of Figure A.4 shows that a similar relationship holds for specific majors, such as medicine. Panel b) is a binned scatter plot relating the share of a province's labor force holding at least a three-year degree (share high-skilled) to the same distance measure.

3.2.1 Mobility scholarship and changes in students' demand

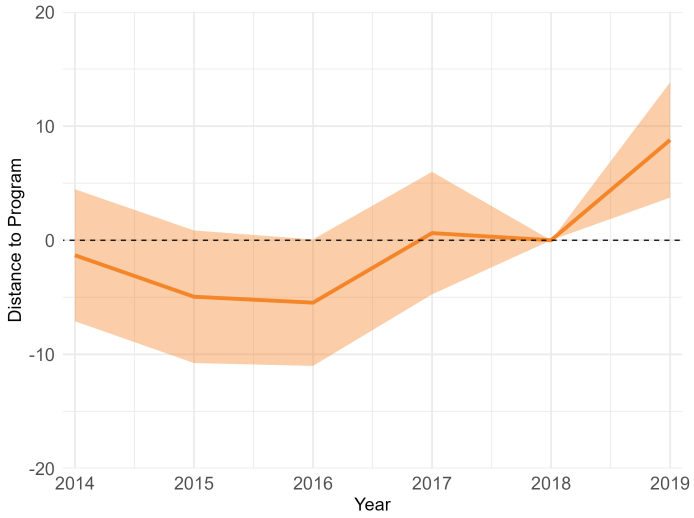
Studying away from home entails direct expenses—such as travel and housing costs—which are likely to be an important component of mobility frictions. Could moving subsidies lead students to apply farther away from home? To answer this question, I evaluate a mobility scholarship introduced in 2019 that guarantees 500 euros to low-income students conditional on enrolling in a program located outside their catchment area. This amount covers a fraction of the costs associated with such a move, corresponding to about 9% of the annual budget of the average student.[9]

The scholarship is available only to recipients of a means-based scholarship in high school, as a way to target low-income students. Using application data from 2014 to 2019, I can thus compare changes in application distance for eligible and non-eligible students before and after the policy implementation. Because other reforms to the admission system were implemented over this period, potentially affecting students with different eligibility statuses differently, I use an additional comparison group. In particular, I compare eligible (E) and non-eligible (NE) students living in provinces with low (L) and high (H) supply of higher education programs.[10] Students in high-supply areas face many options close to home and

[9]Source: Observatoire National de la Vie Etudiante.

[10]I define as high-supply areas the ten provinces at the top of the supply distribution. Low-supply areas

Figure 4: Effect of the Spatial Mobility Scholarship on Application Distance



NOTES. This figure presents parameter estimates of $\beta_{2018+\tau}$ from Equation (1). It shows the effect of the policy offering moving subsidies, implemented in 2019, on the mean application distance, in kilometers. 95% confidence intervals are shown, computed using robust standard errors. These estimates are displayed in Table A.8.

are therefore unlikely to be affected by the reform. I estimate the following model:

$$d_{i,e,s,t} = \alpha_{s,t} + \alpha_{e,t} + \alpha_{s,e} + \sum_{\tau=-3,\tau\neq 0}^{1} \beta_{2018+\tau} D_{e,s,2018+\tau} + \epsilon_{i,s,e,t}, \tag{1}$$

where $d_{i,e,s,t}$ is the distance, in kilometers, between a student and the program to which she submits application i in year t, while being of eligibility status $e \in \{E, NE\}$ and living in a province with supply level $s \in \{H, L\}$. $\alpha_{s,t}$ are supply group-year fixed effects, $\alpha_{e,t}$ are eligibility-year fixed effects, and $\alpha_{s,e}$ are supply-by-eligibility fixed effects. $D_{e,s,2018+\tau}$ is an indicator for eligible students in low-supply areas in year $2018 + \tau$, for $\tau \in \{-3, 1\} \setminus \{0\}$. $\beta_{2018+\tau}$ is thus a triple-difference estimator, estimating the relative treatment effect for eligible students in the low- versus high-supply areas.

The identifying assumption is that, absent the scholarship, the relative evolution of outcomes for eligible and non-eligible students would have been the same in high- and low-supply areas (Olden and Møen, 2022). Details are provided in Appendix C.1. I assess this common-trends assumption using placebo tests in the pre-policy years.

Figure 4 and Table A.8 report estimates of $\beta_{2018+\tau}$. I find that the moving subsidy allowed eligible students from low-supply areas to apply to programs located farther away from their home location: the mean application distance increased by 9 kilometers. I find no evidence

are defined as the bottom five provinces of the supply distribution.

of differential pre-trends prior to the reform. The effects are driven by an increase in out-of-catchment applications (Table A.8). However, this shift does not translate into a significant increase in out-of-catchment offers, such that enrollment patterns are not affected.

These results suggest that demand responds to reductions in mobility costs. While lower mobility costs may allow students to apply to programs they prefer, the policy evaluation above does not allow me to quantify this effect. Furthermore, applications may fail to translate into admissions, as the market is two-sided: programs choose which applicants to admit.

3.2.2 Distributional effects of alleviating barriers to mobility

Alleviating mobility frictions alters students' demand. In a two-sided market with capacity constraints, such changes can entail substantial distributional effects. In this section, I investigate these effects. First, I use a local policy change to show that eliminating barriers to students' mobility may generate an outflow of high-ability students toward higher education hubs in the short run. I then use survey data to illustrate how, if occurring at the scale of the country, such an outflow could affect the distribution of labor supply across regions in the long run.

Distributional Effects — Higher Education Market Application responses by marginal students need not translate into admissions, as illustrated above. A broad reduction in moving frictions, as considered in the counterfactual analysis, may shift application patterns more widely and could therefore affect admissions and enrollment in equilibrium. To illustrate these reallocation effects, I next examine a reform in the Greater Paris area that removed institutional barriers to mobility and show how lower geographic barriers can reshape sorting in the higher education market when capacity is fixed.

The Greater Paris area is composed of three catchment areas: Paris, Versailles, and Créteil.[11] Until 2018, to manage frequent oversubscription of university programs in Paris, geographic priorities—akin to residency-based constraints used by some public universities in the United States—reduced the admission chances of students living outside Paris.[12] Table A.10 shows how these priorities affected the final ranking of students: on average, a student

[11] Figure A.5 presents a map of the area and Table A.9 shows summary statistics of programs located in the region.

[12] In the United States, in 2017, the University of California for example implemented an 18% cap on out-of-state enrollment. See 'A bold plan for UC: Cut share of out-of-state students by half amid huge California demand', Los Angeles Times. Published on May 25^{th}, 2021. Last accessed: October 10^{th}, 2023.

from Versailles applying to a university program in Paris was down-ranked by 483 positions, while the average rank of the last-admitted student was 1,000. In 2019, to level the playing field for students in the region, these geographic priorities were removed.[13]

I evaluate how this reform affected student sorting within the higher education market. In particular, I estimate the following model:

$$y_{j,\ell,t} = \alpha_j + \sum_{\ell} \beta_\ell D_{\ell,t} + \epsilon_{j,\ell,t}, \tag{2}$$

where $y_{j,\ell,t} = \frac{n^H_{j,\ell,t}}{n_{j,\ell,t}}$ is the share of high-ability students admitted by program j located in $\ell \in \{\text{Paris}, \text{Versailles}, \text{Créteil}\}$ in year t, with $n_{j,\ell,t}$ ($n^H_{j,\ell,t}$, respectively) the number of admitted students (high-ability admitted students, respectively). High-ability students are those graduating from high school with honors or with the highest honors. I also consider the share of high-ability applicants to program j to investigate direct changes in demand. α_j are program fixed effects. $D_{\ell,t}$ is a dummy variable equal to one for university programs in location ℓ in $t = 2019$, i.e., after the reform. I allow the reform to have different effects by program location.

Table 2 presents parameter estimates from Equation (2). I find that this reform generated an outflow of students from peripheral regions toward programs located in Paris. In particular, the share of high-ability students from Versailles enrolling in programs in Paris increased by 50% after the reform (Column (1)). In contrast, the share of high-ability students from Versailles enrolling in their local universities decreased by 25%. These students thus enrolled in universities located in Paris instead of their local institutions.[14]

As the number of seats is limited, this inflow of admitted students in Paris had further consequences for overall sorting. Students from peripheral areas crowded out low-ability students from Paris from their local universities: their share decreased by 31%, from a baseline of 22% when they were receiving priority (Column (3)). In turn, the share of low-ability students from Paris enrolled in universities located in the peripheral areas increased.

Table A.11 shows that these enrollment effects were mostly driven by programs being able to express their preferences without being subject to geographic priorities, rather than

[13] See also Bechichi et al. (2021) for descriptive evidence on the reform.

[14] I do not observe comparable movements for students from Créteil. This is consistent with geographic differences in achievement: the share of students classified as high-ability is 6.5 percentage points larger in Versailles than in Créteil.

Table 2: Removing Barriers to Mobility: Sorting Effects

	Share High-Ability		Share Low-Ability
	from Versailles	from Créteil	from Paris
	(1)	(2)	(3)
β_{Paris}	0.025	0.010	-0.067
	(0.007)	(0.007)	(0.015)
$\beta_{\text{Versailles}}$	-0.029	-0.000	0.022
	(0.010)	(0.003)	(0.005)
$\beta_{\text{Créteil}}$	0.003	-0.006	0.022
	(0.002)	(0.004)	(0.006)
Baseline Paris	0.049	0.048	0.213
Baseline Versailles	0.114	0.013	0.029
Baseline Créteil	0.009	0.054	0.046
Program FE	Yes	Yes	Yes
Observations	680	680	680

NOTES. This table presents estimates from Equation (2). I study the effect of removing barriers to mobility on the share of high-ability students from Versailles (Column (1)), the share of high-ability students from Créteil (Column (2)), and the share of low-ability students from Paris (Column (3)) enrolled in programs located in Paris, Versailles, and Créteil. Standard errors clustered at the program level are shown in parentheses.

by changes in demand spurred by the reform. The increase in applications from high-ability students from Versailles to Paris is much smaller than the corresponding enrollment increase. This indicates that programs' admission decisions are key in shaping the equilibrium of the higher education market.

These results are robust to controlling for other factors that may have affected the higher education market over the same period (Tables A.12-A.13). In particular, I use non-university programs, which were never subject to geographic restrictions, as a control group (Appendix C.2). Overall, these results highlight that quantifying the impact of mobility frictions requires accounting for their potential effects on student sorting within the higher education market.

Distributional Effects — Local Labor Markets If occurring at the scale of the country, such a reallocation of students within the higher education market may have consequences for the spatial distribution of workers in the long-run. Using data on individuals' migration flows between home, study, and work locations, I find that many students who graduate outside their location of origin also enter the labor market elsewhere. Table A.5 shows that among the 54% of students graduating outside their home location, more than half do not return at labor market entry. In contrast, only 21% of individuals who study in their home

province enter the labor market in a different location.

Conditional on studying outside their home location, 17% of individuals enter the labor market in their study location. Location choices thus seem to display stickiness, with individuals having a high likelihood of contributing to the labor force of their study location. Higher education hubs benefit from this mechanism, as illustrated by Figure A.6.

The reduced-form evidence above shows that mitigating barriers to mobility can shift applications, but whether these shifts affect sorting within the higher education market depends on programs' admission decisions. It also suggests that the resulting reallocation of students could have persistent consequences for the spatial distribution of workers, since many students who study away do not return at labor market entry. In the next section, I build a model incorporating these channels to quantify the impact of mobility frictions on spatial inequalities.

4 Model

In this section, I build a dynamic model that combines an equilibrium model of the higher education market with forward-looking students and a model of location choice at labor market entry. I consider a two-stage framework. In the first period, a set of high school students $\mathcal{I}$ and a set of higher education programs $\mathcal{J}$ meet on a centralized platform, where they can express their preferences over one another. A match between students and programs is formed on the platform. In the second period, individuals exit the education system with a certain degree and choose the location in which to enter the labor market from a set of locations $\mathcal{L}$. In what follows, I first describe individuals' preferences and programs' priorities. I then discuss how they map into choices and outcomes by detailing the equilibrium assumptions.

4.1 Preferences

Individuals' preferences On the centralized platform, students observe program j's characteristics, z_j, for all $j \in \mathcal{J}$. Student i's payoff from enrolling in program j is:

$$\mathcal{U}_{ij} = u(x_i, z_j) - c(x_i, d_{ij}) + \sigma\epsilon_{ij} + \bar{U}(x_i, h_i, z_j). \tag{3}$$

This payoff has two components. First, students derive a flow utility from enrolling in program j. The function $u(x_i, z_j)$ captures the value of enrolling in program j as a function of student characteristics x_i and program characteristics z_j. Mobility frictions are captured by $c(x_i, d_{ij})$, which depends on student characteristics and the distance d_{ij} between student i and program j. I denote the outside option by $j = 0$ and normalize $u(x_i, z_0)$ and $c(x_i, d_{i0})$ to zero.[15] ϵ_{ij} is an idiosyncratic preference shock, independently distributed Type I Extreme Value. The parameter σ controls the scale of this shock.

Second, students are forward-looking: they take into account that their choice on the higher education market affects their outcomes in stage two, which takes place after graduation. In this stage, individuals choose a labor-market location $\ell \in \mathcal{L}$ and obtain payoff $\mathcal{W}_{i\ell}$, defined below. I capture these dynamic considerations through the continuation value $\bar{U}(x_i, h_i, z_j)$, defined as the discounted expected value of the optimal stage-two location choice for a student with characteristics (x_i, h_i) given the characteristics of program j. The variable h_i only enters the stage-two payoff.[16] We have:

$$\bar{U}(x_i, h_i, z_j) = \int \beta^{t(y_i)} \mathbb{E}\left[\max_{\ell \in \mathcal{L}} \mathcal{W}_{i\ell} \mid x_i, h_i, y_i\right] f\left(y_i \mid x_i, h_i, z_j\right) dy_i, \tag{4}$$

where y_i collects the degree obtained and the location at graduation, both influenced by first-stage enrollment decisions. The transition density $f\left(y_i \mid x_i, h_i, z_j\right)$ captures the probability of each graduation outcome y_i, given initial characteristics (x_i, h_i) and program characteristics z_j. Students thus know that, conditional on their characteristics, each program is associated with a different probability of graduating with a given degree in a given location. β is the discount factor, and the exponent $t(y_i)$ captures that individuals enter the labor market at different times depending on the length of their studies.

$\mathcal{W}_{i\ell}$ is individual i's payoff from entering the labor market in location ℓ:

$$\mathcal{W}_{i\ell} = w_{i\ell} + \omega_\ell(x_i, h_i, y_i) + \frac{1}{\theta}\nu_{i\ell}, \tag{5}$$

where $w_{i\ell}$ is the expected monthly wage of individual i in location ℓ, and $\omega_\ell(x_i, h_i, y_i)$ captures the value of location ℓ's attributes for an individual with characteristics (x_i, h_i, y_i). $\nu_{i\ell}$ is the

[15] The outside option consists in not matching to any program available on the platform. As the latter gathers the quasi-universe of higher education programs in France, the outside option can be thought of as dropping out of the higher education system.

[16] I discuss the role of this exclusion restriction in Section 5.1.

idiosyncratic shock of individual i for location ℓ, independently distributed Type I Extreme Value. The parameter θ controls the relative importance of the deterministic component of the utility relative to the shock.

In particular, I define $w_{i\ell} = \mathbb{E}[w^*_{i\ell}|x_i, h_i, y_i]$ where $w^*_{i\ell}$ is a latent variable such that:

$$w^*_{i\ell} = w_\ell(x_i, h_i, y_i) + \tilde{\epsilon}_{i\ell},$$

where $\tilde{\epsilon}_{i\ell}$ is an idiosyncratic shock with zero-mean, independently distributed across i and ℓ, that the individual does not observe prior to choosing location ℓ, as in Kennan and Walker (2011). $w_\ell(x_i, h_i, y_i)$ is a flexible function, defined in Section 5, allowing for differential returns to education across local labor markets.[17]

Together, Equation (5) and the wage specification above pin down the continuation value $\bar{U}(x_i, h_i, z_j)$. As the scale of $\bar{U}(x_i, h_i, z_j)$ is fixed due to the normalization of the wage parameter in $\mathcal{W}_{i\ell}$, I let the scale of the idiosyncratic part of the flow utility associated with program j, σ, free. Normalizing this scale to one, as is typically done, would be a strong assumption in this model as it is a non-stationary framework. In particular, the choices made in the first and second stages are very different, while normalizing the scale of ϵ_{ij} to one would be equivalent to fixing the relative scale of utility across stages one and two. I show how to identify this parameter thanks to an exclusion restriction in Section 5.1.

