amse
école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

Like Father, Like Child: Intergenerational Mobility in the French Grandes Écoles throughout the 20 th Century

Stéphane Benveniste

WP 2023 Nr 18

Like Father, Like Child: Intergenerational Mobility in the French Grandes Écoles throughout the 20TH Century *

Stéphane Benveniste ♣

September 2023

Abstract

While the educational expansion of the 20th century promoted social mobility overall, the top of the social hierarchy may have remained privileged. This paper examines the evolution of intergenerational mobility in admissions to the French elite colleges—the *Grandes Écoles* (GE)—over more than a century. Admission to these institutions is subject to partially anonymous competitive examinations, and their degrees are the ticket to top positions in the public and private sectors. In the growing literature measuring intergenerational mobility through surnames, I design a novel method and apply it to a self-collected dataset on all 285,286 graduates from ten of the most prestigious *Grandes Écoles* between 1886 and 2015. Principally, I find that children of male GE graduates were highly over-represented in the top colleges throughout the 20th century. Importantly, unlike previous studies exploiting fathers' socio-professional categories, I find a stable low level of intergenerational mobility for all cohorts born since 1916: chances of GE admission for children of GE graduates were approximately 80 times higher than for the rest of the population.

JEL Classification: J62, I23, N34.
Keywords: Intergenerational mobility, Higher education, Elites, *Grandes Écoles*, Historical economics.

*I am grateful for helpful comments and suggestions by Renaud, Bourles, Yann Bramoullé, Paolo Brunori, Baptiste Coulmont, Ewen Gallic, Arnaud Lefranc, Thomas Piketty, Morgan Raux, Lorenzo Rotunno, Akiko Suwa-Eisenmann, Alain Trannoy, Louis-André Vallet, as well as many seminar participants. I would also like to thank all those in archive departments and alumni associations that helped me collect data from numerous sources and whose help made this study possible. This work was supported by the French National Research Agency Grant ANR-17-EURE-0020, by the Excellence Initiative of Aix-Marseille University – A*MIDEX, and by the NORFACE transnational research program on the Dynamics of Inequality Across the Life-course.

♣Institut national d'études démographiques (INED), F-75020 Paris, France; Aix-Marseille University (Aix-Marseille School of Economics), CNRS, Marseille, France. Mail: stephane.benveniste@ined.fr.

"One is a *Normalien*, as one is a prince by blood."[1]
Georges Pompidou, Prime Minister

1 Introduction

Education was proclaimed the "great equalizer of conditions of men, the balance wheel of the social machinery" by Horace Mann in 1848. The educational systems of developed countries have since experienced massive expansion, with a documented decrease in the influence of parental socio-economic characteristics on educational attainment over the 20th century (Breen and Müller, 2020). However, Erikson and Goldthorpe (1992) argued that this "neglects the degree to which those in advantaged positions can secure similar positions for their children." Using data from the last two decades, Chetty et al. (2020) find that admissions to US elite colleges are more likely for children of affluent families, while Chetty et al. (2023) show that graduating from an elite institution has causal impacts on the probability of joining a prestigious firm and of reaching the top 1% of the earnings distribution.

This paper contributes to the debate on elite education as an equalizer in the French context. Exploring a wider historical period, I develop a novel surname-based method that leverages the entirety of the dataset instead of a limited sub-sample based on a subjective cutoff for rare surnames. I apply this new method to investigate long-term intergenerational mobility at the most prestigious French higher education institutions—the *Grandes Écoles* (GE). One feature that contrasts strongly with colleges in the English-speaking world is that tuition fees are negligible in most *Grandes Écoles*. Moreover, there is no system of legacy preferences and very limited (not race-based) affirmative action policies for admission to these colleges. Instead, admission is based on a competitive *concours*, consisting of anonymous written tests followed by oral evaluations. While it is commonly argued that the *concours* system selects candidates purely on their abilities and qualifications, theoretically promoting meritocratic admissions, critics argue that it may inadvertently perpetuate existing social inequalities. Factors such as disparities in educational resources, socioeconomic background,

[1] A *Normalien* is a student from *École Normale Supérieure*, one of the oldest *Grandes Écoles*. Quoted in Peyrefitte (1964), translated by Suleiman (1978).

and access to preparatory resources can influence performance in the *concours*, thereby affecting the extent to which meritocracy is truly achieved.

To examine trends in mobility over more than a century, I construct an original dataset of graduates that includes the exhaustive nominative list of the 285,286 students graduating from ten *Grandes Écoles* between 1886 and 2015.[2] This corresponds to the top 0.36% of the educational distribution in France over the period. The main findings highlight a considerable advantage for sons and daughters of *Grande École* graduates across the period, and a remarkable stability of intergenerational mobility over 80 years.[3] Indeed, in the early 20$^{\text{th}}$ century, someone whose father had graduated from a GE was 154 times more likely to be admitted to the colleges of the baseline sample.[4] This advantage was halved in the following cohort, but remained stable for children of GE graduates born between 1916 and 1995, with 72 to 83 times higher chances of admission.

The present study also reveals a dynastical over-representation of families in the French elite colleges: having a grandfather or a great-grandfather who graduated from a *Grande École* made admission 30 to 54 times more likely over the period. Furthermore, heterogeneity analyses reveal a clear pattern of descendants with considerably higher admission rates to the very colleges their fathers graduated from. For instance, children of "*énarques*" (graduates of the *École nationale d'administration*) were 250 to 330 more likely to become "*énarques*" themselves. These findings are far from trivial considering the student population composition over the last century: the very small number of children of graduates accounted for as much as 13 to 17% of GE students across cohorts, a substantial over-representation. Finally, I characterize the late and slow admission of women, who were granted legal access to all *Grandes Écoles* only in the early 1970s. I show that the prospects of admission to the most prestigious *Grandes Écoles* are of comparable magnitude for sons and daughters of GE

[2]Section 3 provides a description of the dataset, and Appendix C.1 supplements it with information on each college. While there are currently about 500 *Grandes Écoles* in France, the study focuses on a restricted set of 10 schools which play (and have played for the past century) an exceptionally important role in the training of the French elite, as described in section 2 on the institutional context.

[3]The terms *son(s)*, *daughter(s)*, *child(ren)*, *father(s)*, etc. characterize probabilistic genealogical links, formalized in section 4.2. Estimated at the surname-cohort level, these links could alternatively be read as *pseudo-father* and *pseudo-child*.

[4]The baseline consists of all 10 colleges except *Sciences Po Paris*, whose annual student intake is four times that of the second biggest college in the sample. Findings are robust to its inclusion.

graduates. In fact, it appears that social reproduction may be slightly higher from fathers to daughters, suggesting that the increasing presence of women in elite colleges might be associated with higher measures of immobility.

This work is related to the broader literature on intergenerational mobility, in which sociologists and economists have extensively explored how socio-economic outcomes are transmitted across generations (see Black and Devereux, 2011 or Torche, 2015 for comprehensive reviews of the numerous national and transnational studies). My contribution concerns the role of education (the *great equalizer*) in intergenerational mobility. Ganzeboom et al. (1991) argued that if "the main role of education is to promote social mobility; [...] education is also the main vehicle of social reproduction." With the major structural transformations of developed societies over the 20$^{\text{th}}$ century (notably the introduction of welfare regimes and the expansion of educational systems), scholars expected a trend towards equalization. While quantitative democratization—i.e., better access to higher levels of education—has indeed operated to a large extent, the achievement of qualitative democratization—i.e., a reduction of association between social origin and educational outcome—is more debatable.[5] Studying cohorts born between 1906 and 1979 in eight countries including France, Breen and Müller (2020) conclude that "the twentieth century saw both educational expansion and educational equalization". In France, the present paper shows that this equalization trend did not extend to the elite colleges. Specifically focusing on the highest echelons of the educational distribution in both the parents' and the children's generations, it challenges previous findings that primarily relied on parental occupation or socioeconomic category to capture social origin.[6]

A subfield of this literature indeed focuses on the stratification of education, and notably on the top of the educational distribution. General *equalization* does not preclude potential non-linearities of intergenerational mobility across educational levels, especially for the

[5]The influential work of Shavit and Blossfeld (1993) found a stable link between socio-economic origin and the level of education, but most recent studies— including by Shavit himself—challenged this result and reported a qualitative equalization over the 20$^{\text{th}}$ century in developed societies (Shavit et al., 2007; Breen et al., 2009; Breen and Müller, 2020), in France in particular (Thélot and Vallet, 2000; Vallet and Selz, 2008), and specifically for the French *Grandes Écoles* (*cf.* footnote 9).

[6]With a slightly increasing share of the population graduating from the GE (see section 3), this has been more consistent over time than most alternatively used background characteristics (like parental occupation or socio-economic status). Moreover, paternal education in a *Grande École* is more linked to cultural capital, which Thélot and Vallet (2000) argue has more influence on educational attainment than social origin.

more selective tracks. Mare (1980) as well as Shavit and Blossfeld (1993) claimed that social background mainly influenced the early years of education, with much less or no effect for residual transitions to advanced degrees of higher education. While both these studies rely on classifications whose granularity at the top remains limited, this result was corroborated in the US for MBA programs (Stolzenberg, 1994; Mullen et al., 2003), although not for first professional degrees or PhD programs (Mullen et al., 2003; Torche, 2018). More generally, the literature has found substantial inequalities in enrollment in tertiary education tracks, especially the most selective.[7] Yet, focusing on historical trends in 26 European countries, Barone and Ruggera (2018) find that, apart from a stabilization observed for cohorts born after 1965, there was equalization overall in higher education. By contrast, my findings indicate that the dual nature of higher education in France (comparable to the UK and the US, *cf.* contextual section 2) is associated with particularly large magnitudes of intergenerational reproduction in elite colleges, maintained throughout most of the 20$^{\text{th}}$ century.

In France, higher education and especially the *Grandes Écoles* have been the focus of major academic contributions, notably by the sociologist Pierre Bourdieu.[8] Other researchers studied intergenerational mobility in the GE over several decades, generally noting an equalization trend over part of the 20$^{\text{th}}$ century.[9] The present paper is to first to cover such a long time frame, with graduates born between 1866 and 1995, and to systematically in-

[7] Piketty (2020) reports that in the US, the probability of pursuing tertiary education rises quasi-linearly, from 25% to more than 90% for the children of the poorest *vs* richest parents. Chetty et al. (2020) find that offspring with parents in the top 1% *vs* in the bottom quintile of the income distribution are 77 times more likely to enroll in 12 elite institutions including the *Ivy League* colleges. Their over-representation in these institutions remains significant when conditioning on test scores (Chetty et al., 2023). Linking generations via surnames, Clark and Cummins (2014) show that descendants of early 19$^{\text{th}}$ century graduates remain more likely to enroll at *Oxford* and *Cambridge*. Reciprocally, Henderson et al. (2020) find that British students whose parents do not hold a university degree ("first-in-family") have lower chances of being admitted to an elite university.

[8] The seminal *Les Héritiers* by Bourdieu and Passeron (1964) documents social inequalities in access to higher education over the period 1960-1963. Bourdieu and de Saint Martin (1987) survey GE students in 1966-1969 and Bourdieu (1989) adds data on enrollment in 84 GE in 1984-1985, notably concluding that there is a "clear frontier" between GE and university students in inherited *economic* and *cultural capital* and a polarization between the more accessible and the more elitist GE.

[9] Euriat and Thélot (1995) use non-exhaustive data on three prestigious colleges (cohorts 1930s-1970s), while Albouy and Wanecq (2003) focus on 19 top *Grandes Écoles* (cohorts 1919-1968). Both papers document admission inequality in favor of sons of teachers and executives, yet declining over time. Falcon and Bataille (2018) use a much broader definition of the GE (up to 5-7% of the population) and also conclude that cohorts 1918-1984 show a "clear equalization trend". Finally, Bonneau et al. (2021) use administrative data on a shorter and more recent period (births from 1988 to 1995), extending the finding of decreasing admission inequalities to a set of 23 highly selective GE.

vestigate the related intergenerational (and multigenerational) mobility. Instead of a less time-consistent measure of circumstances based on fathers' occupations, I use as a background characteristic the probability that ancestors graduated from an elite college.[10] These innovations yield a series of new results on intergenerational mobility in the French elite colleges, for instance a comparison of qualification reproduction between sons and daughters of graduates over time.

Finally, this paper belongs to the literature taking advantage of the rich informational content of surnames, which can compensate for the scarcity of multigenerational datasets. While surnames do not causally influence outcomes, successive cross-sectional nominative data and the distribution of surnames in the population can inform on the probability both of filiation and of a given achievement in each generation. Working on Catalonian data, Güell et al. (2007, 2015) were the first to investigate intergenerational mobility using surnames—notably followed by Collado et al. (2012), Clark et al. (2014), or Barone and Mocetti (2020). While Clark et al. (2014) claimed that surname-based estimates revealed a "law" of high intergenerational social status persistence, constant over time and across countries, Torche and Corvalan (2018) showed that this was more due to a focus on elite groups—which is also the focus of the present paper. Within this literature on name-based estimates of intergenerational mobility (see Santavirta and Stuhler, forthcoming, for a review), the present paper contributes details of a new method which makes it possible to include the whole population in the analysis. Previous studies usually restricted their samples to rare surnames (e.g., Clark et al., 2014), which had the disadvantage of a subjective and arbitrary choice of threshold of rareness, thus introducing sensitivity to the selected cutoff point. Instead, my novel approach uses as a dependent variable the probability of a given outcome in the previous generation (e.g., having a GE-graduate father) computed at the surname-cohort level. Leveraging the comprehensive dataset, it provides for more robust and inclusive analysis, enabling a more comprehensive understanding of intergenerational mobility.

The rest of the paper is organized as follows. Section 2 describes the *Grandes Écoles* and their central importance in French society. It also provides contextual details on struc-

[10] A prior exception was Le Bras (1983), who categorized students admitted to *Polytechnique* in the single year 1979 with respect to fathers' and grand-fathers' characteristics, including graduation from *Polytechnique*.

tural changes to developed countries' educational systems over the 20th century, particularly in France. Section 3 introduces the data on GEgraduates and provides some descriptive statistics. Section 4 specifies the empirical strategy. I first present the new surname-based methodology, through the construction of the probability of having a GE-graduate ancestor. In this section, I also explain my easy-to-interpret measure of intergenerational mobility: relative admission rates, which provides the "relative risk" of admission of those with a GE-graduate ancestor, compared to the rest of the population. Section 5 provides an extensive set of findings: the baseline results are supplemented by a multigenerational perspective, a heterogeneity analysis across colleges, and a gender decomposition. Finally, section 6 concludes with public policy implications.

2 Context: the *Grandes Écoles*, a stable pillar in the educational transformations of the 20th century

France in feudal times and under the *Ancien Régime* was a monarchy ruled by nepotism and the traditional *three estates*. Social positions were defined by the lottery of birth and individual merit played a marginal role. In 1789, the French Revolution abolished the privileges of the aristocracy and overthrew this system. The 22 pre-existing universities—including *La Sorbonne*—were dismantled because of their link to the clergy and the aristocracy, and two elite higher education institutions were founded in 1794: *École Normale Supérieure* and *École Polytechnique*. With *École des Ponts et chaussées* (1753) and *École des Mines de Paris* (1783), they constituted the very first *Grandes Écoles*. Napoléon Bonaparte was highly involved in structuring the GE, which he viewed as an instrument of rulership, to organize and control the training of teachers, engineers, industrialists, and soldiers. Although the universities progressively regained importance, the dual structure of the French higher education system has remained remarkably stable for the last two centuries.[11]

The *Grandes Écoles* are relatively small top-level tertiary education establishments, with

[11] While higher education is more homogenous in Germany, Spain, Italy, or the Scandinavian countries, the duality of the higher education system, divided between elite institutions and standard universities, is a characteristic of numerous countries including Japan, the US, the UK, or France (Brezis and Hellier, 2018).

usually 50 to 300 students annually. Admission is subject to highly competitive examinations called *concours*, taken after two to three years of a dedicated program of post-secondary school preparation—the *classes préparatoires aux Grandes Écoles*. The *concours* takes place in two successive rounds. The first round consists of anonymous written tests used to rank candidates. Only a small proportion of top performers qualify for the second round, consisting of non-anonymous oral evaluations.[12] While in the United States, individuals may finance prestigious institutions, thereby ensuring their offspring admission (Meer and Rosen, 2009), this *legacy student* system does not exist in France. The race-based affirmative action policies recently ruled out by the US Supreme Court have always been forbidden in France, but other types of policies favoring the disadvantaged target scholarship students or people from poor neighborhoods. These, however, only appeared only in the 21st century and have remained of limited magnitude and impact on the recruitment of students (Allouch, 2022; Bonneau et al., 2021).

The *Grandes Écoles* were deliberately designed for and explicitly dedicated to the education of the elite. Over the last two centuries, they have remained the royal road to decision-making positions in the public and private sectors. They have educated what Suleiman (1978) referred to as *state elites*, "trained by the State and destined for State service", be it within the administration or in national public or private industries. While the lack of professional opportunities for university students was partly responsible for the May 1968 outbreaks, the careers of GE graduates are to some extent guaranteed. Seven out of the eight Presidents of the French Fifth Republic studied at the *Grandes Écoles*. In July 2021, thirty-one out of the thirty-five French CEOs of the CAC40—the forty largest companies on the French stock market—were GE gradtuates, including 10 from *École Polytechnique* alone. The *Grandes Écoles* also shape careers slightly less at the top of the occupational distribution (Delefortrie-Soubeyroux, 1961; Barsoux and Lawrence, 1991; Vion et al., 2014). While there is no historical dataset on wages of graduates of the top GE, numerous sources

[12] Belhoste (2002) presents the evolution over two centuries of the admission examination of *École Polytechnique*, which inspired most other schools. It was designed in 1794 and immediately decentralized to 22 cities across the country. Initially, there were only oral examinations, but written tests were progressively introduced during the 19th century. The *concours* evolved very marginally over the 20th century. These examinations were rapidly adopted in most schools, although slightly later in business schools.

document substantial financial returns on these diplomas.[13]

Many of the most prestigious schools were directly founded at the initiative of the State, except for business schools. Public subsidies historically supported at least partially all *Grandes Écoles*, and it was mainly business schools that were fee-paying until recently. Students at *ENA*, *ENS*, or *Polytechnique* even have the status of civil-servant trainees, receiving financial payment during their education. Thus, although private and public institutions coexist, all schools are to some degree supervised by the State. But they are attached to different ministries and are not centrally supervised by the Ministry of higher education, like universities. This system has promoted the autonomy of the GE and their capacity to resist attempts at reforms (Suleiman, 1978; Pasquali, 2021). Therefore, over the last century, the environment of the *Grandes Écoles* has remained exceptionally stable.

By contrast, there have been substantial educational system transformations in all developed societies, including France. Primary schooling rapidly became universal in the United States in the 19$^{\text{th}}$ century, although it only began to be universal around the 1880s in countries like the United Kingdom, Germany, or France (Piketty, 2020). In France, the Ferry laws made primary schooling free and compulsory in 1882. While in countries like Germany or the Netherlands, a substantial share of the population was already receiving secondary schooling at the beginning of the 20$^{\text{th}}$ century, this remained very marginal in countries such as France, Spain, or Italy, with their dominant agricultural agricultural activities (Breen and Müller, 2020). Although the expansion was slower in France or Spain, secondary education was available to a majority of the population from the second half of the century. The share of the population with tertiary education also increased significantly, especially for cohorts born after World War 2. From a small minority, it progressively reached 30 to 50% of the national populations in the most developed countries, and up to 70% in Japan or Korea.

In France, the surge in tertiary education occurred slightly later. Appendix Figure A.1

[13]The newspaper *L'Expansion* had a notorious tradition of reporting wages of graduates from different GE. Both the *Conférence des grandes écoles* (the official association of GE) and, more recently, the *Financial Times* also undertake such reporting. A broad conclusion is that the median first job wage offered to graduates of top colleges is roughly 2-3 times the median wage in France. Although partly informative, these sources usually only cover the last 30 to 50 years; they also rely on partial response rates from graduates' surveys about the first years of careers, and therefore rarely report exceptionally high wages.

shows that until 1950 less than 5% of the population obtained the *baccalauréat* (secondary school diploma, needed to enter higher education); this share progressively increased, really soaring after the turning point of the 1968 uprisings. Alongside the increasing number of students in the universities, many new *Grandes Écoles* were founded; currently, there are about 500 GE in the French educational system. None of the more recent schools, however, offer any competition on key features—selectivity, graduates' access to elite positions, reputation—with the centuries-old schools examined in the present study (Bourdieu, 1989; Vion et al., 2014). Meanwhile, these long-established *Grandes Écoles* hardly underwent any change at all: for instance, their number of graduates scarcely increased, as shown in section 3.