Programs' priorities over students On the platform, programs observe applicants' characteristics (x_i, h_i). While programs can express preferences over applicants, they also have to take into account some equity objectives imposed by the government. For example, scholarship-eligible applicants benefit from an affirmative action policy. Admission decisions are therefore based on programs' preferences after incorporating the government's equity objectives. I capture the resulting priorities through an applicant-program specific score. The final score of applicant i for program j is:

$$\mathcal{V}_{ji} = v(x_i, h_i, z_j) + \eta_{ji}, \tag{6}$$

[17] This function is consistent with a labor demand model allowing for degree-specific and location-specific exogenous productivity, as in Diamond (2016). Note, however, that it does not allow for individuals' location choices to influence wages in equilibrium, which rules out the existence of within-location spillovers across workers and imperfect substitutability of workers with different skills (Katz and Murphy, 1992). These effects are likely to be second-order in the short run, as I consider the location decisions of *entry-level workers* only, which represents a small fraction of the overall labor supply.

where $v(x_i, h_i, z_j)$ is the deterministic component of this score. I normalize to zero the payoff associated with programs' outside option, $i = 0$, which corresponds to leaving a seat empty. η_{ji} is an idiosyncratic preference shock, independently distributed according to a Type I Extreme Value distribution.

4.2 Higher education market equilibrium

In this section, I characterize the equilibrium of the higher education market. I first describe how programs rank applicants on the platform. I then define the equilibrium concept and discuss its implications for enrollment outcomes.

4.2.1 Programs' ranking of applicants

I assume that programs rank applicants according to the admission score defined in Equation (6), which reflects programs' priorities after incorporating the government's equity objectives. This implies that the student ranked r^{th} by program j, denoted $i^*_{j,r}$, has the highest program-specific score among applicants who have not been ranked higher:

$$i^*_{j,r} = \underset{i \in \mathcal{A}_j \setminus \{i^*_{j,k}\}_{k=1}^{r-1}}{\arg\max} \mathcal{V}_{ji}, \tag{7}$$

where $\mathcal{A}_j$ is the set of applicants to program j, including the outside option $i = 0$.[18]

While programs do not disclose their exact ranking rules, this assumption is consistent with interviews with admission committees (Section 2), suggesting that programs typically select weights on observed applicant characteristics and use an algorithm to generate rankings. This behavior is made easier as the platform provides an online tool to generate a truthful ranking based on the program's chosen weights. When needed, rankings are adjusted to satisfy equity objectives.[19]

4.2.2 Higher education market: equilibrium

On the higher education market, each student i is matched to a program $j \in \mathcal{J} \bigcup \{0\}$. I assume that the resulting match is stable. A matching is stable if there exists no student-

[18] I assume that η_{ji} does not enter in students' preferences, such that $\mathcal{A}_j$ is exogenous.

[19] For example, if scholarship-eligible applicants are initially concentrated at the bottom of the ranking, some may be up-ranked. These adjustments preserve (i) the relative order among scholarship-eligible applicants and (ii) the relative order among non-eligible applicants.

program pair that would like to break their current match and re-match with each other (Gale and Shapley, 1962). Formally, this means that there is no (i, j)-pair, with $i \in \mathcal{I} \bigcup \{0\}$ and $j \in \mathcal{J} \bigcup \{0\}$, such that:

$$\mathcal{U}_{ij} > \mathcal{U}_{i\mu(i)} \text{ and } \mathcal{V}_{ji} > \mathcal{V}_{j\mu(j,\mathcal{R}_j)}, \tag{8}$$

where $\mu(i)$ is the program to which student i is matched, $\mathcal{R}_j$ is the rank of the last-admitted student at program j, and $\mu(j, \mathcal{R}_j)$ is the identity of that student. Such an equilibrium exists as long as there are no complementarities between admitted students, which is consistent with programs ranking applicants individually and extending offers in rank order. Regarding uniqueness, Azevedo and Leshno (2016) prove that a continuum model, where a finite number of programs are matched to a continuum of students, typically has a unique stable matching. They further prove that all of the stable matchings of a discrete economy converge to this unique stable matching of the continuum economy. With its very large number of applicants, the French college admission system qualifies as a large discrete economy.

Stability implies that each student enrolls in her preferred program among those willing to admit her. The resulting match can therefore be written as the solution to a discrete choice problem with student-specific choice sets (Fack et al., 2019):

$$\mu(i) = \underset{j \in \mathcal{F}_i \bigcup 0}{\arg\max} \mathcal{U}_{ij},$$

where $\mathcal{F}_i$ is the set of programs that would admit student i, i.e. student i's *feasible set*. It is defined as the set of programs preferring to admit student i over their last-admitted applicant:

$$\mathcal{F}_i = \{j \in \mathcal{J} : \mathcal{V}_{ji} > \mathcal{V}_{j\mu(j,\mathcal{R}_j)}\}. \tag{9}$$

Note that if a program leaves at least one seat empty, the identity of the last admitted student is the outside option, with associated utility $\mathcal{V}_{j0}$. Seats may remain empty either because the program finds remaining applicants undesirable (preferring $i = 0$) or because it receives fewer applications than seats. In both cases, the program would then be feasible for a student i only if the latter is more desirable than the outside option.

Overall, assuming that the equilibrium is stable allows me to express the observed match as a function of the model primitives without having to take a stand on students' applica-

tion behavior. Stability only requires that each student applies to their preferred feasible program. This is consistent with several plausible application strategies, making it a good approximation of what happens on the platform. For example, the equilibrium would be stable if students applied to all programs they prefer to the outside option and accepted their favorite offer among those they receive.[20] While the number of applications is limited to ten in theory, which would make such behavior less plausible, the effective number of applications allowed is much larger in practice: applications to different institutions offering the same program type and major are typically processed as one. As a result, 25% of applicants submit more than 13 applications, while the maximum number of applications observed in the data is 162 (Table A.3).

Stability is also compatible with students being strategic and submitting applications that maximize their expected utility. In this case, a stable equilibrium is likely to arise when students are able to anticipate which programs are feasible. This is made easier by the fact that the market is very large, with many students and many seats per program. Admission probabilities then become close to degenerate, making it unlikely that students with rational expectations omit their favorite feasible program from their application portfolio (Fack et al., 2019). Consistent with this, the rank of a program's last admitted applicant is highly persistent across years: the correlation between 2018 and 2019 is 0.96.

5 Identification and Estimation

In this section, I show how the primitives of the model can be identified from data on individuals' location decisions at labor market entry, programs' rankings of applicants, and realized matches on the higher education market. Importantly, identification does not rely on functional-form assumptions on the deterministic part of the payoffs. I then describe the estimation strategy.

5.1 Identification

I present the identification argument, which proceeds by backward induction. I first show that preferences over locations can be identified from data on location choices at entry on

[20] Several studies of school-choice mechanisms assume students are truthful (Abdulkadiroğlu et al., 2017; Abdulkadiroğlu et al., 2020). The requirement here is weaker, since students do not submit rank-ordered lists on the French platform.

the labor market. I then turn to the first stage of the model and describe how programs' priorities are identified from observed rankings of applicants. Finally, I show how to identify students' preferences over programs from data on realized matches. An exclusion restriction identifies the scale of students' idiosyncratic preference shocks, σ, avoiding the need to fix the relative scale of payoffs across stages of the model.

Individuals' preferences over locations After graduation, individuals choose which local labor market $\ell \in \mathcal{L}$ to enter. This second-stage problem is a static discrete choice model. Let $\boldsymbol{w}_i$ denote individual i's expected wages across locations. The share of individuals with characteristics (x_i, h_i, y_i) working in location ℓ, $\forall \ell \in \mathcal{L}$, denoted $\Pr(\ell^* = \ell | x_i, h_i, y_i, \boldsymbol{w}_i)$, can be mapped into the deterministic part of individuals' preferences over local labor markets:

$$\Pr(\ell^* = \ell | x_i, h_i, y_i, \boldsymbol{w}_i) = \frac{\exp(\theta(w_{i\ell} + \omega_\ell(x_i, h_i, y_i)))}{\sum_{\kappa \in \mathcal{L}} \exp(\theta(w_{i\kappa} + \omega_\kappa(x_i, h_i, y_i)))}. \tag{10}$$

The left-hand side of this equation can be directly identified from data on individuals' location choices at labor market entry. As this mapping can be inverted, I can recover θ and $\omega_\ell(x_i, h_i, y_i)$. Details are provided in Appendix C.3.

From this, the continuation value associated with enrolling in a program with characteristics z_j for a student with characteristics (x_i, h_i) can be identified:

$$\bar{U}(x_i, h_i, z_j) = \frac{1}{\theta} \int \beta^{t(y_i)} \left(\gamma + \log \left[\sum_{\ell \in \mathcal{L}} \exp(\theta(w_{i\ell} + \omega_\ell(x_i, h_i, y_i))) \right] \right) f\left(y_i \mid x_i, h_i, z_j\right) dy_i,$$

where γ is the Euler constant. The transition probabilities $f\left(y_i \mid x_i, h_i, z_j\right)$ are identified from administrative records of students' paths in higher education.

Program priorities over students Programs observe applicants and rank them based on each student's program-specific score defined in Equation (6). Let $\boldsymbol{x}$ and $\boldsymbol{h}$ collect x_i and h_i, respectively, for all i. The probability that student i is ranked r^{th} by program j, denoted $\Pr(i^*_{j,r} = i | \boldsymbol{x}, \boldsymbol{h}, z_j)$, admits a closed-form expression:

$$\Pr(i^*_{j,r} = i | \boldsymbol{x}, \boldsymbol{h}, z_j) = \frac{\exp(v(x_i, h_i, z_j))}{\sum_{a \in \mathcal{A}_j \setminus \{i^*_{j,k}\}_{k=1}^{r-1}} \exp(v(x_a, h_a, z_j))}. \tag{11}$$

The left-hand side of this equation is directly identified from data on programs' rankings. Since this mapping is invertible, $v(x_i, h_i, z_j)$ can be recovered. Details are provided in Appendix C.4.

Students' preferences over programs Finally, given the objects identified above, I show that one can identify students' preferences from realized matches. When the criteria and weights used to rank students are fully observed, as in the majority of school choice systems, one can directly construct each student's feasible set and recover preferences from a discrete choice model with observed, student-specific choice sets (Fack et al., 2019). In many college admission mechanisms, however, admission scores are not observed, so feasible sets cannot be constructed directly.[21]

When feasible sets are unobserved, identification of students' preferences from realized matches typically requires two types of exclusion restrictions—a shifter of students' preferences and a shifter of programs' priorities (Agarwal and Somaini, 2025; He et al., 2024). I show how additional data on programs' rankings of applicants allow me to circumvent the need for such exclusion restrictions. In particular, the share of students with observable characteristics (x_i, h_i) who are matched to program j can be written as:

$$\begin{aligned}\Pr(\mu(i) = j | x_i, h_i, \boldsymbol{z}) = &\left[\sum_{\mathcal{F} \in \mathbf{F}} \Pr(\mathcal{F}_i = \mathcal{F} | x_i, h_i, \boldsymbol{z}) \Pr(\mathcal{U}_{ij} \geq \mathcal{U}_{ik}, \forall k \in \mathcal{F} | x_i, h_i, \boldsymbol{z}, \mathcal{F})\right] \\ &\times \Pr(j \in \mathcal{F}_i | x_i, h_i, \boldsymbol{z}),\end{aligned} \tag{12}$$

where $\mathbf{F}$ is the set of all potential feasible sets and $\boldsymbol{z}$ collects program characteristics for all j. The first term inside the sum is the probability that student i faces feasible set $\mathcal{F}$, which is governed by programs' priorities through Equation (9). The second term is the probability that program j is preferred to all programs in $\mathcal{F}$, which admits a closed-form expression:

$$\Pr(\mathcal{U}_{ij} \geq \mathcal{U}_{ik}, \forall k \in \mathcal{F} | x_i, h_i, \boldsymbol{z}, \mathcal{F}) = \frac{\exp(\frac{1}{\sigma}(u(x_i, z_j) - c(x_i, d_{ij}) + \bar{U}(x_i, h_i, z_j)))}{\sum_{k \in \mathcal{F} \bigcup 0} \exp(\frac{1}{\sigma}(u(x_i, z_k) - c(x_i, d_{ik}) + \bar{U}(x_i, h_i, z_k)))} \tag{13}$$

where $\bar{U}(x_i, h_i, z_j)$ has been identified in a first step.

Students' preferences over programs are identified in two steps. First, because $v(x_i, h_i, z_j)$

[21] This is the case in most college admission processes, such as in the United States, where universities base their decisions on a discretionary combination of test scores, GPA, and extra-curricular information.

is identified from programs' rankings, one can construct students' feasible sets as a function of observables and programs' shocks. For a given draw of programs' shocks, the feasible set for a student with characteristics (x_i, h_i) is:

$$\mathcal{F}(x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta}) = \{j \in \mathcal{J} : v(x_i, h_i, z_j) + \eta_{ji} > v(x_{\mu(j,\mathcal{R}_j)}, h_{\mu(j,\mathcal{R}_j)}, z_j) + \eta_{j\mu(j,\mathcal{R}_j)}\}, \tag{14}$$

where $\boldsymbol{\eta}$ collects η_{ji} for all programs and students.[22] Equation (12) can then be rewritten as an integral over $\boldsymbol{\eta}$:

$$\frac{\Pr(\mu(i) = j | x_i, h_i, \boldsymbol{z})}{\Pr(j \in \mathcal{F}(x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta}) | x_i, h_i, \boldsymbol{z})} = \int \Pr(\mathcal{U}_{ij} \geq \mathcal{U}_{ik}, \forall k \in \mathcal{F}(x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta}) | x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta}) f_\eta(\boldsymbol{\eta}) d\boldsymbol{\eta},$$

where $f_\eta(.)$ is the density of $\boldsymbol{\eta}$. The numerator on the left-hand side is identified from realized matches. The denominator is the probability that program j is feasible for student i, which can be computed from observed applicant sets and Equation (14). The interior of the integral has a closed-form expression given by Equation (13), and the distribution of $\boldsymbol{\eta}$ is known.

Second, this mapping can be inverted to identify, sequentially, the scale parameter σ and the functions $u(x_i, z_j)$ and $c(x_i, d_{ij})$. I identify σ using an exclusion restriction. Specifically, I assume there exists a variable, h_i, that affects the stage-two payoff but is excluded from the stage-one flow utility of enrolling in a given program. I take this excluded variable to be the home province, which shifts the continuation value through home bias in location choice. Because mobility frictions within the higher education market are captured flexibly via $c(x_i, d_{ij})$, this exclusion restriction is fairly weak. Comparing students with the same x_i but different home provinces then identifies σ (Appendix C.5). Given the parameters related to individuals' preferences over locations and the scale parameter previously identified, one can then non-parametrically identify the functions $u(x_i, z_j)$ and $c(x_i, d_{ij})$. This does not rely on exclusion restrictions. The proof is constructive and can be found in Appendix C.6.

5.2 Estimation

This section describes the parameterization choices and the estimation strategy. As in the identification argument, estimation proceeds by backward induction, starting with the second stage of the model.

[22] As in Agarwal and Somaini (2025), equilibrium cutoffs can be taken as fixed as I consider a large economy (Azevedo and Leshno, 2016). Hence, feasible sets only depend on observable characteristics and $\boldsymbol{\eta}$.

Preferences over local labor markets I parametrize the flow utility associated with entering the labor market in location ℓ for individual i as follows:

$$\omega_\ell(x_i, h_i, y_i) = g_i'\boldsymbol{\phi}_1 \mathbb{1}\{\ell = \ell_i^G = h_i\} + g_i'\boldsymbol{\phi}_2 \mathbb{1}\{\ell = h_i \neq \ell_i^G\} + g_i'\boldsymbol{\phi}_3 \mathbb{1}\{\ell = \ell_i^G \neq h_i\} + \delta_\ell,$$

where g_i is a vector of degree-at-graduation dummies. I distinguish three degree types: no tertiary degree, a vocational degree, and a bachelor's degree or more. ℓ_i^G denotes the graduation location and h_i the home location. The coefficients $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$ capture home bias—staying in or returning to the home location depending on whether the individual graduated there. $\boldsymbol{\phi}_3$ reflects individuals' taste for remaining in their graduation location when it differs from their home location, capturing moving costs. These effects are allowed to vary by degree type. δ_ℓ denotes location fixed effects capturing differences in amenities.

I parameterize the systematic wage component, $w_\ell(x_i, h_i, y_i)$, as follows:

$$w_\ell(x_i, h_i, y_i) = g_i'\boldsymbol{\kappa}_\ell + \boldsymbol{\kappa} a_i^w,$$

where $\boldsymbol{\kappa}_\ell$ captures that returns to different degrees may differ across locations. a_i^w is a subset of the applicants' characteristics in x_i, including high school GPA, gender, and scholarship status.[23]

The wage equation is estimated via an auxiliary linear regression. The primitive parameters capturing preferences over locations, $\boldsymbol{\Phi} = \{\theta, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \boldsymbol{\phi}_3, \{\delta_\ell\}_{\forall \ell}\}$, are estimated by maximum likelihood, where the log-likelihood function is constructed using the conditional choice probabilities in Equation (10). I assume that individuals choose which province to enter in metropolitan France, that is, $\mathcal{L}$ is the set of the 96 French provinces.