3 Data: elite college registers and nominative censuses

The paper combines two main types of data: lists of *Grandes Écoles* graduates, and nominative censuses providing the distribution of surnames in France over generations. To my knowledge, this constitutes the first collection of exhaustive lists of graduates of the most prestigious GE. The data covers 10 *Grandes Écoles* over the period 1886-2015 and includes 303,514 curricula of 285,286 distinct students. I directly collected the data from the schools' alumni associations, libraries, archive departments, and from other archival institutions. The 10 colleges are (both historically and currently) among the most prestigious, offering graduates the best career opportunities.[14]

The following provides an overview of the elite colleges in the sample, with further details in Appendix C.1. *Sciences Po Paris* is a school of political science and administration. *École Nationale d'Administration* (*ENA*), founded in 1946, trains senior civil servants. *École Normale Supérieure* (*ENS Ulm*) provides top-level research training in the humanities and science. Five public engineering schools are included: *ESPCI Paris*, *École Polytechnique*,

[14] Appendix C.1 provides supplementary information about each college. The study might have been expected to include *HEC Paris* and *École Centrale Paris*, but these schools refused to share data. Other potential candidates included *École nationale supérieure d'arts et métiers*, *École Nationale Supérieure des Beaux-Arts de Paris*, or military schools, such as *École spéciale militaire de Saint-Cyr*, but these colleges are more specialized, and less representative of the French elite (Suleiman, 1978).

École des Ponts et Chaussées, *Télécom Paris*, and *Mines Paris*. The sample also contains two business schools, for which coverage starts in the early 20th century: *ESSEC* and *ESCP*. These 10 schools are not perfectly homogeneous but can be studied as a whole. Indeed, while Bourdieu (1989) distinguishes between the "intellectual" (e.g., *ENS Ulm*) and the "power" (e.g., *ENA*) traditions, he still advocates for studying the GE as a comprehensive system.

To enhance intertemporal comparability, I restrict the analysis to the most standard curricula, discarding Ph.Ds, MBAs, and executive or specialized masters, as these degrees only emerged in recent decades. I also exclude the international programs for foreign students held at *ENA* since 1964. Table 1 provides summary statistics for each school of the sample, such as period covered by the data, year of admission of the first women, average number of students per year, and share of students with a "native" surname—defined as surnames consistently observed in the censuses over the period.

Appendix Figure A.2 reports for each school the evolution over time of the raw number of graduates (A.2a) and of the share of the French population enrolled (A.2b). We observe troughs followed by peaks in enrollments during the First and the Second World Wars. These modest annual fluctuations should have very limited impacts on intergenerational mobility estimates, since I am analyzing cohorts over 25 years and not short-term variations. Moreover, while the share of the population graduating from the prestigious GE increased over the last 130 years, the proportions were not comparable to the massive rise in university graduates. After the turning point of World War 1, most of the increase in GE graduates can be attributed to business schools and *Sciences Po Paris*, the latter accounting for about half the graduates in the sample. Therefore, I study *Sciences Po* separately and consider the 9 other *Grandes Écoles* as the baseline, to avoid disproportionate influence from a single school in the baseline sample. It can also be seen that data coverage is slightly incomplete after 2010—in particular, business schools are only covered until 2012. I discuss in footnote 19 the minor implications of this small attrition.

Observations systematically include the surname and the first name of each student. At least one middle name is provided for one third of observations. The maiden name is provided for almost all women, but only for 15% do I observe both a maiden and a married

name. 30% of the observations mention gender; I widen this to cover 99.7% of the students by using a gender propensity score for each first name.[15]

Additional care was applied to data processing. To ensure consistency between the different sources and improve matching quality, I implemented token and bigram fuzzy matchings on surnames between the GE data and the censuses. This allowed misspelled or shortened surnames to be detected in some school registers. Using a multicriteria algorithm (see details in Appendix C.2), I identify individuals educated at multiple elite colleges; this is important when examining admissions to a pool of colleges like my baseline 9 GE, to ensure that students with multiple degrees are taken into account only once. After exclusion of the non-standard curricula, I find that among the 303,514 curricula of the full sample, 267,943 individuals attended one school only, 16,465 attended two schools, while 871 graduated from three different schools and 7 from no less than four schools.[16]

Finally, I approximate the birth year of each student. As admissions at Master, MBA, or Ph.D. level were discarded, and due to the required two-year post-secondary school preparation before the admission examination, the standard age of first admission to a *Grande École* is around 20 years old. This is consistent with statistics reported by several schools. Students admitted in 1886 (2010) are therefore assumed to have been born in 1866 (1990). The birth year of students attending multiple GE is based on the first school they attended. An exception is admissions to *ENA*, which occur at older ages.[17]

The second source of data is the French nominative census, as information on the frequency of surnames in the population is also required. The French National Statistics Institute produces for the period 1891-1990 a detailed birth census by surname for 1891-1990, structured in four cohorts of 25 years: 1891-1915, 1916-1940, 1941-1965, and 1966-1990.[18]

[15] I use a gendered birth census by first name over 1900-2016, from the French Institute of Statistics (INSEE). Some first names are gender-neutral but 93% of the first names from the college registers could be gender-categorized. Students enrolled before a school was accessible to women are identified as men.

[16] The most common sequences are *École Polytechnique* – *École des Ponts* (4,128 individuals), and *Sciences Po Paris* – *ENA* (3,559 individuals), which are expected combinations, as explained in Appendix C.1.

[17] As 56% of ENA-enrollees in the data previously attended at least one other GE of the sample, I identify their average age of admission to ENA as 27. I assume that students not previously attending another *Grande École* of the sample were also admitted to ENA at age 27. This is an approximation, notably due to admissions of some on-the-job students who are already civil servants (*concours interne*).

[18] *Fichiers des noms patronymiques de 1891 à 1990*, edition 1999, INSEE (producer), ADISP (distributor).

The 25-year divide roughly corresponds to generations, since Mazuy et al. (2015) show that in 1946, 1966, and 1986, most births occur when parents are between 20 and 30 years old. I therefore consider those born in 1891-1915 as the parents of those born in 1916-1940, the grandparents of those born in 1941-1965, and the great-grandparents of those born in 1966-1990. I create an extra pseudo-cohort (1971-1995) for students born up to 1995, assuming a similar number of births per surname as in 1966-1990.[19] Obviously, not everyone born in the $[Year; Year + 25[$ interval has their children in the $[Year + 25; Year + 50[$ interval. For individuals at the edge of a generation, parents and children may be considered in the same generation on some occasions—when there are less than 25 years between their births—, or sometimes two generations apart. This constitutes a mismeasurement of the intergenerational link that adds noise and therefore somewhat biases estimations. The exact magnitude of the bias is difficult to gauge, but its direction is downward, provided the parent–child association is the strongest within the family.

As the pool of potential applicants to the *Grandes Écoles* is measured according to number of births per surname in France, my analysis has to be restricted to those surnames providing the clearest observations of number of births. This means "native" surnames, defined as those where the latest immigration occurred at the beginning of the period of study, i.e., the late 19^{th} century.[20] Many scholars working with surnames over a long historical timespan have had to make the same choice (e.g., Dupâcquer and Kessler 1992). It also limits the issue of self-selection in migration, i.e., the fact that migrants have different

[19] The total number of births between 1971 and 1995 was simply 3% lower than between 1966 and 1990 (source: INSEE). This assumption on surname distribution and the fact that the graduate lists are less exhaustive after 2010 imply that the results for the 1971-1995 pseudo-cohort should be read more cautiously. Yet, on top of widening the historical panorama of the paper, the 1971-1995 pseudo-cohort adds a glimpse of the evolution of mobility within the last decade, when it is compared to the 1966-1990 cohort.

[20] Among potential GE candidates, this restriction eliminates individuals born abroad who immigrated to France before age 20, for whom data on number of births is partial. By contrast, it includes those born in France who emigrated to study, but this is not a concern for several reasons. Such emigration was historically rare and has remained relatively marginal more recently. Docquier and Marfouk (2006) show that only 3.4% of tertiary educated individuals emigrate outside France. The GE provide high-quality education, at a lower cost than alternatives in the English-speaking world, for instance. More importantly, while people may choose to emigrate to study, this does not remove studies at a GE from the opportunity set.

unobserved characteristics from natives (Borjas, 1987).[21] Appendix C.3 details how "native" surnames are identified, as opposed to "foreign" ones. The categorization operates purely at the surname level, regardless of individuals' nationality or migration history, which are not observed. In the end, surnames classified as "foreign" represent 17% of births in France in the last cohort (1966-1990), which is consistent with the proportions of immigrants and descendants of immigrants at that time (Bouvier, 2012).

Table 2 displays descriptive statistics by cohorts in three parts: (1) on the nominative censuses, (2) on graduates of the baseline 9 GE, and (3) on *Sciences Po* graduates. A third of surnames in the censuses are associated with immigration during the 20$^{\text{th}}$ century, and those immigrant surnames account for about 8% of all births. Over the whole period, there are 118,337 graduates of the 9 baseline GE and 113,085 graduates of *Sciences Po* with native surnames, each accounting for 0.18% of the native French population. The share of the population admitted to the GE progressively increases from the first to the last cohort, both at *Sciences Po Paris* and in the baseline schools—with the exception of the 1971-1995 pseudo-cohort with partially missing data. Studying in a *Grande École* today is not equivalent to doing so at the end of the 19$^{\text{th}}$ century. Still, there has been very little expansion compared to the rise of baccalaureate holders and tertiary education as a whole. Heterogeneity results by school and by school categories are a useful complement to the main analysis, as the composition of the baseline evolved across time, with the business schools producing an increasing share of graduates (Appendix Table B.1).

4 Empirical strategy

This section first presents the innovative methodological approach based on a surname-cohort-dependent probability of a characteristic. Here, I construct the probability of having a father who graduated from an elite college. Using this probability as an independent variable enables me to consider all the observations in the analysis, instead of restricting to

[21] Indeed, in France, Meurs et al. (2006) showed that, in 1999, first- and second-generation immigrants experienced more unemployment and lower access to high-status occupations. More recently, a report by the OECD (2016) similarly found that among those with poorly-educated parents, access to tertiary education is much lower for individuals both of whose parents are foreign-born (26.8%) compared to French-born (42.3%).

a discretionary sub-sample of rare surnames. In a second sub-section, I describe my measure of intergenerational mobility: the relative admission rate (RAR), which relates the admission rate of those with a GE-graduate father to the admission rate of the rest of the population.

4.1 Methodological contribution: surname-based probability of a GE-graduate father

Surnames have little or no direct effect on socio-economic outcomes, especially once foreign surnames (which may be discriminated against) are excluded. Yet, they work as family trackers: all bearers of a surname in generation $t+1$ are descendants of a father sharing the same surname in generation t.[22]

Linking generations via surnames is effective but not perfect. Most surnames share a common stem but subdivide into distinct family branches. Fortunately, the distribution of surnames is highly skewed, with an abundance of rare surnames in many countries, particularly in France (Appendix Figure A.3). Surname-based studies generally restrict their samples to rare surnames (e.g., Clark et al. 2014; Güell et al. 2015). To avoid choosing an arbitrary threshold of rarity, and given the lack of clear father?child linkage in the data, I define as an explanatory variable the surname- and cohort-dependent probability of having a GE-graduate father. The rarer the surname, the more likely their bearers in the older and younger cohorts are to be directly related.