Programs' priorities I parameterize program j's admission score for applicant i as a flexible function of program and student characteristics:

$$v(x_i, h_i, z_j) = q_j'\boldsymbol{\Lambda} a_i + q_j'\boldsymbol{\tau} s_{ij},$$

where q_j is a vector of dummies for program type, included in z_j. I distinguish seven types: three-year university programs, two-year elite programs, engineering and business schools,

[23] High school GPA is grouped into four honors categories.

two-year technology-oriented programs, two-year vocational programs, nursing schools, and other programs.[24] The vector a_i includes the same variables as a_i^w, as well as high school track and socioeconomic status. The vector s_{ij} collects additional indicators for institutional priorities specific to student-program pairs. In particular, it includes indicators for same-catchment matches when program j is subject to place-based admission priorities, and for applicants from the high school in which the program is located.[25] The programs' priority parameters, $\{\boldsymbol{\Lambda}, \boldsymbol{\tau}\}$, can be estimated by maximum likelihood, where the log-likelihood function is constructed using the conditional rank probabilities in Equation (11).

Estimation of students' preferences The value and mobility costs associated with enrolling in program j for student i are parametrized as follows:

$$u(x_i, z_j) = q_j' \boldsymbol{\delta} a_i \qquad\qquad c(x_i, d_{ij}) = d_{ij} \boldsymbol{\alpha}_1 a_i + d_{ij}^2 \boldsymbol{\alpha}_2 a_i,$$

where $u(.)$ is specified as a flexible interaction between program type and student characteristics a_i. The mobility cost term $c(.)$ depends on the distance, in kilometers, between student i's home ZIP code and program j's GPS coordinates, denoted d_{ij}. Including d_{ij}^2 allows for non-linearities in mobility costs. These costs are allowed to vary with student characteristics.

As shown in Section 5.1, the conditional probability that a student with observables (x_i, h_i) is matched to j is an integral over $\boldsymbol{\eta}$ that does not admit a closed-form expression. I therefore estimate these parameters using a simulated method of moments procedure, following Gourieroux and Monfort (1996) and Train (2009), with optimal moments given by the score of the simulated likelihood.[26] This procedure provides a computationally tractable way to obtain consistent and efficient estimates in this two-sided matching framework.

Estimation of transition probabilities I let students' graduation outcomes be the result of a stochastic process, depending on their characteristics and on the characteristics of their enrolled program. As a result, students face different probabilities of leaving higher education with no tertiary degree, a vocational degree, or a bachelor's degree or more, depending on their enrollment choice. I also allow these characteristics to affect graduation location, so I do not impose that students graduate in the same location as their program of enrollment.

[24] High school GPA enters continuously in $v_{ji}(x_i, h_i, z_j)$.
[25] Two-year elite programs and two-year vocational programs are typically located in high schools.
[26] I use ten draws.

The transition probabilities from initial enrollment to degree-location outcomes can be recovered directly from the data. I use administrative records covering the universe of students enrolled in French higher education to compute the joint probability that a student with characteristics x_i and h_i, enrolled in a program with characteristics z_j, graduates with a given degree in a given location, that is, $y_i = (g_i, \ell_i^G)$.

Although transition probabilities are identified non-parametrically, I estimate them parametrically for tractability. Specifically, I use a conditional logit model for the probability that an individual graduates with a given degree-location bundle among $|\mathcal{G}| \times |\mathcal{L}|$ possibilities, where $\mathcal{L}$ denotes the set of graduation locations, which coincides with the set of local labor markets, and $\mathcal{G}$ denotes the set of degree types: no tertiary degree, a vocational degree, and a bachelor's degree or more. I parameterize this joint probability as follows:

$$f(y_i \mid x_i, h_i, z_j) \equiv \Pr(g_i, \ell_i^G \mid x_i, h_i, z_j) = \frac{\exp(\rho_1 \mathbb{1}\{\ell_i^G = \ell_j\} + g_i' \boldsymbol{\rho}_2 q_j + g_i' \boldsymbol{\rho}_3 a_i + \delta_{\ell_i^G}^G)}{\sum_{\ell \in \mathcal{L}} \sum_{g \in \mathcal{G}} \exp(\rho_1 \mathbb{1}\{\ell = \ell_j\} + g' \boldsymbol{\rho}_2 q_j + g' \boldsymbol{\rho}_3 a_i + \delta_\ell^G)}$$

where ℓ_j is the province in which program j is located and δ_ℓ^G are location fixed effects. The set of parameters $\boldsymbol{\rho} = \{\rho_1, \boldsymbol{\rho}_2, \boldsymbol{\rho}_3, \{\delta_\ell^G\}_{\forall \ell}\}$ is estimated by maximum likelihood. Additional details are provided in Appendix D.2.

6 Estimation Results and Model Fit

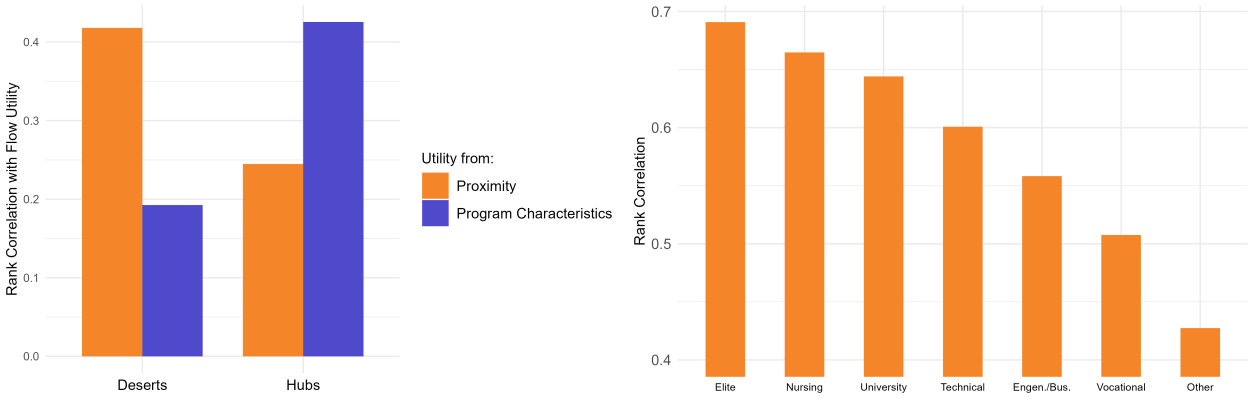
In this section, I first present the parameter estimates and then discuss the fit of the model.

6.1 Estimation results

Students' preferences over programs Estimates of students' preferences over programs show that demand is highly sensitive to distance (Tables A.14-A.16). The average student would be willing to forgo 741 euros in future monthly entry-level wages, equivalent to 45% of the average entry-level wage, to enroll in a program located 100 kilometers closer to her home location. There is also substantial heterogeneity across socioeconomic groups: this amount reaches 886 euros per month for low-SES students, which is 307 euros more than for high-SES students.

Panel a) of Figure 5 further illustrates the importance of mobility frictions in students' higher education choices. For each applicant, I rank programs according to their flow utility

Figure 5: Determinants of Students' and Programs' Preferences



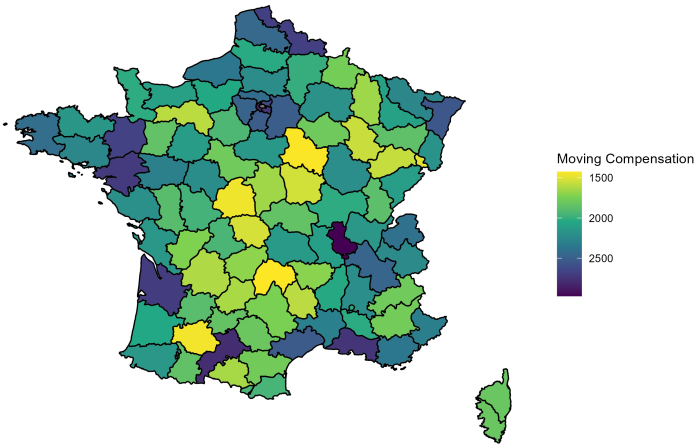
a) Students **b) Programs**

NOTES. Panel a) illustrates the determinants of the payoff students derive from enrolling in a given program. Using data at the student-program level, it plots, in orange, the rank correlation between the flow utility associated with a program and its proximity to the student's home location. It also plots, in purple, the rank correlation between the flow utility associated with a program and the consumption value, $u(x_i, z_j)$, associated with the program's characteristics, z_j. The analysis restricts attention to each student's top fifty programs. The figure reports these correlations separately for students from higher education deserts (the 50 provinces with less than 5% of the total supply of academic seats) and hubs (the ten provinces at the top of the seat supply distribution). Panel b) illustrates the importance of students' GPA in programs' rankings of applicants. It plots the rank correlation between the program-specific score associated with an applicant, $v(x_i, h_i, z_j)$, and the GPA of the student. The figure reports these correlations separately by program type.

and then compute the correlation between this ranking and the ranking of programs by proximity to the student (in orange).[27] I do so separately for applicants from higher education deserts and hubs. I define higher education deserts as the fifty provinces that jointly account for less than 5% of academic seats while being home to 25% of students. Higher education hubs are the ten provinces at the top of the academic-seat distribution and also contain 25% of students. While this correlation is high for all students, it is about 70% higher for students from higher education deserts. This reflects the fact that these applicants have few nearby higher education options, making distance a central component of program payoffs. By contrast, students from higher education hubs have many more nearby options, so their ranking is less sensitive to distance. As a result, their ranking is driven more by the utility associated with other program characteristics. The correlation between the utility rank of a program and its rank based only on the non-distance component, $u(x_i, z_j)$, is twice as large for students from hubs (42%) as for students from higher education deserts (19%).

[27] A higher rank corresponds to a program closer to the student's home location. Since distance enters students' payoffs as a cost, ranking programs by proximity is equivalent to ranking them by $-c(x_i, d_{ij})$.

Figure 6: Location Choices - Home Bias



NOTES. This map uses estimated preferences over local labor markets (Table A.18) to compute the compensation, in monthly euros, that an individual from a given province would require to be willing to move to Paris, conditional on having studied in their home location and having obtained at least a bachelor's degree, abstracting from wage differences.

Programs' priorities Table A.17 presents estimates of programs' priorities. Panel b) of Figure 5 shows how programs' rankings of students correlate with GPA. While there is heterogeneity across programs, this correlation is high for all program types, underscoring the importance of academic achievement in admission decisions. Final rankings also depend on other applicant characteristics. For example, eligibility for an income-based scholarship has an effect on admission to two-year elite programs equivalent to increasing GPA by 0.27 standard deviations. Programs located in high schools also tend to prefer students from their own institution.

Preferences over local labor markets Table A.18 presents estimates of the parameters governing the utility from entering the labor market in a given location. The results show that workers have a strong preference for their home location. In Figure 6, I use these estimates to compute the compensation an individual would require to enter the labor market in Paris rather than in their home location, holding wages fixed at the home-location level. For example, an individual from Creuse—one of the provinces at the bottom of the higher education supply distribution—who graduates with at least a bachelor's degree in Creuse would need to receive 1,528 euros per month to accept a job in Paris rather than in her home location. This reflects a strong home bias, as this amount represents about 60% of the average wage difference between Paris and Creuse. This home bias is even stronger for

individuals with no tertiary degree, as shown in Table A.18. The cross-province variation in Figure 6 reflects differences in amenities: workers from provinces with higher amenity levels require greater compensation to be willing to leave their home location.

While the preference for home is a key determinant of location choice, workers also face large moving costs. This is illustrated in Figure A.7, where I compute the same object as above, but conditional on graduating in Paris rather than at home. A student from Creuse who moved to Paris to study and graduated there with at least a bachelor's degree would now be better off entering the labor market in Paris, holding wages equal across labor markets. This individual would need a monthly tax rebate of 961 euros to be willing to return home upon graduation. Because home bias is stronger for individuals with lower levels of education, an individual who moved to Paris to study but left higher education without a tertiary degree would need to receive only an additional 247 euros per month to return home. Overall, the magnitude of these moving costs likely allows higher education hubs to retain students graduating there, despite the existence of a large home bias.

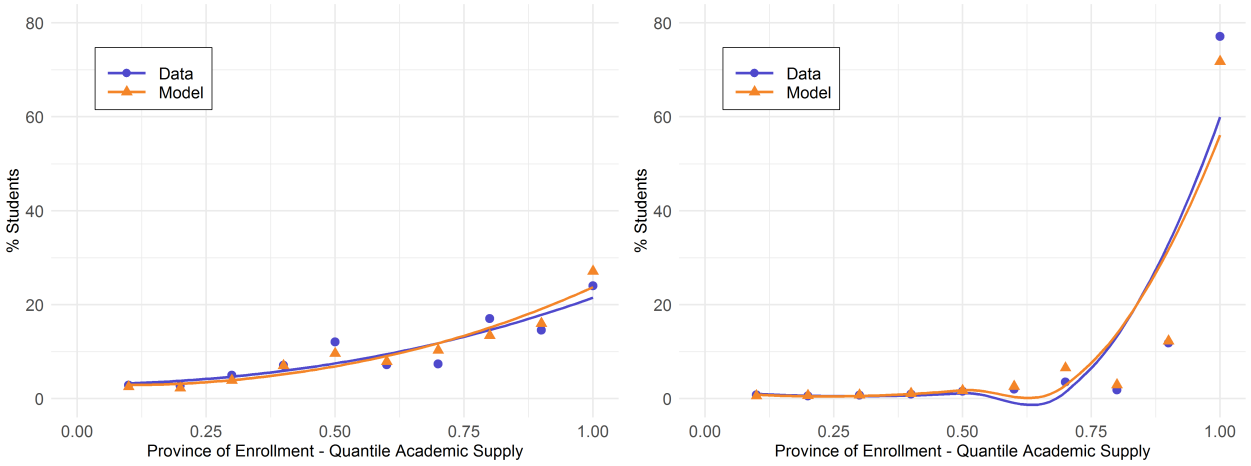
6.2 Model fit

I simulate higher education matches and location choices at labor market entry to evaluate the fit of the model. Table A.20 compares these simulations with the data. Panel A presents characteristics of the observed higher education match in Column (1) and of the simulated match in Column (2). Panel B reports patterns related to location choices.[28] Overall, the simulated shares of students enrolling in different program types, the distance to the matched program, and the shares of individuals entering the labor market in their home or study location closely match their empirical counterparts.

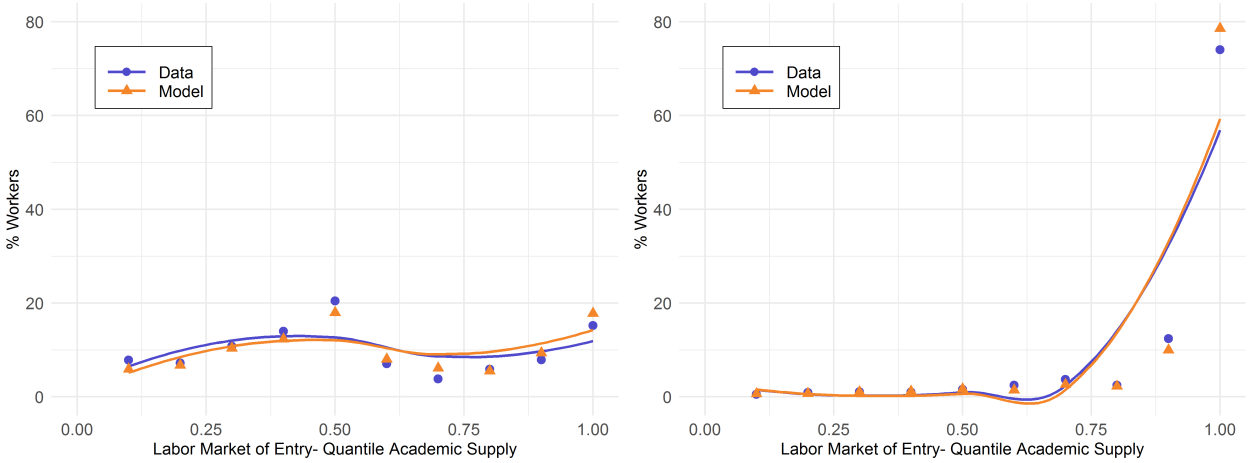
I also evaluate the fit of the model using moments not directly targeted in the estimation. In particular, it is important for the model to replicate migration flows both within the higher education market and across local labor markets. First, I simulate enrollment flows between provinces and compare them with those observed in the data. I rank provinces according to their supply of academic seats and group them into ten bins. Panel a) of Figure 7 plots the share of students from higher education deserts enrolled in each of these bins. About 25% of

[28] I use both students' preference parameters (Tables A.14-A.15) and programs' priority parameters (Table A.17) to compute their payoffs. I then feed them into a Deferred Acceptance algorithm to simulate the stable match. I use individuals' preferences over locations (Table A.18) to simulate their local labor market choice upon graduation.

Figure 7: Model Fit



a) Enrollment Flows - Students from Deserts **b) Enrollment Flows - Students from Hubs**



c) Work Flows - Workers from Deserts **d) Work Flows - Workers from Hubs**

NOTES. I rank provinces according to their supply of academic seats and group them into ten bins of equal size. Panel a) plots the share of students from higher education deserts enrolling in each of these bins. For example, about 25% of these students enroll in a program located in a province in the top 10% of the academic-seat distribution. Panel b) plots the same statistics for students from higher education hubs. Panel c) shows the share of workers from higher education deserts entering the labor market in each bin. For example, about 15% of them enter the labor market in a province belonging to the top 10% of the academic-seat distribution. Panel d) displays the same statistics for workers from higher education hubs. Each panel presents these flows in the data (orange triangles) and as predicted by the model (purple dots).

students from higher education deserts enroll in a higher education hub, that is, a province belonging to the top 10% of the academic-seat distribution. This contrasts sharply with the enrollment patterns of students from higher education hubs, as 75% of them remain in a hub to study (Panel b). The model replicates these enrollment flows very closely for both groups, although they were not directly targeted in the estimation.

I further use simulated enrollment flows at the province level to estimate the same gravity model as in Section 3, regressing them on origin and destination fixed effects and on the distance between each location pair. Table A.21 presents results from the gravity model estimated on enrollment flows observed in the estimation sample (Columns (1) and (2)) and on simulated flows (Columns (3) and (4)). The latter yields an elasticity of enrollment flows with respect to distance of -1.91, which is very close to the elasticity estimated in the data (-1.86).

I then turn to worker flows across local labor markets, which are key objects in the counterfactual analysis. Panel c) of Figure 7 plots the share of individuals from higher education deserts entering the labor market in provinces by decile of the academic-seat distribution. Only about 15% of them enter the labor market in a higher education hub. Again, this contrasts sharply with the behavior of individuals from higher education hubs, who are 60 percentage points more likely to enter the labor market in a hub. Simulated flows match those patterns closely.

Finally, I evaluate the fit of the model by using the preference estimates to predict the effect of a 500 euro moving subsidy, such as the one evaluated in Section 3.2.1. This scholarship corresponds to a monthly payment of 41.66 euros. The preference estimates imply that, for a scholarship-eligible student, increasing the distance to the enrolled program by one kilometer entails a utility cost of about 8.56 euros per month. Receiving 41.66 euros per month would therefore offset the utility cost associated with an additional 4.86 kilometers, a figure that lies within the 95% confidence interval of the estimated effect of the subsidy (Table A.8).

7 Counterfactuals

In this section, I first use the model to quantify how mobility frictions interact with the uneven within-country distribution of higher education options to generate spatial inequalities across

individuals and local labor markets. To do so, I eliminate mobility frictions within the higher education market, holding fixed the distribution of programs. Second, I explore whether mobility scholarships combined with incentives for students to return home upon graduation could help low-opportunity areas invest in the human capital of their local labor force.

7.1 The geography of higher education and spatial inequalities

I first quantify the impact of mobility frictions on spatial inequalities across students. Panel A of Table 3 reports several measures of these inequalities at baseline (Column (1)) and under a counterfactual in which mobility frictions in the higher education market are eliminated (Column (2)).[29] The model captures three key potential drivers of these inequalities: differences in students' test scores, differences in students' preferences, and mobility frictions. Simulating individuals' choices in the absence of mobility frictions, while holding fixed the distribution of programs, therefore isolates the role of these frictions in generating spatial inequalities through their interaction with the geography of the higher education market.

At baseline, the gap in the share of individuals graduating with at least a bachelor's degree between higher education hubs (47.5%) and deserts (41.9%) is 5.59 percentage points. Column (2) shows that this gap would decrease by half in the absence of mobility frictions within the higher education market, reaching 2.82 percentage points. Overall, mobility frictions account for about 13% of the variance in educational attainment across provinces.

The model also allows me to quantify average welfare differences across students, where welfare is computed using Equation (3). At baseline, the welfare gap between students from higher education hubs and deserts amounts to 728 euros per month. While all students would benefit from the removal of mobility frictions, those from higher education deserts would gain twice as much. As a result, the welfare gap between the two groups would decrease by 39%.

Overall, mobility frictions, combined with the uneven distribution of higher education programs, account for a substantial share of spatial inequalities in educational attainment and welfare across students. They also reduce the overall stock of human capital, as eliminating them increases the share of individuals obtaining a bachelor's degree or more by two percentage points.

The model further allows me to quantify the long-run impact of mobility frictions on the spatial distribution of labor supply across local labor markets (Panel B of Table 3). I find

[29] Details on the computation of all measures considered in this section are presented in Appendix E.2.

Table 3: Mobility Frictions and Spatial Inequalities: Quantification

	Baseline	Counterfactual
	(1)	(2)
Panel A: Spatial inequalities between individuals		
Share High-Skilled:		
Gap Hubs-Deserts	5.59	2.82 (-49.63%)
Cross-Province Variance	33.94	29.67 (-12.58%)
Mean Welfare (€/month):		
Gap Hubs-Deserts	728	441 (-39.42%)
Panel B: Inequalities across local labor markets		
Share High-Skilled:		
Gap Hubs-Deserts	9.42	6.13 (-35.00%)
Cross-Province Variance	34.16	31.85 (-6.77%)
Sum Wages (M€):		
Gap Hubs-Deserts	163.77	172.66 (5.42%)
Cross-Province Variance	189.35	218.49 (15.39%)
Panel C: Aggregate Outcomes		
Share High-Skilled	44.61	46.62 (4.51%)
Sum Wages (M€)	999.95	1,012.77 (1.28%)

NOTES. This table presents results from the main quantification exercise: removing mobility frictions within the higher education market. Panel A focuses on spatial inequalities between individuals with respect to the share of individuals from a given province graduating with at least a bachelor's degree (share high-skilled, in percent) and welfare, based on Equation (3), expressed in monthly euros. Panel B focuses on inequalities across local labor markets, measured by the share of their labor force with at least a bachelor's degree (in percent) and by the sum of entry-level workers' wages. Gaps in these variables are computed by comparing hubs and deserts—defined as the provinces at the top and bottom of the academic seat distribution, both gathering 25% of the high school student population — and by measuring their variance across all provinces. Panel C describes aggregate outcomes. Column (1) (Column (2), resp.) presents results for the status quo (counterfactual simulations, resp.). Appendix E.1 and E.2 detail how the counterfactual simulations are performed.

two effects. On the one hand, mobility frictions reduce the human capital available in low-opportunity areas by discouraging the local population from pursuing tertiary education. In particular, they account for 35% of the 9.4 percentage point gap between hubs and deserts in the share of entry-level workers entering the labor force with at least a bachelor's degree. On the other hand, moving costs insulate low-opportunity areas from an outflow of their local population. When these frictions are eliminated, higher education hubs retain a significant share of the additional students they attract, due to moving costs and differences in returns to education across provinces, resulting in a 9% increase in the number of individuals from deserts entering the labor market in hubs (Table A.22). This second effect dominates when considering regional inequalities in output generated by entry-level workers. In the counterfactual, inequalities across local labor markets, measured by the variance of aggregate wages earned by entry-level workers across local labor markets, increase by 15.4% (Column (2)).

The gap between higher education hubs and deserts in this measure similarly rises by 5.42% relative to the status quo. This increase in regional inequality, driven by the reallocation of workers toward hubs, is accompanied by an increase in the aggregate wages of entry-level workers, reflecting both their higher human capital and their expanded access to local labor markets with higher returns to education.

Table 4: Quantification

	No Mobility Frictions within Higher Education Market		
	Only	No Admission Restrictions	No Moving Costs
	(1)	(2)	(3)
Panel A: Spatial inequalities between individuals - Gap Hubs-Deserts			
Share High-Skilled	-49.63%	-34.93%	18.58%
Mean Welfare (€/month):	-39.42%	-48.76%	-19.06%
Panel B: Inequalities across local labor markets - Gap Hubs-Deserts			
Share High-Skilled	-35.00%	-21.98%	6.65%
Sum Wages (M€)	5.42%	20.56%	-10.95%

NOTES. This table reports results from the main counterfactual exercise presented in Table 3 (Column (1)), a counterfactual ignoring the supply side of the higher education market in addition to removing mobility frictions (Column (2)), and a counterfactual eliminating both mobility frictions in the first stage and moving costs between the study and home location in the second stage (Column (3)). Appendix E.1 and E.2 provide more details about the counterfactual simulations.

The magnitude of these effects would be substantially different if the two-sided nature of the higher education market was ignored. Column (2) of Table 4 reports results from a counterfactual in which mobility frictions are eliminated and each student is allowed to enroll in her preferred program, regardless of capacity constraints and programs' admission decisions. Educational inequalities would still decline, but by 15 percentage points less than in the equilibrium counterfactual. This difference reflects two forces. First, some students from higher education deserts would now enroll in programs they prefer but from which they are less likely to graduate. Second, students from higher education hubs would gain access to previously congested programs from which they graduate. On the labor market side, inequalities across provinces would increase much more than when the higher education market clears in equilibrium. Because any student can now enroll in a program located in a hub, a larger share of workers leave their home province (Table A.22). In particular, the share of workers entering the labor market in Paris would increase by 34% relative to the status quo (Column (3) of Table A.23). Accounting for capacity constraints and programs'

admission decisions is therefore central to capturing the distributional effects of mobility frictions.

Overall, eliminating mobility frictions within the higher education market leads to an outflow of students toward higher education hubs and, although human capital is portable, this initial effect persists: many of the students who leave to study do not return home at labor market entry. In the model, this persistence reflects both location-specific returns to education and costs of moving from the graduation location. To isolate the role of the latter, I simulate study and location choices after removing moving costs between graduation and home, in addition to mobility frictions within the higher education market. Educational attainment in low-opportunity areas would still increase in this scenario (+4.13% relative to the status quo), but less than in the main counterfactual, because lower costs of returning home reduce the incentive for students from these areas to obtain a bachelor's degree or more, given spatial differences in returns to education.[30] However, because more students return home, inequalities across local labor markets decline. This suggests that low-opportunity areas could benefit from outsourcing the education of their local labor force to higher education hubs, provided they can also incentivize students to return.

7.2 Outsourcing education through mobility scholarships

Mobility frictions within the higher education market hurt the educational attainment of students from low-opportunity areas, but the results above show that eliminating these frictions would lead to an outflow of workers from these areas. Could policymakers in higher education deserts design policies that expand access to distant higher education opportunities while strengthening the human capital of their local labor force?

To answer this question, I study a policy that combines mobility scholarships with incentives for students to return home upon graduation. In particular, I consider a program targeted at students from higher education deserts that offers a subsidy covering the average utility cost of enrolling 100 kilometers away from home. In exchange, beneficiaries must either enter the labor market in their home province or repay the subsidy.[31]

[30] In contrast, students from higher education hubs are now even more likely to graduate with a bachelor's degree: even those crowded out of their local programs face strong incentives to pursue their studies and return, due to high local returns to education.

[31] Similar programs exist in Chile, Colombia, and Singapore, where students receive scholarships to study abroad in exchange for returning to their home country upon graduation or repaying the subsidy. I consider a within-country version of these programs. Anecdotal evidence suggests that some rural French communes

I simulate this policy counterfactual by allowing students from higher education deserts to decide at enrollment whether to take up the scholarship when enrolling in a program located more than 100 kilometers from home. If they do so, the continuation value associated with that program incorporates the repayment condition.

Table 5: Policy Counterfactual

Panel A: Policy Effects (relative to status quo)	
Number High-Skilled from Deserts	+3.29%
Number High-Skilled Working in Deserts	+16.04%
Sum Wages in Deserts (M€)	+11.92%
Panel B: Policy Outcomes	
Share of Scholarship Recipients (from Deserts)	0.55
Share of Scholarship Recipients Who Repay	0.42
Cost (Share of Total Wages)	2.97%

NOTES. Panel A presents results from the policy counterfactual, expressed in percentage change relative to the status quo. Panel B reports key statistics from the policy, namely the share of students from deserts taking the scholarship, the share repaying it, and the total cost.

Table 5 summarizes the effects of this policy. About half of the students from higher education deserts choose to take up the scholarship. This allows them to enroll in programs located farther from home, increasing by 1.5 percentage points (3.29%) the share graduating with at least a bachelor's degree and reducing spatial gaps in educational attainment by 26%. Because the policy incentivizes recipients to enter their home labor market upon graduation, this initial effect translates into a 16% increase in the number of high-skilled workers in low-opportunity areas, raising aggregate wages of entry-level workers in those provinces by 12%. Since 42% of scholarship recipients prefer to repay the subsidy, the cost of the policy remains low: it represents less than 3% of the cohort's total wages. Overall, mobility scholarships tied to return incentives would allow low-opportunity areas to leverage regional trade in education services to invest in their local human capital.

8 Conclusion

To what extent does the geography of the higher education market contribute to spatial inequalities across students and local labor markets? Using data on the universe of college applicants and higher education programs in France, I document that higher education op-

implement similar policies, for example to secure the presence of a general practitioner in their area.

tions are unevenly distributed across space and that students' demand is highly sensitive to distance. To quantify the role of these features in shaping spatial inequalities, I build and estimate a dynamic model linking equilibrium sorting in the higher education market to location choices at labor market entry. I find that the interaction between mobility frictions and the uneven spatial distribution of higher education programs accounts for a substantial share of spatial inequalities in educational attainment and welfare across students. It also shapes the long-run spatial distribution of labor supply. On the one hand, mobility frictions within the higher education market contribute to spatial skill sorting by limiting investment in higher education among individuals from low-opportunity areas. On the other hand, they insulate these areas from an outflow of workers: relaxing mobility frictions increases the number of individuals from low-opportunity areas entering the labor market in higher education hubs, thereby widening inequalities across local labor markets. Low-opportunity areas could invest in their local human capital by offering mobility scholarships to college applicants tied to incentives to return home at labor market entry.

References

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak, "The welfare effects of coordinated assignment: Evidence from the New York City high school match," *American Economic Review*, 2017, *107* (12), 3635–3689.

_ , Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters, "Do parents value school effectiveness?," *American Economic Review*, 2020, *110* (5), 1502–1539.

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

_ and _ , "Demand Analysis under Latent Choice Constraints," *The Review of Economic Studies*, 2025.

Azevedo, Eduardo M and Jacob D Leshno, "A supply and demand framework for two-sided matching markets," *Journal of Political Economy*, 2016, *124* (5), 1235–1268.

Bechichi, Nagui, Julien Grenet, and Georgia Thebault, "Ségrégation à l'entrée des études supérieures en France et en région parisienne: quels effets du passage à Parcoursup?," 2021.

Calsamiglia, Caterina, Chao Fu, and Maia Güell, "Structural estimation of a model of school choices: The boston mechanism versus its alternatives," *Journal of Political Economy*, 2020, *128* (2), 642–680.

Card, David, Jesse Rothstein, and Moises Yi, "Location, location, location," *American Economic Journal: Applied Economics*, 2025, *17* (1), 297–336.

Chetty, Raj and Nathaniel Hendren, "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects," *The Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

_ , John N Friedman, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter, "The opportunity atlas: Mapping the childhood roots of social mobility," *American Economic Review*, 2026, *116* (1), 1–51.

_ , Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Chyn, Eric and Lawrence F Katz, "Neighborhoods matter: Assessing the evidence for place effects," *Journal of Economic Perspectives*, 2021, *35* (4), 197–222.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "Spatial wage disparities: Sorting matters!," *Journal of urban economics*, 2008, *63* (2), 723–742.

_ , _ , _ , and Sébastien Roux, "Sorting and local wage and skill distributions in France," *Regional Science and Urban Economics*, 2012, *42* (6), 913–930.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Matching in cities," *Journal of the European Economic Association*, 2022, *20* (4), 1478–1521.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000," *American Economic Review*, 2016, *106* (3), 479–524.

_ **and Cecile Gaubert**, "Spatial sorting and inequality," *Annual Review of Economics*, 2022, *14*, 795–819.

Epple, Dennis, Richard Romano, and Holger Sieg, "Admission, tuition, and financial aid policies in the market for higher education," *Econometrica*, 2006, *74* (4), 885–928.