Let me first define $AR_{GE,c,S}$ as the admission rate (AR) to a given (set of) GE of the bearers of surname S in cohort c:

$$AR_{GE,c,S} = \frac{St_{GE,c,S}}{N_{c,S}}$$

with $St_{GE,c,S}$ the number of students with surname S born in cohort c graduating from the

[22] In France, surnames have been hereditarily transmitted through the patriarchal line since the 12^{th} century, although surname selection and mutations were very common in the middle ages. From 1474 on, surnames could not be modified without the King's approval. In 1539, the order of Villers-Cotterêts made it compulsory to register family names, whose spelling was stabilized. Since 1870, the spelling of surnames has been definitively settled through the *Livret de famille* (family register). Two recent laws of 2003 and 2008 state that the father's name, the mother's name, or a combination of both may be chosen. The study focuses on individuals born between 1866 and 1995, which makes patronyms a reliable intergenerational link between fathers and sons, and between fathers and daughters through their maiden names.

GE and $N_{c,S}$ the number of births of bearers of surname S in the French population in cohort c. GE will be, in turn, the whole set of schools, the 9 schools in the baseline, a single school, or specific categories (like engineering or business schools). Denoting male students as St^M, and under the assumption that as many men as women bear each surname in each cohort, I am able to compute a male admission rate:

$$AR^M_{GE,c,S} = \frac{St^M_{GE,c,S}}{N_{c,S}/2}$$

I then define $X_{GE_c,S}$ as the probability that the father of someone bearing surname S born in cohort c studied in a single (or group of) GE. It takes value 0 for those whose surnames do not appear in the GE in the previous generation, and the probability of value 1 that the father is a GE graduate for the others.[23] $X_{GE_c,S}$ is a function of the number of male births and of male graduates bearing the same surname in the previous cohort, specifically being the GE admission rate of male bearers of the same surname in the previous cohort:

$$X_{GE,c,S} = \frac{St^M_{GE,c-1,S}}{N_{c-1,S}/2} = AR^M_{GE,c-1,S}$$

I focus on transmission from fathers to both sons and daughters. I focus on transmission from fathers to both sons and daughters. Although an OECD report (2016) stated that upward mobility when only one parent holds a higher qualification is about the same regardless of which, Beller (2009) argues that excluding mothers—as most studies do—is not trivial for estimations of intergenerational mobility. In my setting, however, women remain a small minority of the sample of graduates, only 7%, 6%, 6%, and 18% of the four ancestors' cohorts. In addition, the homogamy, common among the French elite, implies that fathers and mothers have very similar characteristics (Goux and Maurin, 2003; Bouchet-Valat, 2014; Frémeaux and Lefranc, 2020).

Another important dimension of intergenerational mobility is fertility, especially if birthrates depend on socio-economic background.[24] As I observe education, fertility, and the probabil-

[23] Appendix Table B.2 reports descriptive statistics on these variables by cohort and father's school(s). For the multigenerational analysis, Appendix C.4 subsequently defines the probability that the paternal grandfather, great-grandfather and great-great-grandfather studied at a *Grande École*.

[24] A negative socio-economic and educational gradient of fertility was theoretically conceptualized through differences in opportunity costs and a children quantity–quality trade-off (Becker and Lewis, 1973; Becker and Tomes, 1976). Yet, this has been empirically questioned. Notably, Kravdal and Rindfuss (2008) use Norwegian data to demonstrate that although they are a few years older when they give birth, better-educated women do not have fewer children, while better-educated men are even less likely to remain childless.

ity of having a GE-graduate father at surname level, the accuracy of $X_{GE,c,S}$ depends on the underlying assumption that there are no major differences in fertility between GE-graduate fathers and other fathers bearing a similar surname. While I cannot test this assumption with the available data, demographical studies provide evidence of relatively uniform fertility in France.[25]

4.2 Measure of intergenerational mobility: relative admission rates

My measure of intergenerational mobility is simple to understand and interpret. I first define the admission rate to a given GE (or group of GE) for those born in cohort c whose father is a GE graduate, relating presence in the GE registers ($St_{GE,c,X_{GE,c,S}}$) to occurrence in the French population ($N_{c,X_{GE,c,S}}$). This comes down to the share of those born in cohort c whose father was a GE graduate and who were themselves admitted to the GE.[26]

$$AR_{GE,c,X_{GE,c,S}} = \frac{St_{GE,c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}}$$

I then define a risk ratio, a tool close to odds ratios, albeit more straightforward to interpret: the relative admission rate (RAR) of those with a GE-graduate father. This relates the GE admission rate of those with a GE-graduate father (identified with probability $X_{GE,c,S}$) to the admission rate of the rest of the population, i.e., the admission rate of the complementary group (identified with probability $X'_{GE,c,S}$):

$$RAR_{GE,c,X_{GE,c,S}} = \frac{AR_{GE,c,X_{GE,c,S}}}{AR_{GE,c,X'_{GE,c,S}}}$$

This relative admission rate is therefore the factor by which they are over- or under-represented in the GE compared to those without a GE-graduate father. If $RAR_{GE,c,X_{GE,c,S}}$ equals 1, the descendants of GE graduates are proportionately represented in the GE.

[25] The variance in fertility rates has progressively decreased in developed countries. In France, socio-economic differences in fertility are even lower than in other European countries (Toulemon et al., 2008), and the desired number of children has remained stable at 2 or 3 for decades (Sobotka and Beaujouan, 2014). While there was more volatility at the end of the 19$^{\text{th}}$ century (Dupâquier, 1988), the vast majority of 20$^{\text{th}}$ French families had 2 or 3 children (Toulemon, 2001). Finally, although extinction of family lines was common in the early 19$^{\text{th}}$ century, it became much rarer over the 20$^{\text{th}}$ century, and therefore only extremely rare surnames were threatened with vanishing (Dürr, 1992; Dupâquier, 1992).

[26] The focus is principally on admissions to the same (group of) school(s) as the fathers (e.g., in the baseline GE for those whose father graduated from a baseline GE); however in some cases, I measure admissions to a given GE for children with GE'-graduate fathers (e.g., admission to *ENA* for children of *ENS* graduates).

When $RAR_{GE,c,X_{GE,c,S}}$ is below 1, they are under-represented, while above 1, they are over-represented. The coefficients of the relative admission rates (RAR) and their corresponding confidence intervals are estimated with a log-binomial specification (Wacholder, 1986). More precisely, I estimate for the successive cohorts the probability of a binary outcome—having studied in a *Grande École* ($GE = 1$) or not ($GE = 0$)—, as a function of a univariate explanatory variable—the probability that the father is a GE graduate $X_{GE,c,S}$. The $RAR_{GE,c,X_{GE,c,S}}$ is simply the exponential of β in the following equation of the log-binomial model:

$$\log[P(GE = 1|c, X_{GE,c,S})] = \alpha_c + \beta X_{GE,c,S}$$

5 Results

The results presented in this section show that, over the last century, descendants of graduates enjoyed better prospects of admission to the *Grandes Écoles* than the rest of the population (sub-section 5.1). This intergenerational persistence runs across multiple generations (sub-section 5.2), with a particularly large advantage in the very same school from which fathers graduated (sub-section 5.3). Moreover, even the slowly increasing admissions of women show that daughters' advantage was comparable to sons' (sub-section 5.4).

5.1 Main result: like father, like child in the *Grandes Écoles*

An intuitive framework for intergenerational mobility consists in linking an achievement in one generation with the same achievement in the previous generation. Table 3 summarizes my main results on intergenerational mobility, displaying relative admission rates (RAR) to any of the baseline 9 *Grandes Écoles* for children of graduates of any of the 9 elite colleges.

For the 1891-1915 cohort (first line), 14,619 sons and daughters with 5,502 different surnames had a father born between 1866 and 1890 who was a GE graduate.[27] 1,766 of those children were admitted to one of the GE, a 12.1% admission rate compared to the overall admission rate of 0.13%. This implies a hefty relative admission rate for children

[27] With 7,996 fathers in the paternal generation bearing these 5,502 surnames, this corresponds to 1.83 children per graduate on average.

of graduates born between 1891 and 1915, whose likelihood of enrolling in these prestigious schools was 154 times higher than the rest of the population's.

Although the magnitude of social reproduction significantly reduced in the next generation, it has remained amazingly stable for all subsequent cohorts born since World War 1. Those born over the 80-year period between 1916 and 1995 still had a 72- to 83-times greater chance of enrolling in one of the very prestigious *Grandes Écoles* if his or her father was a GE graduate, with no statistical difference between the point estimates of the successive cohorts. This has had substantial consequences on the student body composition of the *Grandes Écoles*, with between 13 and 17% across cohorts being children of graduates, i.e., one every 6 to 8 students.

As a sensitivity analysis, Appendix Table B.3 provides comparable findings for the complete sample including *Sciences Po Paris*, with point estimates only slightly lower. Appendix Table B.4 also reports a series of robustness tests for the colleges of the baseline. I show that both the trend and the orders of magnitude of the estimates are robust to restrictions to rarer surnames, for which lineages can be tracked more accurately. Results are also robust to the inclusion of "immigrant" surnames, although the evolution pattern obviously fluctuates more sharply for the more recent cohorts, with the GE welcoming more international students.

To explore whether the moderate growth in size of the elite colleges (from 0.13% to 0.25% of the population) influenced trends in mobility, Appendix Table B.5 reports relative admission rates to a subset of elite colleges with stable student populations across all cohorts: *École Polytechnique*, *ENS*, and *ESPCI*. While the pattern of decreasing relative admission rates is confirmed between the first two cohorts, the trend in the following cohorts appears to be contrastingly upward, although statistical power is limited for these small schools.[28] Bearing in mind that the three colleges may differ other than simply in the evolution of student population size, this rather U-shaped pattern suggests that the stability of intergenerational mobility documented throughout most of the 20$^{\text{th}}$ century in the baseline sample may stem from increased admissions to the other GE. While it remained modest compared

[28] Whereas the decline from 1891-1915 to 1916-1940 is significant at the 2% level, the increase between 1916-1940 and 1971-1995 is only significant at the 11% level.

to the universities, a larger increase could appear as a step towards greater *qualitative democratization*.

5.2 Dynastical analysis: like (great-)grandfather, like child

The surname-based methodology also enables me to study multigenerational mobility. An important finding is that the high admission prospects transmitted by the GE graduates of the first cohort do not carry on to their descendants two, three, or even four generations later. Figure 1 presents a multi-generational perspective, with the relative admission rates to the 9 baseline *Grandes Écoles* of descendants of GE-graduate fathers, grandfathers, great-grandfathers, and great-great-grandfathers.[29] By construction, the dark solid curve representing father-child associations recalls uni-generational results from Table 3.