Fabre, Anaïs, "Robustness of Two-Way Fixed Effects Estimators to Heterogeneous Treatment Effects," *TSE Working Paper*, 2025.

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

Fu, Chao, "Equilibrium tuition, applications, admissions, and enrollment in the college market," *Journal of Political Economy*, 2014, *122* (2), 225–281.

_ **, Junjie Guo, Adam J Smith, and Alan Sorensen**, "Students' heterogeneous preferences and the uneven spatial distribution of colleges," *Journal of Monetary Economics*, 2022, *129*, 49–64.

Gale, David and Lloyd S Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, 1962, *69* (1), 9–15.

Gibbons, Stephen and Anna Vignoles, "Geography, choice and participation in higher education in England," *Regional science and urban economics*, 2012, *42* (1-2), 98–113.

Gourieroux, Christian and Alain Monfort, *Simulation-based econometric methods*, Oxford university press, 1996.

He, YingHua, Shruti Sinha, and Xiaoting Sun, "Identification and Estimation in Many-to-One Two-Sided Matching Without Transfers," *Econometrica*, 2024, *92* (3), 749–774.

Hillman, Nicholas W, "Geography of college opportunity: The case of education deserts," *American Educational Research Journal*, 2016, *53* (4), 987–1021.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow, "The allocation of talent and us economic growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Ishimaru, Shoya, "Geographic Mobility of Youth and Spatial Gaps in Local College and Labor Market Opportunities," *Journal of Labor Economics*, 2025, *43* (4), 1251–1294.

Kapor, Adam, "Transparency and percent plans," *Econometrica*, 2025, *93* (6), 2123–2157.

_ , Mohit Karnani, and Christopher Neilson, "Aftermarket frictions and the cost of off-platform options in centralized assignment mechanisms," *Journal of Political Economy*, 2024, *132* (7), 2346–2395.

Katz, Lawrence F and Kevin M Murphy, "Changes in relative wages, 1963–1987: supply and demand factors," *The quarterly journal of economics*, 1992, *107* (1), 35–78.

Kelchtermans, Stijn and Frank Verboven, "Program duplication in higher education is not necessarily bad," *Journal of Public Economics*, 2010, *94* (5-6), 397–409.

Kennan, John and James R Walker, "The effect of expected income on individual migration decisions," *Econometrica*, 2011, *79* (1), 211–251.

Kim, Kimin and Myoung jae Lee, "Difference in differences in reverse," *Empirical Economics*, 2019, *57* (3), 705–725.

Olden, Andreas and Jarle Møen, "The triple difference estimator," *The Econometrics Journal*, 2022, *25* (3), 531–553.

Silva, JMC Santos and Silvana Tenreyro, "The log of gravity," *The Review of Economics and statistics*, 2006, *88* (4), 641–658.

Train, Kenneth E, *Discrete choice methods with simulation*, Cambridge university press, 2009.

Appendices

A Additional tables and figures

Table A.1: Summary Statistics: Students

	Mean	St. Dev.
	(1)	(2)
Female	0.5397	0.4984
Socio-Economic Characteristics		
With Scholarship	0.1938	0.3953
High SES	0.3138	0.4640
Medium High SES	0.1565	0.3633
Medium Low SES	0.3120	0.4633
Low SES	0.2177	0.4127
End-of-High School Exam		
With Highest Honors	0.0945	0.2925
With High Honors	0.1740	0.3791
With Honors	0.2962	0.4566
Pass	0.4353	0.4958
High School Track		
General Track	0.6247	0.4842
Technological Track	0.2224	0.4159
Vocational Track	0.1529	0.3599
Observations	507,605	

NOTES. This table presents summary statistics for college applicants in 2019. I restrict attention to high school students living in metropolitan France.

Table A.2: Summary Statistics: Programs

	Mean (1)	St. Dev. (2)
Panel A: Program Type and Major		
Three-Year University	0.2232	0.4164
Humanities	0.4146	0.4928
Economics - Law	0.1639	0.3702
Social Sciences	0.1942	0.3957
STEM	0.2274	0.4192
Two-Year Elite	0.0743	0.2622
STEM	0.5259	0.4996
Economics - Law	0.3140	0.4644
Humanities	0.1601	0.3669
Engineering/Business	0.0716	0.2578
Engineering	0.5160	0.5001
Business	0.0613	0.2400
Arts & Architecture	0.4227	0.4943
Two-Year Technical	0.0710	0.2569
Production	0.5997	0.4903
Services	0.4003	0.4903
Two-Year Vocational	0.4397	0.4964
Production	0.3498	0.4769
Services	0.5492	0.4976
Agriculture	0.1011	0.3014
Nursing & Social Work	0.0519	0.2219
Nursing	0.6391	0.4807
Social Work	0.3609	0.4807
Other	0.0682	0.2521
Panel B: Other Characteristics		
Number Seats	61.1547	113.0219
With Dorm	0.0572	0.2321
Observations	10,936	

NOTES. This table presents summary statistics of programs available on the platform in 2019, located in metropolitan France.

Table A.3: Summary Statistics: Applications, Offers, and Enrollment

	Mean	St. Dev.	25^{th} Pctile	50^{th} Pctile	75^{th} Pctile
	(1)	(2)	(3)	(4)	(5)
Number of Applications	10.035	9.076	5.000	8.000	13.000
Number of Offers	3.745	4.108	1.000	3.000	5.000
Enrollment Distance	65.374	109.777	8.809	25.984	73.001

NOTES. This table presents summary statistics of the applications, offers, and enrollment of students in high school in 2019, living in metropolitan France. Characteristics of this sample are summarized in Table A.1.

Table A.4: Summary Statistics: Enrollment by Program Type

	Mean	St. Dev.
	(1)	(2)
Three-Year University	0.439	0.496
Two-Year Elite	0.078	0.268
Engineering/Business	0.055	0.228
Two-Year Technical	0.108	0.311
Two-Year Vocational	0.256	0.436
Nursing	0.040	0.196
Other	0.024	0.154
Unmatched	0.193	0.395
Observations	507,605	

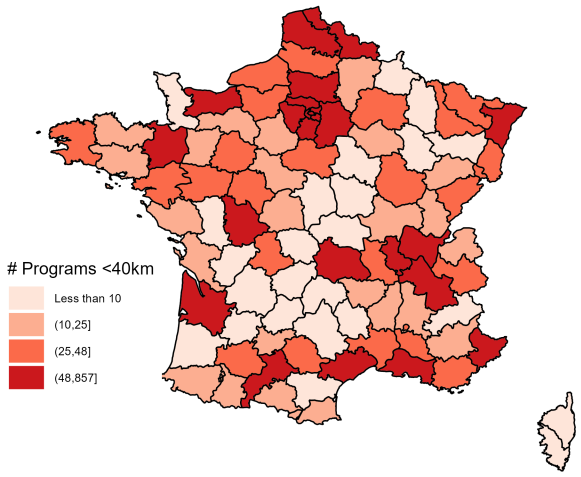
NOTES. This table presents summary statistics of the program of enrollment of students in high school in 2019, living in metropolitan France. Characteristics of this sample are summarized in Table A.1.

Table A.5: Summary Statistics: Entry-Level Workers

	Mean	St. Dev.	Min	Max
	(1)	(2)	(3)	(4)
Panel A: Work and Study Locations				
Study Home	0.46			
Work Home	0.61			
Work and Study Home	0.37			
Work Home if Studied Home	0.79			
If Studied Outside Home Location				
Work Home	0.44			
Work in Study Location	0.17			
Work Elsewhere	0.38			
Panel B: Earnings				
Monthly Earnings	1,627	819	0	20,000
Number Individuals	38,621			

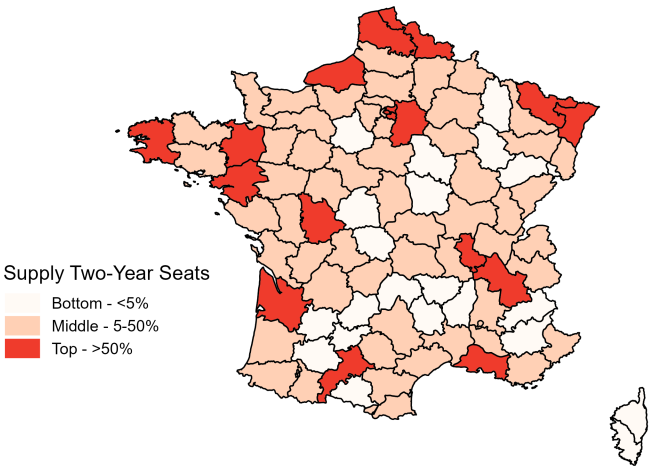
NOTES. This table provides summary statistics for the study and work locations of individuals who exited the education system, as well as for their earnings. Locations are defined at the province level. Study Home (Work Home, respectively) refers to the share of individuals who studied (work, respectively) in their home province. Work and Study Home refers to the share of individuals who both studied and work in their home province. The table then restricts the sample to individuals who studied outside their home province and reports the share who returned home to work, who stayed in their study location, or who entered the labor market outside both their study and home provinces (Work Elsewhere).

Figure A.1: Number of Academic Programs Within Commuting Distance



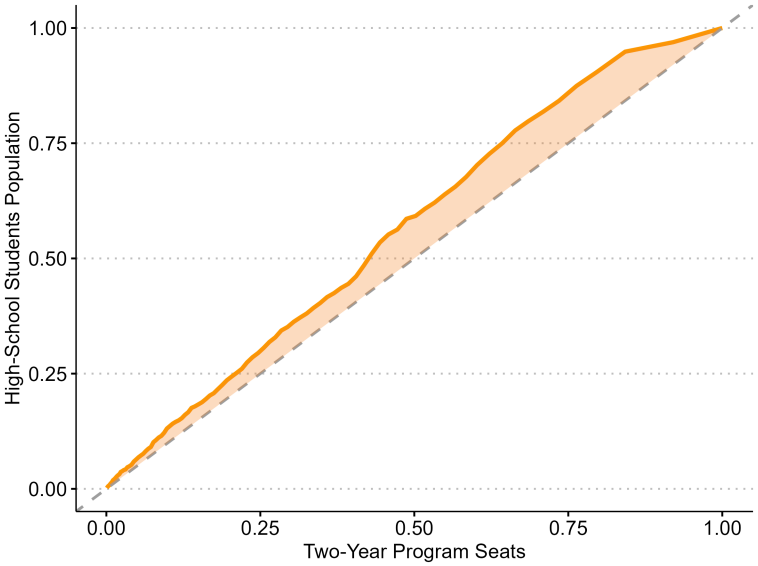
NOTES. This map plots the number of academic programs the average student of a given province has access to within a 40-kilometer radius from her home.

Figure A.2: Spatial Distribution of Seats in Two-Year Technological and Vocational Programs



NOTES. This figure shows the distribution of seats in two-year technological and vocational programs at the province level. Provinces are ranked according to their position in the distribution of these seats. Provinces in white are at the bottom of this distribution: they jointly account for less than 5% of the total supply. In contrast, provinces in red are at the top of this distribution: together, they account for half of the total supply of seats in two-year technological and vocational programs.

Figure A.3: Cumulative Distribution of Seats in Two-Year Technological and Vocational Programs & Student Population



NOTES. This figure plots the relationship between the spatial distributions of seats in two-year technological and vocational programs and of high school students. It depicts the cumulative distribution of the supply of these seats across provinces on the x-axis and the corresponding cumulative distribution of the high school student population on the y-axis. Provinces are ranked according to the share of vocational and technological seats they offer.

Table A.6: Minimum Distance to Programs

	Quantile			
	95^{th}	90^{th}	75^{th}	50^{th}
	(1)	(2)	(3)	(4)
Three-Year University Programs				
Law	62	51	33	15
Economics	74	62	41	18
Medicine	109	93	63	31
Sports	77	64	43	19
Mathematics	91	74	46	19
Physics & Chemistry	97	78	48	20
Two-Year Elite Programs				
Sciences	58	46	27	11
Humanities	85	69	43	17
Economics & Business	67	56	37	15
Two-Year Technical Programs				
Legal Careers	256	210	150	74
Marketing	74	63	42	18
Chemistry	189	144	94	52
IT	98	82	56	27
Two-Year Vocational Programs				
Real Estate	110	93	61	29
Accounting	39	31	19	8
Construction	86	74	51	24
Engineering Schools	70	57	35	14
Management Schools	107	91	61	28
Nursing Schools	33	26	17	7

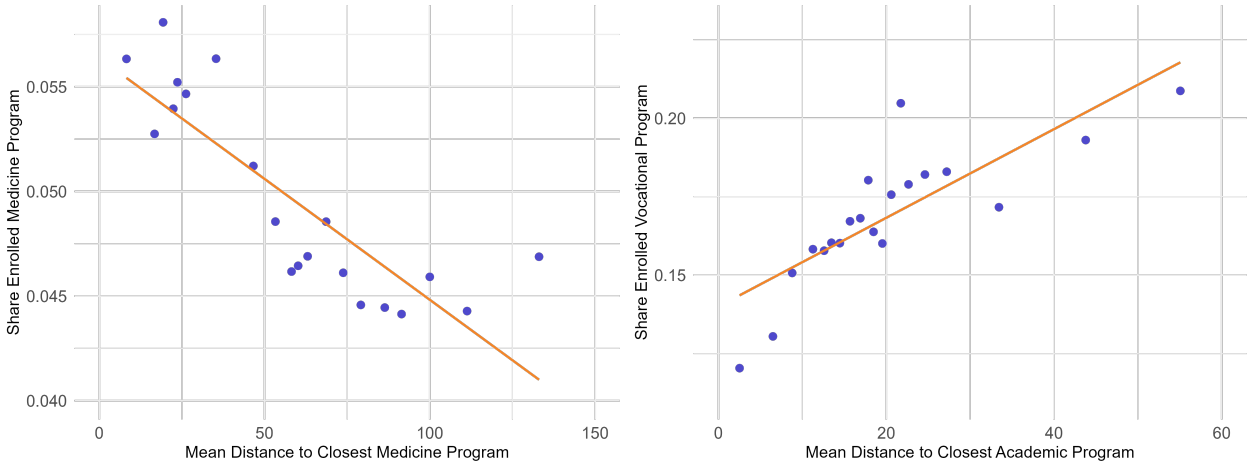
NOTES. Minimum distance to travel, by program type. For example, 10% of students must travel at least 93 kilometers to reach a medicine program.

Table A.7: Gravity Model

	Enrollment Flows			
	OLS	OLS	PPML	PPML
	(1)	(2)	(3)	(4)
$\text{Distance}_{o,d}$	-1.917	-1.910	-2.483	-1.856
	(0.022)	(0.021)	(0.043)	(0.130)
Origin FE	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes
Origin-Destination Pairs	8,930	9,025	8,930	9,025
Adjusted/Pseudo R-Squared	0.752	0.794	0.873	0.925

NOTES. This table presents estimates of a gravity equation. It regresses students' enrollment flows between origin o and destination province d on the logarithm of distance (in kilometers) between the centroids of o and d, controlling for origin and destination fixed effects. Columns (1) and (2) show results from an estimation using OLS, while Columns (3) and (4) present results using PPML, as recommended by Silva and Tenreyro (2006). Columns (1) and (3) omit observations for which $o = d$, while Columns (2) and (4) include a dummy equal to one when $o \neq d$ as control variable. Robust standard errors are shown in parentheses.

Figure A.4: Higher Education Market Geography and Spatial Inequalities



a) Enrollment in medicine programs **b) Enrollment in two-year vocational programs**

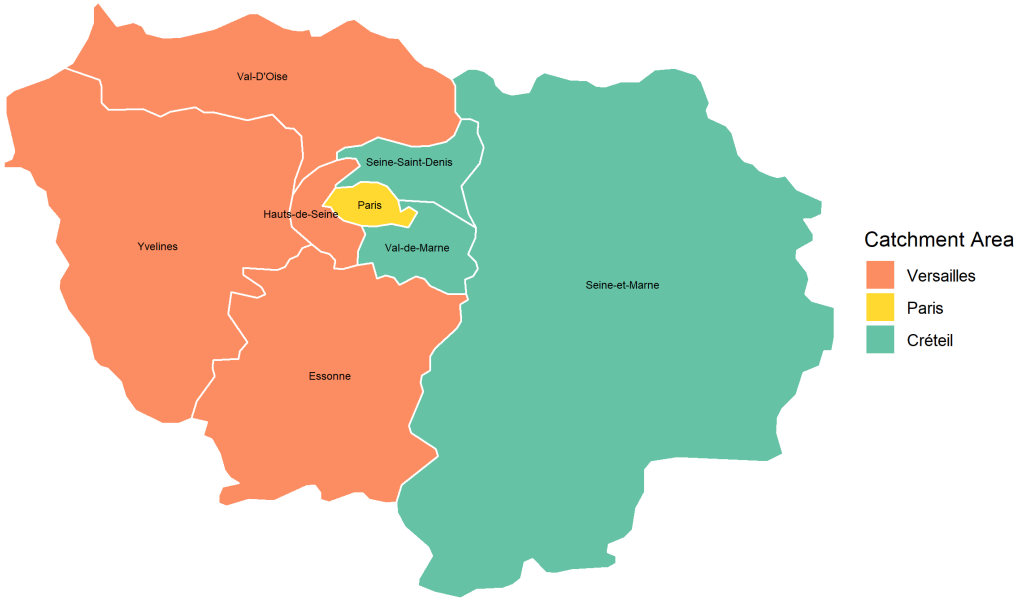
NOTES. Panel a) is a binned scatter plot, at the province level, of the relationship between the share of students from a province enrolled in a medicine program and the mean distance to the closest medicine program for the average student of the province. Panel b) is a binned scatter plot, at the province level, of the relationship between the share of students from a province enrolled in a two-year vocational program and the mean distance to the closest academic program for the average student of the province.

Table A.8: Effect of the Mobility Scholarship

		Out-of-Catchment Area	
	Distance	Applications	Offers
	(1)	(2)	(3)
β_{2019}	9.018	0.024	0.009
	(2.568)	(0.008)	(0.006)
β_{2017}	2.031	-0.000	-
	(2.729)	(0.009)	-
β_{2016}	-5.251	-0.027	-
	(2.826)	(0.009)	-
β_{2015}	-4.301	-0.009	-
	(2.959)	(0.010)	-
β_{2014}	-1.925	-0.020	-
	(2.949)	(0.010)	-
Controls	Yes	Yes	Yes
Observations	4,519,699	4,521,568	2,041,859

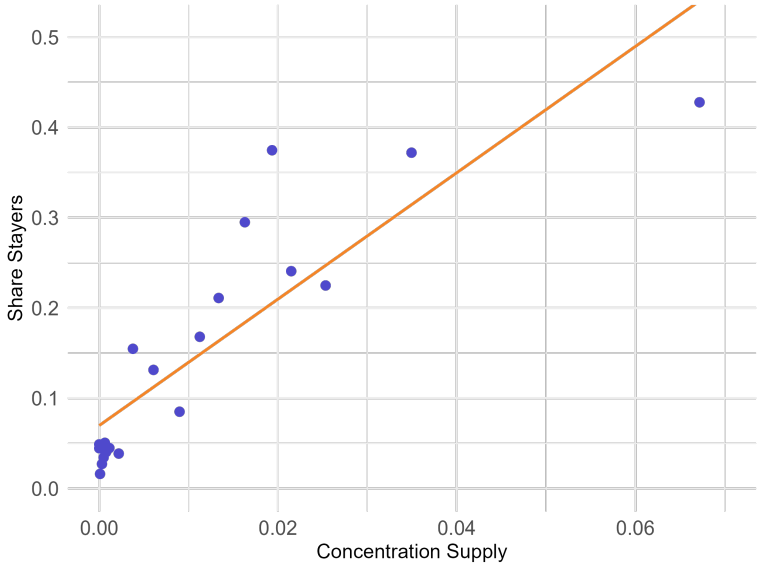
NOTES. This table presents parameter estimates of $\beta_{2018+\tau}$ from Equation (1). It shows the effect of the mobility subsidy, implemented in 2019, on the mean application distance (Column (1), see also Figure 4), on the probability to submit an out-of-catchment area application (Column (2)), and to receive an out-of-catchment area offer (Column (3)). Results shown correspond to defining low-supply areas as the five locations at the bottom of the distribution of higher education seats. Controls include honors on the end-of-high-school exam, high school track, province of origin, household income, and the number of applications submitted interacted with the year of application. Changes in the system implied that students could receive multiple offers only from 2018 onward. Effects are robust to extending this definition to the ten locations at the bottom of this distribution, although the magnitude of the estimates is slightly smaller. High-supply areas are defined as the ten locations at the top of this distribution. Results are robust to excluding Paris from the sample. Robust standard errors are shown in parentheses.

Figure A.5: Map of the Catchment Areas - Greater Paris Area



NOTES. The Greater Paris Area is composed of three catchment areas: Versailles, Paris, and Créteil.

Figure A.6: Contribution of the Student Population to the Non-Local Labor Force



NOTES. This binned scatter plot shows the relationship between the concentration of academic seat supply and provinces' ability to retain graduates. The x-axis reports a province's share of academic seats. The y-axis reports the share of the province's non-local labor force that graduated there.

Table A.9: Summary Statistics - Greater Paris Area Programs

	Catchment Area			
	Paris	Versailles	Créteil	Greater Paris Area
University				
Number University Programs	167	92	81	340
Number University/Branches	10	8	7	25
Number Ranked Applicants	1,356	1,306	1,231	1,313
% High-Ability Overall	27.17	14.38	8.33	19.22
% High-Ability from Paris	5.81	0.92	0.48	3.22
% High-Ability from Versailles	4.92	11.40	0.94	5.73
% High-Ability from Créteil	4.82	1.31	5.35	4.00
Non-University				
Number Programs	437	499	446	1,382
Number Institutions	113	186	158	457
Number Ranked Applicants	442	326	301	355
% High-Ability Overall	48.32	29.24	21.41	32.75
% High-Ability from Paris	13.62	2.21	1.20	5.49
% High-Ability from Versailles	10.94	19.85	3.38	11.72
% High-Ability from Créteil	6.92	2.28	14.01	7.53

NOTES. Descriptive statistics of the programs offered in the Greater Paris area in 2018. The shares correspond to the mean percentage of high-ability students overall and from Paris, Versailles and Créteil admitted by each type of programs.

Table A.10: Institutional Barriers to Mobility: Greater Paris Area

	Program's Catchment Area		
	Paris	Versailles	Créteil
All Students			
From Versailles	-483	197	-180
From Créteil	-157	-373	171
From Paris	943	-400	-111
High-Ability Students			
From Versailles	-807	73	-200
From Créteil	-470	-472	91
From Paris	407	-522	-93

NOTES. Average difference between the rank of an applicant before and after the implementation of geographic priorities in 2018. On average, an application from a student in Versailles to a university program located in Paris loses 483 ranks after the implementation of geographic priorities.

Table A.11: Removing Barriers to Mobility: Application Effects

	Share High-Ability		Share Low-Ability
	from Versailles	from Créteil	from Paris
	(1)	(2)	(3)
β_{Paris}	0.008	0.003	-0.005
	(0.003)	(0.002)	(0.004)
$\beta_{\text{Versailles}}$	-0.016	0.004	0.008
	(0.003)	(0.001)	(0.001)
$\beta_{\text{Créteil}}$	0.004	-0.009	0.007
	(0.001)	(0.002)	(0.002)
Baseline Paris	0.069	0.046	0.162
Baseline Versailles	0.147	0.019	0.047
Baseline Créteil	0.026	0.077	0.062
Program FE	Yes	Yes	Yes
Observations	680	680	680

NOTES. This table presents estimates from Equation (2). I study the effect of removing barriers to mobility on the share of high-ability students from Versailles (Column (1)), the share of high-ability students from Créteil (Column (2)), and the share of low-ability students from Paris (Column (3)) applying to university programs located in Paris, Versailles, and Créteil. Standard errors clustered at the program level are shown in parentheses.

Table A.12: Removing Barriers to Mobility: Sorting Effects

	Share High-Ability		Share Low-Ability
	from Versailles	from Créteil	from Paris
	(1)	(2)	(3)
$\beta^{DD}_{\text{Paris}}$	0.037	0.014	-0.078
	(0.007)	(0.008)	(0.017)
$\beta^{DD}_{\text{Versailles}}$	-0.005	0.005	0.023
	(0.010)	(0.005)	(0.006)
$\beta^{DD}_{\text{Créteil}}$	0.010	0.001	0.023
	(0.004)	(0.007)	(0.007)
Baseline Paris	0.049	0.048	0.213
Baseline Versailles	0.114	0.013	0.029
Baseline Créteil	0.009	0.054	0.046
Program FE	Yes	Yes	Yes
Observations	3,444	3,444	3,444

NOTES. This table presents estimates from Equation (16). I study the effect of removing barriers to mobility on the share of high-ability students from Versailles (Column (1)), the share of high-ability students from Créteil (Column (2)), and the share of low-ability students from Paris (Column (3)) enrolled in programs located in Paris, Versailles, and Créteil. Standard errors clustered at the program level are shown in parentheses.

Table A.13: Removing Barriers to Mobility: Application Effects

	Share High-Ability		Share Low-Ability
	from Versailles	from Créteil	from Paris
	(1)	(2)	(3)
$\beta^{DD}_{\text{Paris}}$	0.018	0.001	-0.015
	(0.002)	(0.002)	(0.003)
$\beta^{DD}_{\text{Versailles}}$	-0.006	0.004	0.004
	(0.003)	(0.001)	(0.002)
$\beta^{DD}_{\text{Créteil}}$	0.005	-0.008	0.004
	(0.001)	(0.002)	(0.002)
Baseline Paris	0.069	0.046	0.162
Baseline Versailles	0.147	0.019	0.047
Baseline Créteil	0.026	0.077	0.062
Program FE	Yes	Yes	Yes
Observations	3,444	3,444	3,444

NOTES. This table presents estimates from Equation (16). I study the effect of removing barriers to mobility on the share of high-ability students from Versailles (Column (1)), the share of high-ability students from Créteil (Column (2)), and the share of low-ability students from Paris (Column (3)) applying to university programs located in Paris, Versailles, and Créteil. Standard errors clustered at the program level are shown in parentheses.

Table A.14: Students' Preferences - Parameter Estimates

		Heterogeneity								
		Female	Medium High SES	Medium Low SES	Low SES	Schol-arship	High Honors	Honors	Pass	Voc/Tech
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
University	-1.949	0.248	0.726	1.001	1.015	0.484	0.725	0.855	0.697	-2.205
	(0.166)	(0.087)	(0.136)	(0.136)	(0.128)	(0.121)	(0.121)	(0.173)	(0.165)	(0.094)
Two-Year Elite	0.480	-0.896	0.563	0.895	0.531	-0.003	-0.114	-0.499	-1.807	-2.146
	(0.186)	(0.151)	(0.215)	(0.215)	(0.284)	(0.272)	(0.272)	(0.219)	(0.286)	(0.260)
Two-Year Technical	-2.326	-0.612	0.675	0.920	0.691	0.434	1.322	1.539	1.174	-0.402
	(0.279)	(0.126)	(0.191)	(0.191)	(0.187)	(0.165)	(0.165)	(0.271)	(0.271)	(0.130)
Nursing Schools	-6.390	2.340	1.001	1.163	1.077	0.419	1.434	1.895	1.657	0.926
	(0.574)	(0.344)	(0.316)	(0.316)	(0.299)	(0.247)	(0.247)	(0.429)	(0.446)	(0.224)
Two-Year Vocational	-5.080	-0.412	1.023	1.408	1.445	0.298	1.097	1.693	1.419	1.236
	(0.309)	(0.097)	(0.164)	(0.164)	(0.146)	(0.125)	(0.125)	(0.298)	(0.294)	(0.115)
Engineering/Business	-0.243	-1.103	0.384	0.551	0.614	-0.255	0.787	0.559	-0.152	-2.196
	(0.234)	(0.170)	(0.251)	(0.251)	(0.275)	(0.310)	(0.310)	(0.253)	(0.259)	(0.230)
Other	-3.364	0.088	1.371	1.261	1.212	0.683	-0.186	0.918	0.244	-0.227
	(0.501)	(0.234)	(0.375)	(0.375)	(0.368)	(0.288)	(0.288)	(0.542)	(0.520)	(0.250)

NOTES. Standard errors are shown in parentheses. Parameters correspond to $\frac{1}{\sigma}u(x_i, z_j)$. The estimated parameter for σ can be found in Table A.15.

Table A.15: Students' Preferences - Mobility Frictions

		Heterogeneity				
		Female	Medium High SES	Medium Low SES	Low SES	Scholarship
	(1)	(2)	(3)	(4)	(5)	(6)
Distance (100km)	2.086	-0.004	0.538	0.761	1.039	0.209
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Distance^2	-0.166	0.003	-0.034	-0.054	-0.078	-0.011
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
σ^{-1}	3.330					
	(0.220)					

NOTES. Standard errors are shown in parentheses. Parameters correspond to $\frac{1}{\sigma}c(x_i, d_{ij})$.

Table A.16: Students' Preferences - Compensation to Enroll 100km Away from Home

	Average	Female	High SES	Low SES	Scholarship
Compensation (in €/month)	741.3	749.0	579.3	886.4	856.3

NOTES. This table presents the monthly compensation, in euros, required for students to be willing to enroll in a program located 100 kilometers away from their home location. It is computed using the parameter estimates presented in Table A.15. Note that the continuation value is expressed in monthly euros, while its associated parameter is normalized to one in $\mathcal{U}_{ij}$. Parameters of $u(.)$ and $c(.)$ can thus be interpreted in monthly euros.

Table A.17: Programs' Preferences - Parameter Estimates

	University	Two-Year Elite	Two-Year Technical	Nursing	Two-Year Vocational	Engineering / Business	Other
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	-3.213	-6.741	-4.743	-6.489	-2.224	-4.750	-1.641
	(0.062)	(0.081)	(0.063)	(0.063)	(0.015)	(0.049)	(0.049)
GPA	0.512	0.628	0.473	0.527	0.331	0.379	0.251
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female	0.219	0.111	0.256	0.132	0.232	0.216	0.200
	(0.004)	(0.006)	(0.005)	(0.005)	(0.003)	(0.007)	(0.007)
Technological Track	-1.057	-1.031	-0.951	-0.364	-0.495	0.009	-0.476
	(0.008)	(0.051)	(0.075)	(0.075)	(0.006)	(0.009)	(0.009)
Vocational Track	-1.195	-1.784	-1.737	-1.172	-1.047	-0.599	-0.630
	(0.014)	(0.162)	(0.018)	(0.018)	(0.011)	(0.015)	(0.015)
Medium-High SES	0.031	-0.005	0.049	-0.084	0.033	0.040	0.003
	(0.005)	(0.008)	(0.006)	(0.006)	(0.004)	(0.009)	(0.009)
Medium-Low SES	0.008	0.034	0.033	-0.103	0.036	0.036	-0.019
	(0.004)	(0.007)	(0.005)	(0.005)	(0.003)	(0.008)	(0.008)
Low SES	0.032	0.080	0.039	-0.140	0.032	0.047	-0.035
	(0.005)	(0.010)	(0.006)	(0.006)	(0.003)	(0.010)	(0.010)
With Scholarship	0.420	0.399	0.244	0.164	0.064	0.172	0.096
	(0.005)	(0.009)	(0.006)	(0.006)	(0.003)	(0.009)	(0.009)
Same Catchment Area	0.313						
	(0.004)						
High School with HE Program		0.131			0.001		
		(0.006)			(0.003)		
Same High School		0.452			0.362		
		(0.012)			(0.004)		

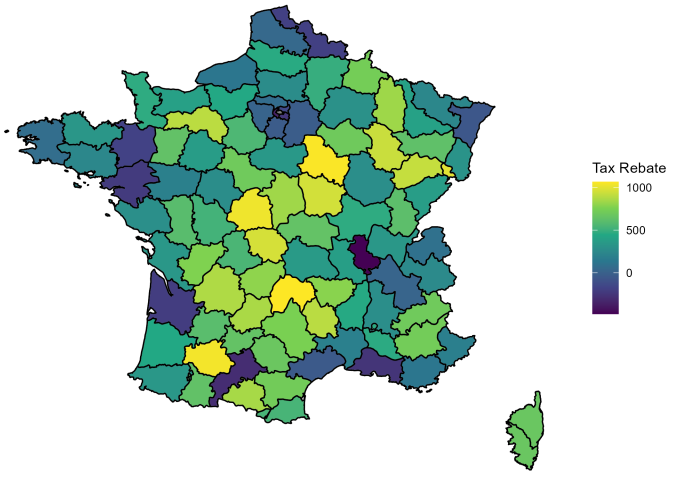
NOTES. Standard errors are shown in parentheses.

Table A.18: Preferences over Local Labor Markets

Home Bias		
High School Diploma × Home Location	5.699	(0.045)
Two-Year Degree × Home Location	5.105	(0.030)
Bachelor's Degree and More × Home Location	4.000	(0.023)
Moving Costs		
High School Diploma × Graduation Location	3.479	(0.064)
Two-Year Degree × Graduation Location	3.261	(0.043)
Bachelor's Degree and More × Graduation Location	2.847	(0.027)
High School Diploma × Home Location × Graduation Location	-3.066	(0.100)
Two-Year Degree × Home Location × Graduation Location	-2.556	(0.072)
Bachelor's Degree and More × Home Location × Graduation	-1.976	(0.053)
Scale		
θ	1.494	(0.090)
Location FE	Yes	
Observations	3,668,900	

NOTES. Standard errors are shown in parentheses.