In each cohort, multi-generational transmission is lower than father-child transmission. The lighter curve represents the relative admission rates of descendants of male ancestors born between 1866 and 1890. Their children had 154 times more chance of admission. While their grandchildren remained privileged, it was 3.4 times less, with 45 times more chance of admission. The RAR of their great-grandchildren born between 1941 and 1965 was still 33, while their great-great-grandchildren—born a century after them—still had 15 times more chance of enrolling in a *Grande École* than the rest of the population. Children, grandchildren, and great-grandchildren of graduates born in 1891-1915 had 81, 54, and 30 times more chance of admission, respectively. Children and grandchildren of graduates born in 1916-1940 had 72 and 34 times more chance of admission, respectively.

Overall, across cohorts, a descendant of a GE graduate still had a 30- to 54-times higher likelihood of admission to a *Grande École* two to three generations later. First, this shows that the very high father-child transmission of the first cohort is a one-off, not passed on in comparable magnitudes to later descendants. Secondly and more importantly, it constitutes evidence of a persistent multigenerational over-representation of families, who dynastically enroll in the French elite colleges.

[29] The dynastical setting does not account for the mediating role of intermediate generations, such that it constitutes a gross measure of persistence in the long run.

5.3 Heterogeneity analysis: in the exact paternal footsteps

The weight of the different colleges in the baseline has evolved over the last century, with the business schools in particular accounting for an increasing share of graduates (see Appendix Table B.1). Figure 2 therefore reports relative admission rates of children of GE graduates to the father's school or school category. A first salient result is that relative admission rates are significantly higher than for the baseline overall, showing greater social reproduction.

These results also indicate that the changing composition of student bodies in the baseline schools is not the underlying factor in the drop in intergenerational persistence between the first two cohorts (1891-1915 and 1916-1940). This is confirmed for all colleges or college categories where ancestors were involved, albeit the difference is not significant for *ENS*. With respectively only 112 and 130 students per year on average, confidence intervals of the RAR for *ENA* and *ENS Ulm* are indeed large, but intergenerational reproduction appears particularly high in these two colleges.[30] As for *Sciences Po Paris*, admissions to the college remain more likely for children of graduates of the institution across all cohorts, but this advantage has greatly decreased, and is lower than what has been observed at the engineering schools, *ENS* or *ENA* since the 1916-1940 cohort. Yet, comparisons of intergenerational reproduction in one school to what happens within a group of several may be misleading.

I therefore deepen the characterization of intergenerational mobility processes over the period by college. Table 4 is a heat matrix, which reports for the cohort born in 1971-1995 the 100 combinations of relative admission rates at the school of origin (father) – school of destination (child) level. The main objective is to identify the intergenerational inter-college dynamics among the most prestigious *Grandes Écoles*, thereby providing a typology of elite education trajectories for children of graduates of each of these elite institutions. The lines contain the relative admission rates to each of the 10 schools, based on the father

[30] While Euriat and Thélot (1995) found that admission to *ENS* was as much influenced by parental occupation for cohorts born around 1970 as for those born in the 1930s, I suggest that there is decreasing inequality in admissions. Although confidence intervals are too large to conclude with certainty, relative admission rates progressively fell from 458 to 350, 319, and 244. Contrastingly, they found decreasing inequality in admissions to *ENA*. However, I find relatively stable point estimates of RAR to *ENA* for children of *ENA* graduates between the 1941-1965 (254) and 1966-1990 (249) cohorts, whereas the pseudo-cohort results point to increasing intergenerational reproduction in recent decades (330).

having studied in one particular school. Each column reports the relative admission rate to a given school depending on the school from which the father graduated. As this analyzes admissions to a given college conditionally on the father's being a graduate of one particular college, sample sizes are reduced, especially for the first cohort with fewer students, or for smaller colleges such as *ESPCI*. With those limitations in mind, results in Table 4, combined with similar matrices for each cohort in Appendix Tables B.6a, b, c, and d, remain very informative regarding intergenerational patterns of mobility.

The most salient result is that descendants tend to mimic their ancestors: the diagonals are among the darkest areas for all cohorts, with particularly high magnitudes for *ENS Ulm*, *École Polytechnique*, and *ENA*. As "intellectual" schools are grouped on the upper-left (*ENS Ulm* and engineering schools), and schools operating within the "power tradition" on the bottom-right (business schools, *Sciences Po Paris*, and *ENA*), the wider dark area around the diagonal indicates that this typology by Bourdieu (1989) also applies when it comes to intergenerational mobility dynamics.

The darker lines in Table 4 indicate those schools where having a graduate father improves the prospects of admission the most. This is particularly true of *ENA* and *École Polytechnique*. Darker columns indicate those schools which are less accessible to those whose fathers are not graduates of any of the *Grandes Écoles*. This includes *Ponts*, *École Polytechnique* again, as well as business schools and *ENS Ulm* in the more recent cohorts.[31]

The exercise also informs on college-specific trends. It highlights the singularity of *ENS*: across cohorts, children of *ENS* graduates have a considerable advantage in admissions to *ENS*, but less so to other schools. This especially contrasts with children of graduates from engineering schools or *ENA*, , who have strong probabilities of being admitted to any of the top colleges. Regarding admissions to *ENA*, an apparent improvement in admission equality for those born in 1941-1965 compared to the 1916-1940 cohort suggests relatively early democratization (last columns of Appendix Tables B.6b, c). However, having a GE-graduate

[31] As a benchmark of the RAR of 296 to *Polytechnique* for children of *Polytechnique* graduates, Le Bras (1983) provided a monograph of admissions to *Polytechnique* in 1979. He showed that 25% of students had a *Polytechnique* graduate in their extended family, and that graduates were 400 times more likely to be the children of engineers than of skilled workers.

father weighs more heavily on admissions to the school for both the 1966-1990 cohort (B.6d) and the 1971-1995 pseudo-cohort, confirming that family background has become increasingly important at *École Nationale d'Administration* in the most recent period.

The dynamics of the engineering schools are also worth noting. *ESPCI* is a relatively special case and shows lower intergenerational associations with the other engineering schools. By contrast, *Ponts*, *Mines*, and *Télécom* are much more strongly linked to *Polytechnique*. While the fact that top-ranked *Polytechnique* students often go on to one of these applied engineering schools may explain the *intra*generational link between these colleges, a striking result of this study is the very strong *inter*generational reproduction between different engineering schools. Actually, for cohorts born since 1941, the chances of admission to *École des Ponts* appear higher for someone whose father is a graduate of *École Polytechnique* than of *École des Ponts* itself—although this is based solely on differences in point estimates.

The matrices complement the findings concerning *Sciences Po Paris*, where intergenerational reproduction was substantial over the whole period, , yet among the lowest of the 10 schools after the First World War. Indeed, the penultimate column is very dark for the first cohort (Appendix Table B.6a) and lightens progressively (B.6b, c, d). In addition, the relative homogeneity of estimates for *Sciences Po*'s column across cohorts indicates that *Sciences Po Paris* is a natural destination for children of graduates, whatever GE their father attended. By contrast, for most cohorts, children of graduates of *Sciences Po Paris* had noticeably lower RAR to engineering schools, while their RAR to business schools was of similar magnitude to their RAR to *Sciences Po* itself.

Except for *Sciences Po Paris*, confidence intervals limit the significance of these heterogeneity analyses. Nevertheless, the set of matrices enables clear conclusions to be drawn about how accessible specific colleges are to the offspring of graduates, depending on their father's *alma mater*. For instance, for the 1971-1995 cohort, admission to *Polytechnique* was 296 [95% confidence interval: 209-420] times more likely for those with a *Polytechnique* graduate father, while the advantage was significantly lower if the father graduated from *Sciences Po* (88 [66;119]) or from ESCP (102 [71;146]). Likewise, the RAR to *École des Ponts* of a child born in 1971-1995 was 279 [175;447] if his or her father graduated from *Polytechnique*,

but only 42 [20;86] if the father graduated from *ESSEC*.

5.4 Gender analysis: like fathers, like (outnumbered) daughters

"Daddy gladly said: *Simone has the brain of a man; Simone is a man.* I was nevertheless treated as a girl", wrote Simone de Beauvoir (1958) in her autobiography *Mémoires d'une jeune fille rangée*. She suffered from being held back and forbidden to pursue a "male" education, as her father actually stated that he would have registered her for *Polytechnique*'s examination if she had been a man. A crucial structural change across the 20th century in tertiary education was the increasingly common admission of women. This process was slower in the French *Grandes Écoles*, for which Appendix Figure A.4 reports the share of female students in relevant school categories by year of admission.[32]

In this gender analysis, I measure male and female relative admission rates separately. More precisely, I relate admission rates of sons of graduates to admission rates of sons of non-graduates, and relate admission rates of daughters of graduates to those of daughters of non-graduates.[33] Because the share of women in the different colleges differs sharply, I analyze colleges singly rather than grouping the baseline colleges as a whole.

Figure 3 reports the relative admission rates to *Sciences Po* of sons versus daughters of *Sciences Po* graduates (Figure 3a), and to *ENS* of sons versus daughters of *ENS* graduates (Figure 3b)—the two schools where admission of women began the earliest. Trends in in-

[32]The increasing share of women in GE is more extensively discussed in Appendix C.1. In short, only *École Normale Supérieure* educated women across the whole period, albeit in separate schools for men and women until 1985. During World War 1, women sporadically attended some *Grandes Écoles*. From 1919 onward, they were admitted to *Sciences Po Paris*, although under distinct entry requirements until World War 2. Women were entitled to apply to *ENA* from its foundation in 1945, though they remained a minority. But it was only in the early 1970s that women were finally granted legal access to all GE. While their presence rapidly increased at *Sciences Po Paris* and in business schools, progress was much slower in engineering schools or at *ENA*. This persistent under-representation of women in most GE, even in the early 21st century, contrasts with the reversed gender gap in access to higher education observed in most developed countries, including France (Buchmann and DiPrete, 2006, 2013).

[33]Here again, I am examining admission probabilities for those with a GE-graduate father. Although outside of *ENS* and *Sciences Po*, women constituted only 1.3% of graduates until 1971 (ancestors' generations), I would have liked to investigate transmission from mothers. However, data limitations and selection on spouse names made this impossible, despite several attempts to hand-collect them for subsets of schools and periods—notably through wedding announcements in alumni magazines, or announcements of official administrative appointments.

tergenerational mobility appear relatively parallel for sons and daughters of graduates, with no significant difference. Confidence intervals of estimates are wider for the smaller *ENS*, and while the insignificant differences in point estimates are more pronounced in the first two cohorts at *ENS*, it should be remembered that there were two separate single-gender institutions until 1985.

Relative admission rates to each of the 10 schools for sons and daughters separately are reported by cohort in the appendix, whether their father studied in the same school (Appendix Table B.7a), or in any of the 9 colleges of the baseline (Appendix Table B.7b). Out of the 71 comparable pairs of gendered estimates, only 3 differ significantly. One difference is irrelevant due to the infinitesimal number of (female) students at *ESPCI*.

A second difference concerns the 1891-1915 cohort at *Sciences Po*. While there was no difference between sons and daughters of *Sciences Po* graduates for admissions to *Sciences Po* (first line of Appendix Table B.7a and Figure 3a), the RAR was significantly higher for daughters than for sons of graduates of any of the baseline 9 GE (first line of Appendix Table B.7b). This suggests that in the early 20$^{\text{th}}$ century, elite college graduates could secure admission for their sons to all colleges, while their daughters were sent to the schools they were entitled to apply to, notably *Sciences Po* (the *École Libre des Sciences Politiques* at that time).

A third significant difference is the higher RAR to *ENS* in the first cohort for sons of graduates of any of the baseline colleges. It should be recalled that there were two single-gender *ENS* at that time. Contrastingly, point estimates (although not statistically different) indicate higher RAR to *ENS* for daughters once the school became coeducational. More generally, I find evidence suggestive of (non-significant) higher RAR for daughters of *Polytechnique* and *Ponts* graduates across cohorts, as well as for most colleges in the 1971-1995 pseudo-cohort.

Taken together, these results indicate that sons and daughters of graduates have benefited from relatively comparable advantages. If anything, intergenerational reproduction in the *Grandes Écoles* is slightly higher from fathers to daughters, possibly contributing to the slight and non-significant upturn in social reproduction over the last decade. Indeed, if

social reproduction is higher for daughters, the rising share of women among graduates would mechanically increase measures of intergenerational persistence. However, this cannot be the sole factor: a comparison of the last columns of Appendix Tables B.7a and B.7b also suggests a slightly increasing persistence over the last decade when considering sons alone. In any case, the increasing admission of women to GE constitutes a progress toward gender equality.

6 Conclusion

If graduates of highly selective private colleges are "disproportionately" over-represented in top positions in both the public and private sectors in the United States (Chetty et al., 2023), graduates of the top *Grandes Écoles* in France are even more so. This paper provides the first estimations of intergenerational mobility regarding these elite colleges, with a historical perspective covering more than a century. The sample, born between 1866 and 1995, covers graduates of 10 of the most selective and prestigious colleges, thereby accounting for the top 0.36% of the educational distribution in France.

Despite the stated objective of the partly anonymous *concours* (admission examinations), I show that the meritocratic promises rooted in the French Revolution—when several GE were founded—have not been fully kept. Descendants of graduates of the French elite colleges have a decisive and durable advantage over the rest of the population when it comes to admission. Those born between 1971 and 1995 whose fathers graduated from *École Nationale d'Administration* had a 330-times higher likelihood of studying at *ENA* themselves. These advantages are long-lasting and dynastical, as descendants of graduates continue to be 30 to 54 times more likely to graduate from a *Grand École* two to three generations after their ancestors. All this highlights how the French elites secure the education of their offspring through what can be called a "glass floor".

Previous studies (using occupation as a background characteristic) claimed there had been a decline in intergenerational reproduction in access to higher education over the 20$^{\text{th}}$ century in France (Euriat and Thélot, 1995; Vallet and Selz, 2008; Falcon and Bataille, 2018). Yet, the present paper shows that, after a *qualitative democratization* at the beginning of the

century, intergenerational mobility regarding the most prestigious *Grandes Écoles* has not improved since World War 1. If anything, intergenerational persistence in elite education has surreptitiously risen in recent decades. This may partly be a mechanical consequence of the increasing admission of women, since social reproduction from fathers to daughters at the top colleges appears slightly higher than from fathers to sons.

While my findings cannot be interpreted as showing that the GE are not accessible without these advantages, they indicate that French citizens do not share a common starting line. Although there is some intergenerational reallocation, the tiny minority who are children of *Grande École* graduates account for up to 17% of admissions. Some findings in the paper suggest that creating more places in the GE would help to diversify admissions.

Not only is studying at some of the *Grandes Écoles* virtually free, but students are even considered trainees and paid by the State in a subset of colleges. Yet, intergenerational mobility is lower in these colleges than in the more expensive business schools. One of the limitations of the study is that I cannot disentangle the channels of transmission. However, while wealth and economic capital may play a non-negligible role at different stages of the educational process, it appears that financial constraints may have less impact than cultural capital on intergenerational reproduction of elite education.

References

Albouy, V. and Wanecq, T. (2003). Les inégalités sociales d'accès aux grandes écoles. *Économie et statistique*, 361(1):27–52.

Allouch, A. (2022). *Les nouvelles portes des grandes écoles*. PUF.

Barone, C. and Ruggera, L. (2018). Educational equalization stalled? Trends in inequality of educational opportunity between 1930 and 1980 across 26 European nations. *European Societies*, 20(1):1–25.

Barone, G. and Mocetti, S. (2020). Intergenerational mobility in the very long run: Florence 1427–2011. *The Review of Economic Studies*.

Barsoux, J.-L. and Lawrence, P. (1991). The making of a French manager. *Harvard Business Review*, 69(4):58–60.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, 81(2, Part 2):S279–S288.

Becker, G. S. and Tomes, N. (1976). Child endowments and the quantity and quality of children. *Journal of political Economy*, 84(4, Part 2):S143–S162.

Belhoste, B. (2002). Anatomie d'un concours. L'organisation de l'examen d'admission à l'École polytechnique de la Révolution à nos jours. *Histoire de l'éducation*, 94:141–175.

Beller, E. (2009). Bringing intergenerational social mobility research into the twenty-first century: Why mothers matter. *American Sociological Review*, 74(4):507–528.

Black, S. E. and Devereux, P. J. (2011). Recent developments in intergenerational mobility. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, pages 1487–1541. Elsevier. edition 1. volume 4. chapter 16.

Bonneau, C., Charrousset, P., Grenet, J., and Thebault, G. (2021). Quelle démocratisation des grandes écoles depuis le milieu des années 2000 ? *Rapport Institut des politiques publiques (IPP)*, N°30.

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American economic review*, 77(4):531–553.

Bouchet-Valat, M. (2014). Les évolutions de l'homogamie de diplôme, de classe et d'origine sociales en France (1969-2011) : ouverture d'ensemble, repli des élites. *Revue française de sociologie*, 55(3):459–505.

Bourdieu, P. (1989). *La Noblesse d'État. Grandes écoles et esprit de corps*. Les Éditions de Minuit, Paris.

Bourdieu, P. and de Saint Martin, M. (1987). Agrégation et ségrégation. Le champ des grandes écoles et le champ du pouvoir. *Actes de la recherche en sciences sociales*, 69:2–50.

Bourdieu, P. and Passeron, J.-C. (1964). *Les Héritiers. Les étudiants et la culture.* Les Éditions de Minuit.

Bouvier, G. (2012). Les descendants d'immigrés plus nombreux que les immigrés : une position française originale en Europe. *Immigrés et descendants d'immigrés en France*, pages 11–28.

Breen, R., Luijkx, R., Müller, W., and Pollak, R. (2009). Nonpersistent inequality in educational attainment: Evidence from eight European countries. *American journal of sociology*, 114(5):1475–1521.

Breen, R. and Müller, W. (2020). *Education and intergenerational social mobility in Europe*

and the United States. Stanford University Press.

Brezis, E. S. and Hellier, J. (2018). Social mobility at the top and the higher education system. *European Journal of Political Economy*, 52:36–54.

Buchmann, C. and DiPrete, T. A. (2006). The growing female advantage in college completion: The role of family background and academic achievement. *American sociological review*, 71(4):515–541.

Buchmann, C. and DiPrete, T. A. (2013). *The rise of women: The growing gender gap in education and what it means for American schools*. Russell Sage Foundation.

Chetty, R., Deming, D. J., and Friedman, J. N. (2023). Diversifying society's leaders? The causal effects of admission to highly selective private colleges. *National Bureau of Economic Research Working Paper Series*, WP 31492.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the United States. *The Quarterly Journal of Economics*, 135(3):1567–1633.

Clark, G. and Cummins, N. (2014). Surnames and social mobility in England, 1170–2012. *Human Nature*, 25(4):517–537.

Clark, G., Cummins, N., Diaz Vidal, D., Hao, Y., Ishii, T., Landes, Z., Marcin, D., Mo Jung, K., Marek, A., and Williams, K. (2014). *The Son also Rises: 1,000 Years of Social Mobility*. Princeton University Press, Princeton.

Collado, M. D., Ortuño-Ortín, I., and Romeu, A. (2012). Intergenerational linkages in consumption patterns and the geographical distribution of surnames. *Regional Science and Urban Economics*, 42(1-2):341–350.

de Beauvoir, S. (1958). *Mémoires d'une jeune fille rangée*. Gallimard.

Delefortrie-Soubeyroux, N. (1961). *Les dirigeants de l'industrie française*, volume 6. Armand Colin.

Docquier, F. and Marfouk, A. (2006). International migration by education attainment, 1990–2000. In Ozden, C. and Schiff, M., editors, *International migration remittances and the brain drain*, pages 151–199. World Bank and Palgrave Macmillan.

Dupâcquer, J. and Kessler, D. (1992). *La société française au XXIème siècle*. Fayard.

Dupâquier, J. (1988). *Histoire de la population française (tome 3): De 1789 à 1914*. PUF.

Dupâquier, J. (1992). Nos patronymes vont-ils disparaître ? In Dupâcquer, J. and Kessler, D., editors, *La société française au XXIème siècle*, pages 461–488. Fayard.

Dürr, M. (1992). La reproduction différentielle des familles. In Dupâcquer, J. and Kessler, D., editors, *La société française au XXIème siècle*, pages 439–459. Fayard.

Erikson, R. and Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford University Press.

Euriat, M. and Thélot, C. (1995). Le recrutement social de l'élite scolaire en France : évolution des inégalités de 1950 à 1990. *Revue française de sociologie*, pages 403–438.

Falcon, J. and Bataille, P. (2018). Equalization or reproduction? Long-term trends in the intergenerational transmission of advantages in higher education in France. *European Sociological Review*, 34(4):335–347.

Ferrand, M., Imbert, F., and Marry, C. (1999). *L'excellence scolaire, une affaire de famille : le cas des normaliennes et normaliens scientifiques*. Editions L'Harmattan.

Frémeaux, N. and Lefranc, A. (2020). Assortative mating and earnings inequality in France. *Review of Income and Wealth*, 66(4):757–783.