Figure A.7: Location Choices - Moving Costs



NOTES. This map uses estimated preferences over local labor markets to compute the tax rebate, in monthly euros, an individual from a given province would need to receive to be willing to move back home rather than stay in Paris, conditional on having studied in Paris and having obtained at least a bachelor's degree. It abstracts from wage differences across locations.

Table A.19: Transition Probabilities

	Vocational Degree	Bachelor's Degree or More
	(1)	(2)
Entry Location	4.928 (0.003)	
Program Choice		
University	-0.006 (0.027)	2.945 (0.023)
Two-Year Elite	0.731 (0.044)	4.468 (0.037)
Two-Year Technical	1.903 (0.031)	3.872 (0.028)
Nursing	1.480 (0.034)	1.287 (0.043)
Two-Year Vocational	1.706 (0.029)	2.813 (0.026)
Engineering/Business	1.265 (0.044)	4.500 (0.039)
Other	0.292 (0.038)	2.533 (0.037)
Outside Option	-0.915 (0.028)	1.666 (0.023)
Baseline Characteristics		
Female	0.100 (0.009)	0.149 (0.009)
With Scholarship	0.024 (0.011)	-0.114 (0.011)
Medium-High SES	0.068 (0.014)	-0.259 (0.013)
Medium-Low SES	-0.012 (0.012)	-0.372 (0.011)
Low SES	0.065 (0.013)	-0.432 (0.013)
High Honors	0.211 (0.028)	-0.369 (0.024)
Honors	-0.043 (0.026)	-1.052 (0.022)
Pass	-0.512 (0.026)	-2.115 (0.021)
Technical Track	-0.400 (0.011)	-1.623 (0.011)
Vocational Track	-1.262 (0.014)	-3.291 (0.019)
Location FE	Yes	
Observations	159,656,175	

NOTES. Standard errors are shown in parentheses.

Table A.20: Model Fit

	Data	Model
	(1)	(2)
Panel A: Higher Education Market		
University	0.482	0.491
Two-Year Elite	0.091	0.080
Two-Year Technical	0.103	0.105
Two-Year Vocational	0.216	0.211
Nursing School	0.026	0.033
Engineering/Business	0.057	0.058
Distance	67.411	65.025
Panel B: Work Location		
Same as Home & Study Location	0.366	0.367
Same as Home Location Only	0.240	0.239
Same as Study Location Only	0.092	0.092
Neither	0.302	0.302

NOTES. This table evaluates the fit of the model by comparing statistics related to the match on the higher education market (Panel A) and to location choices at labor market entry (Panel B). Column (1) and Column (2) present the statistics observed in the data and in the simulated status quo, respectively.

Table A.21: Model Fit: Gravity Model

	Enrollment Flows			
	Data		Model	
	(1)	(2)	(3)	(4)
Distance$_{o,d}$	-2.483	-1.856	-2.314	-1.907
	(0.043)	(0.130)	(0.040)	(0.148)
Origin FE	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes
Origin-Destination Pairs	8,930	9,025	8,930	9,025

NOTES. This table presents estimates of a gravity equation. It regresses students' enrollment flows between origin o and destination province d on the logarithm of distance (in kilometers) between the centroids of o and d, controlling for origin and destination fixed effects. Columns (1) and (2) (Columns (3) and (4), respectively) show results from an estimation using the observed (simulated, respectively) enrollment flows of students in the estimation sample. Estimation is performed using PPML, as recommended by Silva and Tenreyro (2006). Columns (1) and (3) omit observations for which $o = d$, while Columns (2) and (4) include a dummy equal to one when $o \neq d$ as control variable. Robust standard errors are shown in parentheses.

Table A.22: Quantification - Worker Flows

	Counterfactual	Change
	(1)	(2)
Panel A: No Mobility Frictions		
Work Home	0.53	-10.76%
From Desert - Work outside Desert	0.50	8.95%
Panel B: No Mobility Frictions + No Admission Restrictions		
Work Home	0.51	-15.57%
Panel C: No Mobility Frictions + No Moving Costs		
Work Home	0.67	11.34%

NOTES. This table presents changes in the share of workers entering the labor force in their home province and in the share of workers from higher education deserts entering the labor market in a province that is not a higher education desert. Panel A presents results from a counterfactual in which mobility frictions within the higher education market are removed. Panel B presents results from a counterfactual additionally eliminating admission restrictions: all students are admitted to their preferred programs in stage 1. Panel C presents results from a counterfactual in which both mobility frictions within the higher education market and moving costs between the study and home location in the second stage are removed.

Table A.23: Quantification - Labor Force in Paris

	Status Quo	No Mobility Frictions	No Admission Restrictions	No Moving Costs
	(1)	(2)	(3)	(4)
Share Work Paris	0.09	+ 12.74%	+ 34.43%	- 5.33%

NOTES. This table presents the share of workers entering the labor force in Paris in the status quo (Column (1)) and the corresponding changes under three counterfactuals. Column (2) considers the counterfactual in which mobility frictions within the higher education market are removed. Column (3) focuses on the counterfactual in which all students are admitted to their most preferred program. Column (4) considers the counterfactual in which both mobility frictions within the higher education market and moving costs between the study and home location in the second stage are removed.

B Context: additional details

In the French high school system, students have three main educational tracks to choose from: general, technological, and vocational.

In the general track, students can choose among three broad fields: sciences, social sciences, or humanities. While students in the general track share common classes, such as English, they also have specialized courses tailored to their chosen field. For instance, there is a strong emphasis on mathematics, biology, and physics for students in the sciences field. The technological track focuses on the application of scientific and technological knowledge to practical skills. Students can specialize in areas such as information technology, chemistry, or agronomy. Finally, the vocational track is designed to provide students with practical skills, with a strong emphasis on internships. They can specialize in various areas, such as automotive technology, hospitality, construction, agriculture, or healthcare.

A student's track and specialization determine the centralized exam (*baccalauréat*) they take at the end of high school. For example, students in the science track within the general track take tests specifically designed to evaluate their knowledge of mathematics, biology, and physics. The final score ranges from 0 to 20 and is associated with an honors system (*mentions*). The passing score is 10. A score between 12 and 14 corresponds to graduating with honors (*Mention Assez Bien*). A score between 14 and 16 corresponds to graduating with high honors (*Mention Bien*). A student graduates with highest honors (*Mention Très Bien*) with a score of 16 or above. Table A.1 summarizes the share of students receiving each honors category, conditional on passing.

C Proofs

C.1 Evaluation of the mobility scholarship: identifying assumptions

The identifying assumption for the triple-difference specification used in Section 3.2.1 is:

$$\begin{aligned} &E\left[\left(Y_{L,E,t}(0)-Y_{L,E,t-1}(0)\right)-\left(Y_{L,NE,t}(0)-Y_{L,NE,t-1}(0)\right)\right] \\ =&E\left[\left(Y_{H,E,t}(0)-Y_{H,E,t-1}(0)\right)-\left(Y_{H,NE,t}(0)-Y_{H,NE,t-1}(0)\right)\right], \end{aligned}$$

where $Y_{s,e,t}(0)$ denotes the average potential outcome of students with eligibility status $e \in \{E, NE\}$ living in provinces with supply level $s \in \{H, L\}$ in period t when not treated. Under this parallel trends assumption, which can be tested, β_{2019} identifies the following object:

$$\begin{aligned} &E\left[[Y_{L,E,2019}(1) - Y_{L,NE,2019}(0)] - [Y_{H,E,2019}(1) - Y_{H,NE,2019}(0)]\right] \\ &= ATT_{L,E,2019} - ATT_{H,E,2019}. \end{aligned} \tag{15}$$

That is, it identifies the difference in treatment effects of the policy in low-supply versus high-supply areas. However, it seems reasonable to assume that the reform should not have affected the application behavior of scholarship-eligible students living in high-supply areas. The scholarship amount was small, while these students are low-income students living in places where they can easily access higher education programs. The policy is therefore unlikely to have an effect on students already living in a high-supply area, so the ATT is likely to be zero in those locations. In that case, β_{2019} would identify the average treatment effect of the moving subsidy for students living in low-supply areas.

To limit concerns regarding a potential non-null effect of the subsidy on students in the comparison group, I consider areas at the very top and bottom of the academic seat supply distribution. I define high-supply areas as the ten locations at the top of the supply distribution. Low-supply areas are the five locations at the bottom of the supply distribution. The effects displayed in Table A.8 are robust to extending the definition of low-supply areas to the ten locations at the bottom of the supply distribution, although the magnitude of the estimates is slightly smaller.

If the ATT in high-supply areas were non-zero but positive, β_τ would then identify the ATT in low-supply areas net of the one in high-supply areas. As long as the latter is positive, a positive β_τ implies that the treatment effect is positive for students living in low-supply areas, and larger than for those living in high-supply areas.

C.2 Reverse difference-in-differences

Factors other than the reform discussed in Section 3.2.2 may have affected admission outcomes between 2018 and 2019. To control for them, I use as a comparison group a set of programs whose admission criteria remained the same over time: non-university programs,

for which geographic priorities were never implemented. I thus augment the model as follows:

$$y_{j,\ell,t} = \alpha_j + \lambda_{\ell,t} + \sum_{\ell} \beta_{\ell}^{DD} \times D_{\ell,t} + \epsilon_{j,\ell,t}, \tag{16}$$

where $\lambda_{\ell,t}$ captures location-specific trends in admission outcomes common to university and non-university programs. In particular, the specification compares the evolution of admission outcomes in university programs—for which institutional barriers to mobility were removed in 2019—to that in non-university programs, for which such barriers were never implemented.

In standard difference-in-differences, the treatment status of the two groups is the same at baseline and differs after treatment. In contrast, here, the treatment status of the two groups differs at baseline but is the same after treatment. Kim and Lee (2019) and Fabre (2025) show that one can still identify the treatment effect of interest in this framework, under the assumption that the potential outcomes of the two groups would have evolved similarly over time under non-geographic admission criteria.

Note that, given the nature of the market, one of the assumptions required for any difference-in-differences to be valid might not be satisfied in practice: the comparison group may be affected by the reform. Indeed, student flows might respond to the removal of barriers to mobility by shifting away from other types of programs toward university programs located within Paris. To limit such concerns, I implement this difference-in-differences at the catchment area level.[32]

C.3 Identification of preferences over locations

Taking the logarithm on both sides of Equation (10), we have:

$$\log(\Pr(\ell^* = \ell | x_i, h_i, y_i, \boldsymbol{w}_i)) = \theta(w_{i\ell} + \omega_\ell(x_i, h_i, y_i)) - \log(\sum_{\kappa \in \mathcal{L}} \exp(\theta(w_{i\kappa} + \omega_\kappa(x_i, h_i, y_i)))). \tag{17}$$

The same equation can be evaluated for $\ell = \ell_0$, the location choice for which $\omega_{\ell_0}(x_i, h_i, y_i)$ and $w_{i\ell_0}$ are normalized to zero. Taking the difference of Equation (17) evaluated at locations

[32] For example, if an increase in the number of students from Versailles admitted to programs in Paris is compensated by a decrease in the number of students from Versailles admitted to non-university programs in Versailles, the comparison would still be valid. Note that it would, however, be a concern if this increase is rather compensated by a decrease in the number of high-ability students from Versailles admitted to non-university programs in Paris.

ℓ and ℓ_0, we have:

$$\log(\Pr(\ell^* = \ell | x_i, h_i, y_i, \boldsymbol{w}_i)) - \log(\Pr(\ell^* = \ell_0 | x_i, h_i, y_i, \boldsymbol{w}_i)) = \theta(w_{i\ell} + \omega_\ell(x_i, h_i, y_i)),$$

where the left-hand side of the equation is directly observed in data from realized location choices. In particular, taking the derivative of this object with respect to $w_{i\ell}$ allows us to identify θ. Given θ, ω can then be identified.

C.4 Identification of program-specific admission priorities

Taking the logarithm on both sides of Equation (11), we have:

$$\log(\Pr(i^*_{j,r} = i | \boldsymbol{x}, \boldsymbol{h}, z_j)) = v(x_i, h_i, z_j) - \log\Big(\sum_{a \in \mathcal{A}_j \setminus \{i^*_{j,k}\}_{k=1}^{r-1}} \exp(v(x_a, h_a, z_j))\Big). \tag{18}$$

The same equation can be evaluated for $i = 0$, the program's outside option, for which the program-specific score is normalized to zero. Taking the difference between Equation (18) evaluated for student i and 0, we have:

$$\begin{aligned}
&\log(\Pr(i^*_{j,r} = i | \boldsymbol{x}, \boldsymbol{h}, z_j)) - \log(\Pr(i^*_{j,r} = 0 | \boldsymbol{x}, \boldsymbol{h}, z_j)) \\
&= v(x_i, h_i, z_j) - v(x_0, h_0, z_j) \\
&= v(x_i, h_i, z_j),
\end{aligned}$$

where the left-hand side of the equation is directly observed in data from realized rankings of applicants, such that $v(.)$ can be recovered.

C.5 Identification of σ

This section provides the proof of identification of σ. Consider two students, i and o, with $x_i = x_o$ and $h_i \neq h_o$. Consider a program j such that $u(x_i, z_j) - c(x_i, d_{ij}) = u(x_o, z_j) - c(x_o, d_{oj})$.

Consider first student i. Taking the ratio of the probabilities associated with enrolling in

program j and in the outside option, we obtain:

$$\begin{aligned}
&\frac{\Pr(\mu(i)=j|x_i,h_i,\boldsymbol{z})}{\Pr(j\in\mathcal{F}(x_i,h_i,\boldsymbol{z},\boldsymbol{\eta})|x_i,h_i,\boldsymbol{z})\Pr(\mu(i)=0|x_i,h_i,\boldsymbol{z})}\\
=&\frac{\int \frac{\exp(\frac{1}{\sigma}(u(x_i,z_j)-c(x_i,d_{ij})+\bar{U}(x_i,h_i,z_j)))}{\sum_{k\in\mathcal{F}(x_i,h_i,\boldsymbol{z},\boldsymbol{\eta})\bigcup 0}\exp(\frac{1}{\sigma}(u(x_i,z_k)-c(x_i,d_{ik})+\bar{U}(x_i,h_i,z_k)))}f_\eta(\boldsymbol{\eta})d\boldsymbol{\eta}}{\int \frac{\exp(\frac{1}{\sigma}(u(x_i,z_0)-c(x_i,d_{i0})+\bar{U}(x_i,h_i,z_0)))}{\sum_{k\in\mathcal{F}(x_i,h_i,\boldsymbol{z},\boldsymbol{\eta})\bigcup 0}\exp(\frac{1}{\sigma}(u(x_i,z_k)-c(x_i,d_{ik})+\bar{U}(x_i,h_i,z_k)))}f_\eta(\boldsymbol{\eta})d\boldsymbol{\eta}}\\
=&\frac{\exp(\frac{1}{\sigma}(u(x_i,z_j)-c(x_i,d_{ij})+\bar{U}(x_i,h_i,z_j)))}{\exp(\frac{1}{\sigma}(u(x_i,z_0)-c(x_i,d_{i0})+\bar{U}(x_i,h_i,z_0)))}.
\end{aligned} \tag{19}$$

Taking logarithms and subtracting the same expression evaluated at (x_o,h_o), we obtain:

$$\begin{aligned}
&\log\left(\frac{\Pr(\mu(i)=j|x_i,h_i,\boldsymbol{z})}{\Pr(j\in\mathcal{F}(x_i,h_i,\boldsymbol{z},\boldsymbol{\eta})|x_i,h_i,\boldsymbol{z})\Pr(\mu(i)=0|x_i,h_i,\boldsymbol{z})}\right)\\
-&\log\left(\frac{\Pr(\mu(o)=j|x_o,h_o,\boldsymbol{z})}{\Pr(j\in\mathcal{F}(x_o,h_o,\boldsymbol{z},\boldsymbol{\eta})|x_o,h_o,\boldsymbol{z})\Pr(\mu(o)=0|x_o,h_o,\boldsymbol{z})}\right)\\
=&\frac{1}{\sigma}(u(x_i,z_j)-c(x_i,d_{ij})-(u(x_o,z_j)-c(x_o,d_{oj})))\\
&+\frac{1}{\sigma}\left[\left[\bar{U}(x_i,h_i,z_j)-\bar{U}(x_i,h_i,z_0)\right]-\left[\bar{U}(x_o,h_o,z_j)-\bar{U}(x_o,h_o,z_0)\right]\right]\\
=&\frac{1}{\sigma}\left[\left[\bar{U}(x_i,h_i,z_j)-\bar{U}(x_i,h_i,z_0)\right]-\left[\bar{U}(x_o,h_o,z_j)-\bar{U}(x_o,h_o,z_0)\right]\right],
\end{aligned}$$

where the first equality follows from the fact that $u(x_i,z_0)-c(x_i,d_{i0})=0,\forall x_i$, and the second equality from the fact that $u(x_i,z_j)-c(x_i,d_{ij})=u(x_o,z_j)-c(x_o,d_{oj})$. $\bar{U}(.)$ can be identified from the parameters identified from data on location choices and transition probabilities. We can thus express σ as a function of objects observed in the data and parameters previously identified, as long as $\bar{U}(x_i,h_i,z_j)-\bar{U}(x_i,h_i,z_0)\neq\bar{U}(x_o,h_o,z_j)-\bar{U}(x_o,h_o,z_0)$. This is made plausible by the fact that the home location, h_i, enters $\bar{U}(.)$, while students i and o have different home locations.