Ganzeboom, H. B., Treiman, D. J., and Ultee, W. C. (1991). Comparative intergenerational stratification research: Three generations and beyond. *Annual Review of sociology*, 17(1):277–302.

Goux, D. and Maurin, E. (2003). Who marries whom in France? In Blossfeld, H.-P. and Timm, A., editors, *Who Marries Whom? Educational Systems as Marriage Markets in Modern Societies*, pages 57–78. Springer.

Güell, M., Rodríguez Mora, J. V., and Telmer, C. I. (2007). Intergenerational mobility and the informative content of surnames. *CEPR Discussion Papers*. nº6316.

Güell, M., Rodríguez Mora, J. V., and Telmer, C. I. (2015). The informational content of surnames, the evolution of intergenerational mobility, and assortative mating. *The Review of Economic Studies*, 82(2 (291)):693–735.

Henderson, M., Shure, N., and Adamecz-Völgyi, A. (2020). Moving on up: 'first in family' university graduates in England. *Oxford Review of Education*, 46(6):734–751.

Kravdal, Ø. and Rindfuss, R. R. (2008). Changing relationships between education and fertility: A study of women and men born 1940 to 1964. *American sociological review*, 73(5):854–873.

Le Bras, H. (1983). Les origines d'une promotion de Polytechniciens. *Population*, pages 491–502.

Mare, R. D. (1980). Social background and school continuation decisions. *Journal of the american statistical association*, 75(370):295–305.

Mazuy, M., Barbieri, M., Breton, D., and d'Albis, H. (2015). L'évolution démographique récente de la France et ses tendances depuis 70 ans. *Population*, 70(3):417–486.

Meer, J. and Rosen, H. S. (2009). Altruism and the child cycle of alumni donations. *American Economic Journal: Economic Policy*, 1(1):258–86.

Meurs, D., Pailhé, A., and Simon, P. (2006). The persistence of intergenerational inequalities linked to immigration: Labour market outcomes for immigrants and their descendants in France. *Population*, 61(5):645–682.

Mullen, A. L., Goyette, K. A., and Soares, J. A. (2003). Who goes to graduate school? Social and academic correlates of educational continuation after college. *Sociology of Education*, 76(2):143–169.

OECD (2016). Education at a glance 2016: OECD indicators. *OECD Publishing*.

Pasquali, P. (2021). *Héritocratie : Les élites, les grandes écoles et les mésaventures du mérite (1870-2020)*. La Découverte.

Passant, A. J.-G. (2020). *L'origine des écoles de commerce : ESCP Business School, la passion d'entreprendre*. Editions L'Harmattan.

Peyrefitte, A. (1964). *La Rue d'Ulm*. Flammarion, Paris.

Picon, A., Belhoste, B., and Dahan-Dalménico, A. (1994). *La formation polytechnicienne 1794-1994*. Dunod.

Piketty, T. (2020). *Capital and Ideology*. Harvard University Press.

Santavirta, T. and Stuhler, J. (forthcoming). Name-based estimators of intergenerational mobility. *The Economic Journal*.

Shavit, Y., Arum, R., Gamoran, A., and Menahem, G. (2007). *Stratification in higher education: A comparative study*. Stanford University Press.

Shavit, Y. and Blossfeld, H.-P. (1993). *Persistent Inequality: Changing Educational Attainment in Thirteen Countries. Social Inequality Series.* Westview Press.

Sobotka, T. and Beaujouan, É. (2014). Two is best? The persistence of a two-child family ideal in Europe. *Population and Development Review*, 40(3):391–419.

Stolzenberg, R. M. (1994). Educational continuation by college graduates. *American Journal of Sociology*, 99(4):1042–1077.

Suleiman, E. N. (1978). *Elites in French society: the politics of survival*. Princeton University Press.

Thélot, C. and Vallet, L.-A. (2000). La réduction des inégalités sociales devant l'école depuis le début du siècle. *Économie et statistique*, 334(1):3–32.

Torche, F. (2015). Intergenerational mobility and equality of opportunity. *European Journal of Sociology/Archives Européennes de Sociologie*, 56(3):343–371.

Torche, F. (2018). Intergenerational mobility at the top of the educational distribution. *Sociology of Education*, 91(4):266–289.

Torche, F. and Corvalan, A. (2018). Estimating intergenerational mobility with grouped data: A critique of Clark's The son also rises. *Sociological Methods & Research*, 47(4):787–811.

Toulemon, L. (2001). How many children and how many siblings in France in the last century. *Population and Societies*, 374(4).

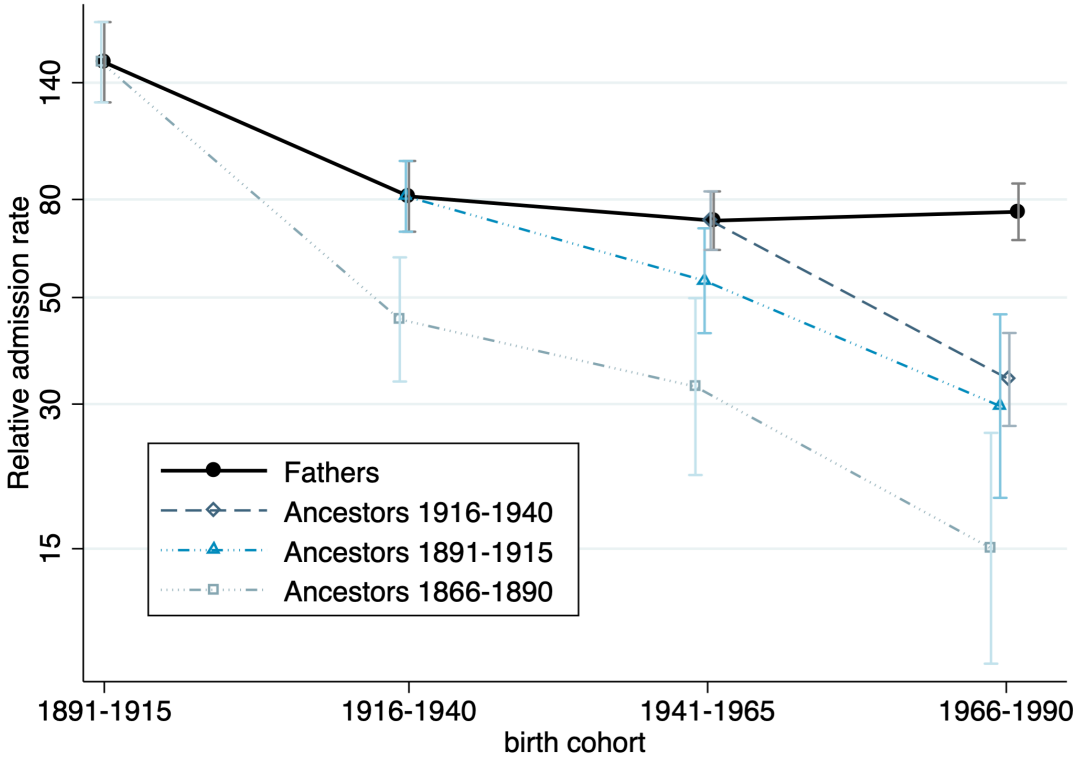
Toulemon, L., Pailhé, A., and Rossier, C. (2008). France: High and stable fertility. *Demographic research*, 19:503–556.

Vallet, L.-A. and Selz, M. (2008). Évolution historique de l'inégalité des chances devant l'école : des méthodes et des résultats revisités. *Idées économiques et sociales*, 4(154):59–68.

Vion, A., Joly, H., Grémont, É., and Dudouet, F.-X. (2014). Retour sur le champ du pouvoir économique en France : l'espace social des dirigeants du CAC 40. *Revue Française de Socio-Economie*, 1(13):23–48.

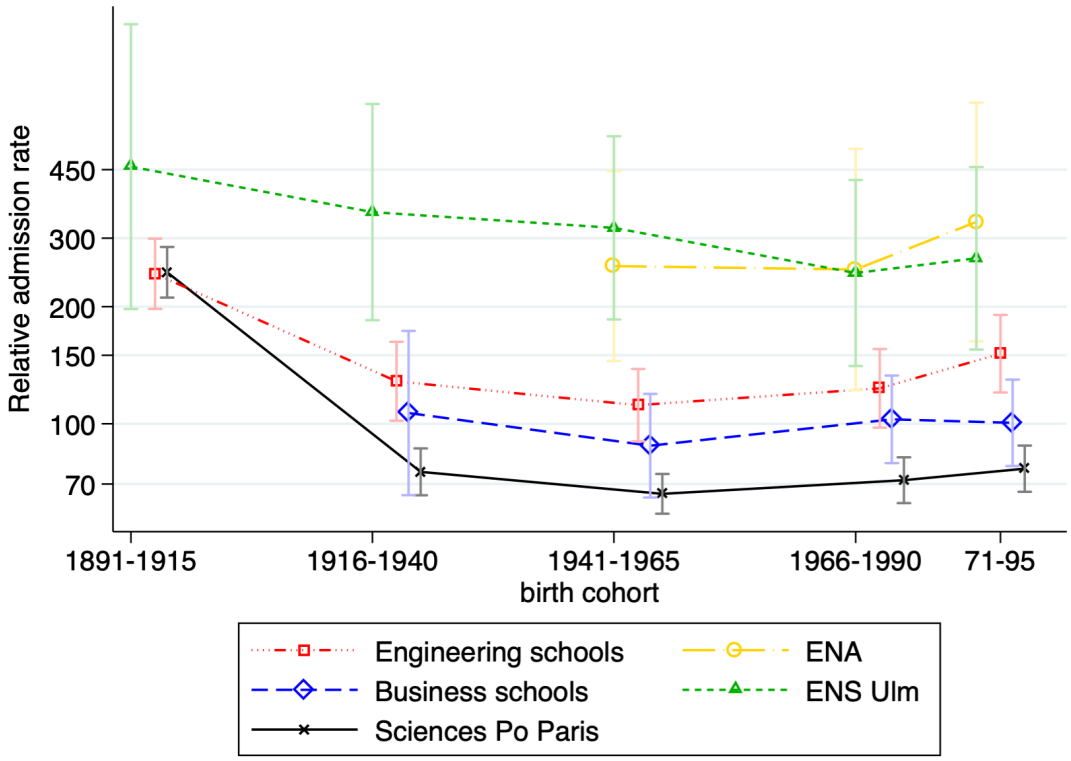
Wacholder, S. (1986). Binomial regression in GLIM: estimating risk ratios and risk differences. *American journal of epidemiology*, 123(1):174–184.

Figure 1: Dynastical admissions to the *Grandes Écoles*.