Note that I include the distance between student i and program j in $c(.)$. The identification argument thus relies on considering a program j located at the same distance from the home locations of students i and o, possibly in the home province of one of them. The flow utility associated with a program can also depend freely on the location of the program, for example through location fixed effects.

The identification argument would also hold by using two different programs, j' and j'', as long as $u(x_i,z_{j'})-c(x_i,d_{ij'})=u(x_o,z_{j''})-c(x_o,d_{oj''})$ and $\bar{U}(x_i,h_i,z_{j'})-\bar{U}(x_i,h_i,z_0)\neq$

$\bar{U}(x_o, h_o, z_{j''}) - \bar{U}(x_o, h_o, z_0)$. In particular, this requires the characteristics of programs j' and j'' to be the same, i.e. $z_{j'} = z_{j''}$. It also requires the distance between student i and program j' to be equal to that between student o and program j''.

C.6 Identification of $u(.)$ and $c(.)$

Taking the ratio of $\Pr(\mu(i) = j|x_i, h_i, \boldsymbol{z})$ and $\Pr(\mu(i) = 0|x_i, h_i, \boldsymbol{z})$, we have:

$$\begin{aligned}
&\frac{\Pr(\mu(i) = j|x_i, h_i, \boldsymbol{z})}{\Pr(j \in \mathcal{F}(x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta})|x_i, h_i, \boldsymbol{z}) \Pr(\mu(i) = 0|x_i, h_i, \boldsymbol{z})} \\
=&\frac{\int \frac{\exp(\frac{1}{\sigma}(u(x_i,z_j)-c(x_i,d_{ij})+\bar{U}(x_i,h_i,z_j)))}{\sum_{k\in\mathcal{F}(x_i,h_i,\boldsymbol{z},\boldsymbol{\eta})\bigcup 0}\exp(\frac{1}{\sigma}(u(x_i,z_k)-c(x_i,d_{ik})+\bar{U}(x_i,h_i,z_k)))} f_\eta(\boldsymbol{\eta})d\boldsymbol{\eta}}{\int \frac{\exp(\frac{1}{\sigma}(u(x_i,z_0)-c(x_i,d_{i0})+\bar{U}(x_i,h_i,z_0)))}{\sum_{k\in\mathcal{F}(x_i,h_i,\boldsymbol{z},\boldsymbol{\eta})\bigcup 0}\exp(\frac{1}{\sigma}(u(x_i,z_k)-c(x_i,d_{ik})+\bar{U}(x_i,h_i,z_k)))} f_\eta(\boldsymbol{\eta})d\boldsymbol{\eta}} \\
=&\frac{\exp(\frac{1}{\sigma}(u(x_i, z_j) - c(x_i, d_{ij}) + \bar{U}(x_i, h_i, z_j)))}{\exp(\frac{1}{\sigma}(u(x_i, z_0) - c(x_i, d_{i0}) + \bar{U}(x_i, h_i, z_0)))}.
\end{aligned} \tag{20}$$

Taking logarithms, this writes:

$$\begin{aligned}
&\log(\Pr(\mu(i) = j|x_i, h_i, \boldsymbol{z})) - \log(\Pr(\mu(i) = 0|x_i, h_i, \boldsymbol{z}) \Pr(j \in \mathcal{F}(x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta})|x_i, h_i, \boldsymbol{z})) \\
&= \frac{1}{\sigma}\left[u(x_i, z_j) - c(x_i, d_{ij}) + \bar{U}(x_i, h_i, z_j)\right] - \frac{1}{\sigma}\left[u(x_i, z_0) - c(x_i, d_{i0}) + \bar{U}(x_i, h_i, z_0)\right] \\
&\iff \sigma\left[\log(\Pr(\mu(i) = j|x_i, h_i, \boldsymbol{z})) - \log(\Pr(\mu(i) = 0|x_i, h_i, \boldsymbol{z}) \Pr(j \in \mathcal{F}(x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta})|x_i, h_i, \boldsymbol{z}))\right] \\
&- \left[\bar{U}(x_i, h_i, z_j) - \bar{U}(x_i, h_i, z_0)\right] \\
&= u(x_i, z_j) - c(x_i, d_{ij}),
\end{aligned} \tag{21}$$

where $\Pr(\mu(i) = j|x_i, h_i, \boldsymbol{z})$ and $\Pr(\mu(i) = 0|x_i, h_i, \boldsymbol{z})$ are directly identified from data on realized matches, while $\bar{U}(x_i, h_i, z_j) - \bar{U}(x_i, h_i, z_0)$ and σ have been identified in a first step. $\Pr(j \in \mathcal{F}(x_i, h_i, \boldsymbol{z}, \boldsymbol{\eta})|x_i, h_i, \boldsymbol{z})$ is also identified as we identified programs' preferences, while we know the set of applicants to each program and the distribution of η. Thus, the left-hand side of this equation is observed. Normalizing $c(x_i, d_{ij}) = 0$ when $d_{ij} = 0$ ensures that we can separately identify $u(.)$ and $c(.)$.

D Estimation details

D.1 Estimation sample

For computational tractability, students' preferences are estimated using a random subsample of 5,159 students. Tables D.1-D.3 present the same summary statistics of students' characteristics and their enrollment outcomes as in Table A.1 and Tables A.3-A.4. All characteristics are very similar in the two samples.

Table D.1: Summary Statistics: Students

	Mean (1)	St. Dev. (2)
Female	0.5427	0.4982
Socio-Economic Characteristics		
With Scholarship	0.1939	0.3959
High SES	0.3117	0.4633
Medium High SES	0.1520	0.3591
Medium Low SES	0.3160	0.4650
Low SES	0.2202	0.4144
End-of-High School Exam		
With Highest Honors	0.0962	0.2948
With High Honors	0.1578	0.3646
With Honors	0.3054	0.4606
Pass	0.4407	0.4965
High School Track		
General Track	0.6190	0.4857
Technological Track	0.2208	0.4148
Vocational Track	0.1601	0.3668
Observations	5,159	

NOTES. This table presents summary statistics of college applicants in 2019, for the estimation sample.

D.2 Estimation of transition probabilities

I use the universe of applicants to the online platform in 2019, regardless of their match status, and match them to the higher education census using each student's unique identifier. I am able to match 82% of students. Among students matched to a program on Parcoursup in 2019, 94% are recovered in the higher education census dataset. This implies that 6% of the students who are matched to an on-platform program either do not enroll in their matched

Table D.2: Summary Statistics: Applications, Offers, and Enrollment

	Mean	St. Dev.	25^{th} Pctile	50^{th} Pctile	75^{th} Pctile
	(1)	(2)	(3)	(4)	(5)
Number of Applications	10.052	9.175	5.000	8.000	12.000
Number of Offers	3.650	3.830	1.000	3.000	5.000
Enrollment Distance	65.352	110.875	8.878	25.987	71.645

NOTES. This table presents summary statistics of the applications, offers, and enrollment of students in high school in 2019, living in metropolitan France, for the estimation sample. Characteristics of the full sample are summarized in Table A.1.

Table D.3: Summary Statistics: Program Type Enrollment

	Mean	St. Dev.
	(1)	(2)
Three-Year University	0.449	0.497
Two-Year Elite	0.074	0.262
Engineering/Business	0.051	0.219
Two-Year Technical	0.105	0.307
Two-Year Vocational	0.261	0.439
Nursing	0.038	0.192
Other	0.022	0.146
Unmatched	0.192	0.394
Observations	5,159	

NOTES. This table presents summary statistics of the program of enrollment of students in high school in 2019, living in metropolitan France, for the estimation sample. Characteristics of the full sample are summarized in Table A.1.

program, or do so but cannot be recovered in the census.

Note that the higher education census also allows me to track some unmatched students, as I observe whether they (i) match to one of the few programs off-platform but present in the census or (ii) match to an on-platform program in the following years, implying they participated again in the admission process. Given that the administrative census gathers the large majority of higher education programs in France, I consider that a student who was unmatched on the platform in 2019 and who cannot be found in the 2019–2024 censuses remained outside the higher education system. I assign the home location of these students as their graduation location.

E Counterfactuals: details

E.1 Counterfactuals: simulation details

This section describes how both the higher education market equilibrium and individuals' location choices are computed in counterfactual exercises.

Higher education market equilibrium I use the estimated parameters and a draw of idiosyncratic shocks to simulate students' utility associated with each program, from Equation (3), as well as their program-specific admission score, from Equation (6). In particular, when removing mobility frictions within the higher education system, the utility of student i from enrolling in program j becomes:

$$\mathcal{U}_{ij}^{\mathcal{C}^1} = u(x_i, z_j) + \sigma\epsilon_{ij} + \bar{U}(x_i, h_i, z_j). \tag{22}$$

This allows me to generate each student's ranking of programs, as well as each program's ranking of students. I feed these two rankings, together with programs' number of seats, into a student-proposing Deferred Acceptance algorithm, which matches each student to a unique program. This algorithm generates the student-proposing stable match. The properties of this market entail that it essentially has a unique stable equilibrium.

Higher education market equilibrium - Change in second stage payoffs Note that some counterfactual exercises directly modify payoffs in the second stage of the model

— such as the counterfactual in which mobility costs between the individual's graduation and home location are eliminated. In such cases, one needs to update the continuation value associated with each program before computing students' utility from Equation (3). I therefore start by computing the new utility from entering each location ℓ in stage 2 from Equation (5), before using it to compute the log-sum formula, which is used to derive the continuation value. Taking the example of the counterfactual where mobility costs between the individual's graduation and home location are eliminated, denoted $\mathcal{C}^2$, the new utility from entering the labor market in location ℓ for individual i writes:

$$\mathcal{W}_{i\ell}^{\mathcal{C}^2} = w_{i\ell} + \omega_\ell(x_i, h_i, y_i) + g_i'\boldsymbol{\phi}_3 \mathbb{1}\{\ell = h_i, h_i \neq \ell_i^G\} + \frac{1}{\theta}\nu_{i\ell}, \tag{23}$$

where the term $g_i^{'}\boldsymbol{\phi}_3 \mathbb{1}\{\ell = h_i, h_i \neq \ell_i^G\}$ is added to remove moving costs between the graduation location of individual i, ℓ_i^G, and her home location, h_i. The continuation value of a program with characteristics z_j now becomes:

$$\bar{U}^{\mathcal{C}^2}(x_i, h_i, z_j) = \int \beta^{t(y_i)} \mathbb{E}\left[\max_{\ell \in \mathcal{L}} \mathcal{W}_{i\ell}^{\mathcal{C}^2} | x_i, h_i, y_i\right] f\left(y_i | x_i, h_i, z_j\right) dy_i. \tag{24}$$

In this case, we can then derive each student's ranking of programs by computing the following object, for each i:

$$\mathcal{U}_{ij}^{\mathcal{C}^2} = u(x_i, z_j) + \sigma\epsilon_{ij} + \bar{U}^{\mathcal{C}^2}(x_i, h_i, z_j).$$

Higher education market equilibrium - No admission restrictions One of the counterfactual exercises eliminates mobility frictions, while preventing programs from rejecting any students. In this case, each student can be admitted to their preferred program, regardless of capacity constraints and programs' admission decisions. The equilibrium is such that each student i is matched to a program $\mu(i)$, such that:

$$\mu(i) = \arg\max_{j \in \mathcal{J} \bigcup 0} \mathcal{U}_{ij}^{\mathcal{C}^1}.$$

Location Decisions at Entry on the Labor Market Once the new higher education market equilibrium is simulated, one can compute the counterfactual transition probabilities, and use Equation (5) to simulate the counterfactual location choice of individual i.

Note that all idiosyncratic shocks are kept fixed across the different simulation exercises

and are the same as those used to simulate the status quo paths.

E.2 Counterfactuals: outcomes

Inequalities between students with different home locations Each counterfactual exercise is evaluated by comparing outcomes between the new equilibrium and the status quo simulation. In order to capture inequalities across individuals with different home locations, I compute the share of individuals from a given province graduating with at least a bachelor's degree as follows:

$$s_\ell^{Home} = \frac{\sum_i \mathbb{1}\{h_i = \ell, g_i = \text{Bachelor}\}}{\sum_i \mathbb{1}\{h_i = \ell\}},$$

where $\mathbb{1}\{h_i = \ell, g_i = \text{Bachelor}\}$ is a dummy equal to one if ℓ is the home location of individual i and individual i graduates with at least a bachelor's degree. $\mathbb{1}\{h_i = \ell\}$ is a dummy equal to one if ℓ is the home location of individual i. I therefore have 96 observations of the variable s_ℓ^{Home}, one for each province, and I compute the variance of these values.

In addition to the variance, I also compare the share of individuals from higher education deserts and hubs graduating with at least a bachelor's degree as follows:

$$\frac{\sum_{i\in\{i:h_i\in Hubs\}} \mathbb{1}\{g_i = \text{Bachelor}\}}{n_{Hubs}} - \frac{\sum_{i\in\{i:h_i\in Deserts\}} \mathbb{1}\{g_i = \text{Bachelor}\}}{n_{Deserts}},$$

where $\mathbb{1}\{g_i = \text{Bachelor}\}$ is a dummy equal to one if individual i graduates with at least a bachelor's degree, *Hubs* is the set of higher education hubs, *Deserts* is the set of higher education deserts, n_{Hubs} is the number of individuals from higher education hubs, and $n_{Deserts}$ is the number of individuals from higher education deserts.

I also compute the mean welfare of students in different provinces as follows:

$$\bar{W}_\ell = \frac{\sum_{i\in\{i:h_i=\ell\}} \mathcal{U}_{ij^*}}{\sum_i \mathbb{1}\{h_i = \ell\}},$$

where j^* corresponds to the program to which student i is matched in the scenario considered. I also compare the mean welfare of students from hubs and deserts:

$$\frac{\sum_{i\in\{i:h_i\in Hubs\}} \mathcal{U}_{ij^*}}{n_{Hubs}} - \frac{\sum_{i\in\{i:h_i\in Deserts\}} \mathcal{U}_{ij^*}}{n_{Deserts}}.$$

Inequalities across local labor markets In order to capture inequalities across local labor markets, I first look at the composition of their labor force. In particular, I compute the share of individuals who graduated with at least a bachelor's degree and who entered the labor market in province ℓ, that is:

$$s_{\ell}^{Work} = \frac{\sum_i \mathbb{1}\{\ell_i^* = \ell, g_i = \text{Bachelor}\}}{\sum_i \mathbb{1}\{\ell_i^* = \ell\}},$$

where $\mathbb{1}\{\ell_i^* = \ell, g_i = \text{Bachelor}\}$ is a dummy equal to one if ℓ is the work location of individual i and individual i graduates with at least a bachelor's degree. $\mathbb{1}\{\ell_i^* = \ell\}$ is a dummy equal to one if ℓ is the work location of individual i. To measure inequalities across local labor markets, I compute the variance of s_{ℓ}^{Work} across the 96 provinces of metropolitan France. Additionally, I compare the share of workers in hubs and deserts as follows:

$$\frac{\sum_{i \in \{i:\ell_i^* \in Hubs\}} \mathbb{1}\{g_i = \text{Bachelor}\}}{L_{Hubs}} - \frac{\sum_{i \in \{i:\ell_i^* \in Deserts\}} \mathbb{1}\{g_i = \text{Bachelor}\}}{L_{Deserts}},$$

where L_{Hubs} is the number of entry-level workers entering the labor market in a higher education hub, and $L_{Deserts}$ is the number of entry-level workers entering the labor market in a higher education desert.

The second measure I use to capture inequalities across local labor markets is the aggregate wages of entry-level workers in a given location:

$$Y_{\ell} = \sum_{i \in \{i:\ell_i^* = \ell\}} w_{i\ell}.$$

I then compare aggregate wages of entry-level workers in hubs and deserts:

$$\sum_{\ell \in Hubs} Y_{\ell} - \sum_{\ell \in Deserts} Y_{\ell}.$$