Notes: This figure reports by birth cohort the relative admission rates to the baseline 9 *Grandes Écoles* of those whose male ancestors graduated from these schools in different cohorts. Depending on the ancestor's birth cohort, they may be fathers, grand-fathers, great-grand-fathers, or even great-great-grandfathers. Brackets refer to 95% confidence intervals. I use a logarithmic scale for the ordinate. The 1971-1995 pseudo-cohort, less suitable for multi-generational analysis, is not reported.

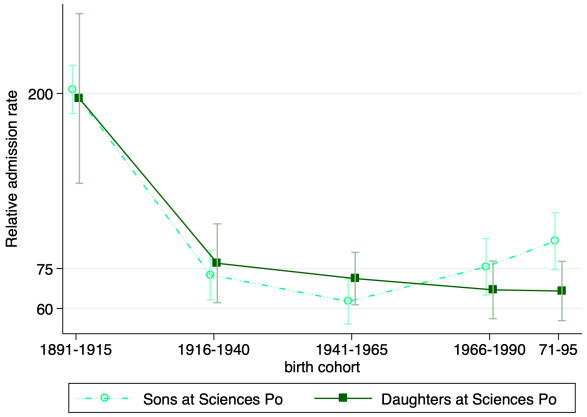
Figure 2: Relative admission rates of children of GE graduates to the same school (category) as their fathers.



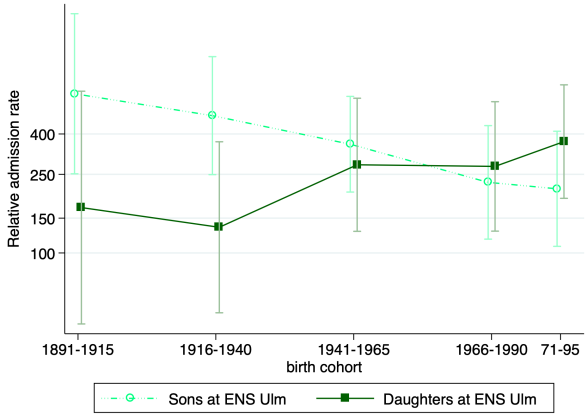
Notes: This figure reports, by birth cohort, the relative admission rates for children with a father who graduated from the same GE or GE category. Brackets refer to 95% confidence intervals. I use a logarithmic scale for the ordinate.

Figure 3: Admissions of sons *vs* daughters of *Sciences Po* graduates to *Sciences Po*, and of sons *vs* daughters of *ENS Ulm* graduates to *ENS Ulm*.

(a) *Sciences Po Paris*.



(b) *ENS Ulm*.



Notes: This figure reports, by birth cohort, the relative admission rates to *Sciences Po Paris* of sons of *Sciences Po* graduates (panel a) and the RAR to *ENS Ulm* of sons of *ENS Ulm* graduates (panel b), relative to sons of non-graduates, as compared to the same figures for daughters of graduates relative to daughters of non-graduates. Brackets refer to 95% confidence intervals. I use a logarithmic scale for the ordinate.

Table 1: Description of the dataset per *Grande École*.

Category	*Grande École*	Data coverage period	Entry of women	Total # of graduates	Average # of annual graduates	Share of native graduates
Admin. and research	Sciences Po Paris	1886-2015	1919	152,578	1,183	74%
	ENA	1946-2015	1946	7,714	112	92%
	ENS Ulm	1886-2015	1886	16,826	130	91%
Engineering	ESPCI Paris	1886-2015	1919	5,978	46	91%
	École Polytechnique	1886-2013	1972	37,823	293	89%
	Ponts et Chaussées	1886-2014	1962	13,567	105	77%
	Télécom Paris	1889-2012	1963	11,829	94	74%
	Mines Paris	1921-2012	1969	8,476	90	85%
Business	ESSEC	1905-2010	1969	20,327	185	84%
	ESCP	1906-2011	1972	28,396	261	78%

Notes: *Data coverage period* reports the earliest and latest year of admission in the data. *Entry of women* reports the date at which women were admitted to the schools on a regular basis, although there may have been earlier very sporadic appearances of one or very few women, for example during World War 1. *Average # (number) of annual graduates* is simply the *Total # (number) of graduates* during the whole period divided by the timespan in the data. *Share of native graduates* corresponds to the share of individuals bearing a "native" surname, as specified in Appendix C.3.

Table 2: Descriptive statistics per cohort: nominative census and *Grande École* registers.

		Full period		Cohorts					
		Number of surnames	Total # of births		Births 1891-1915	Births 1916-1940	Births 1941-1965	Births 1966-1990	
Census	Census	807,229	65,423,121		10,686,923	14,149,274	20,099,847	20,487,077	
	Census of "native"	541,426	59,938,195		10,574,454	13,700,315	18,655,322	17,008,104	
	Share of "native"	67%	92%		99%	97%	93%	83%	
		Number of surnames	Total # of Students 1866-1995	Students cohort 1866-1890	Students cohort 1891-1915	Students cohort 1916-1940	Students cohort 1941-1965	Students cohort 1966-1990	Students pseudo-cohort 1971-1995
9 *Grandes Écoles*	Students	63,155	141,761	8,398	13,962	20,331	36,287	57,437	53,407
	"Native" students	48,073	118,337	7,997	13,291	19,202	31,920	42,438	38,240
	Share of "native"	76%	83%	95%	95%	94%	88%	74%	72%
	Top educ %		0.18%	-	0.13%	0.14%	0.17%	0.25%	0.22%
	Women %		19%	7%	6%	6%	18%	32%	33%
Sciences Po Paris	Students	77,496	152,552	7,035	14,364	26,273	52,752	45,025	43,346
	"Native" students	49,956	113,085	5,467	11,088	21,508	39,993	30,837	28,592
	Share of "native"	64%	74%	78%	77%	82%	76%	68%	66%
	Top educ %		0.18%	-	0.10%	0.16%	0.21%	0.18%	0.17%
	Women %		33%	0%	6%	20%	36%	51%	53%

Notes: This table provides the number of births and number of students for all surnames (first lines of each section of the table) and for native surnames only (second lines), as well as the share of those native surnames (third lines). The first two columns cover the full period, while the following 6 columns cover each cohort separately. Restricting to native surnames, all columns also report the share of the population admitted (*Top educ %*), and the share of women (*Women %*). The sum of the total number of students at *Sciences Po Paris* and at the 9 schools in the baseline is necessarily higher than the total number of students reported overall (285,286), and for native students taken alone (224,264), some students having studied both at *Sciences Po Paris* and other GE. In addition, some students appear both in the 1966-1990 cohort and the 1971-1995 pseudo-cohort. The 1866-1890 cohort exclusively indicates GE-graduate ancestors.

Table 3: Admissions to any of the baseline 9 *Grandes Écoles* of children of graduates of any of the 9.

(1) Cohort	(2) Overall admiss. rate	(3) Number of surnames	(4) Number of births*	(5) Popula -tion share*	(6) Number of students*	(7) Share of students*	(8) Group admiss. rate*	(9) Relative admiss. rate**
1891-1915	0.13%	5,502	14,619	0.14%	1,766	13.3%	12.1%	**154** [127-187]
1916-1940	0.14%	8,602	31,205	0.23%	2,500	13.0%	8.0%	**81** [69-96]
1941-1965	0.17%	12,072	49,234	0.26%	4,483	14.0%	9.1%	**72** [63-83]
1966-1990	0.25%	16,651	49,542	0.29%	6,777	16.0%	13.7%	**75** [66-86]
1971-1995	0.22%	16,972	50,223	0.30%	6,503	17.0%	12.9%	**83** [73-96]

Notes: *admiss.* stands for admission. This table reports by cohort the *relative admission rates* to the baseline 9 GE for children of graduates of these 9 elite colleges, along with the *number of surnames* with at least one graduate ancestor, the *number of births* of those with a GE-graduate father, and the corresponding *population share* within the French "native" population, as defined in the text of the paper. *Share of students* consists of the share of individuals with a GE-graduate father among graduates of the 9 schools. *Group admiss. rate* is the fraction of individuals with a GE-graduate father that is enrolled in any of the 9 GE. *Overall admission rate* reports overall admissions of the French population to the 9 GE. * These figures are recomputed (details in Appendix C.5) to account for the share of births both of those with and those without a GE-graduate father and the estimated relative admission rate comparing GE graduates' descendants to the rest of the population ** 95% confidence intervals are reported between brackets.

Table 4: College of origin – college of destination matrix of intergenerational mobility, 1971-1995 cohort.

FATHER IN \ CHILDREN IN	ENS Ulm	Polytech -nique	Ponts	Mines	Telecom	ESPCI	ESCP	ESSEC	Sciences Po Paris	ENA
Cohort 1971-1995										
ENS Ulm	266 [155;457]	107 [63;180]	46 [21;101]	47 [13;164]	73 [30;178]	26 [7;94]	41 [21;83]	53 [26;107]	67 [41;109]	125 [36;436]
Polytech -nique	141 [91;220]	296 [209;420]	279 [175;447]	191 [113;322]	137 [83;225]	82 [38;174]	102 [71;146]	127 [86;188]	88 [66;119]	134 [62;291]
Ponts	103 [49;217]	166 [100;277]	216 [118;394]	87 [41;184]	76 [37;154]	17 [5;57]	105 [55;197]	87 [46;166]	81 [51;128]	157 [45;553]
Mines	121 [55;265]	195 [110;345]	181 [87;375]	328 [120;899]	87 [40;190]	63 [20;199]	63 [35;115]	94 [46;192]	78 [44;140]	157 [48;514]
Telecom	108 [47;249]	174 [100;302]	199 [90;439]	198 [86;453]	154 [67;353]	98 [28;348]	107 [57;199]	46 [24;89]	60 [36;99]	164 [46;587]
ESPCI	139 [41;473]	108 [44;264]	85 [32;225]	133 [14;1293]	55 [11;275]	365 [95;1402]	124 [35;447]	149 [42;529]	51 [20;135]	390 [44;3441]
ESCP	37 [20;68]	58 [35;94]	63 [31;126]	20 [6;61]	10 [3;32]	36 [10;138]	86 [58;128]	123 [81;186]	56 [41;77]	25 [8;72]
ESSEC	84 [46;153]	57 [35;95]	42 [20;86]	63 [30;129]	86 [39;191]	99 [30;326]	99 [67;146]	107 [71;160]	52 [37;72]	32 [12;85]
Sciences Po Paris	58 [43;79]	55 [44;69]	44 [31;63]	53 [36;78]	36 [23;54]	15 [6;36]	80 [66;98]	76 [61;95]	77 [67;88]	78 [51;120]
ENA	122 [68;219]	116 [55;243]	128 [49;334]	172 [82;359]	76 [26;224]	59 [14;260]	115 [67;199]	137 [84;222]	128 [86;190]	330 [163;669]

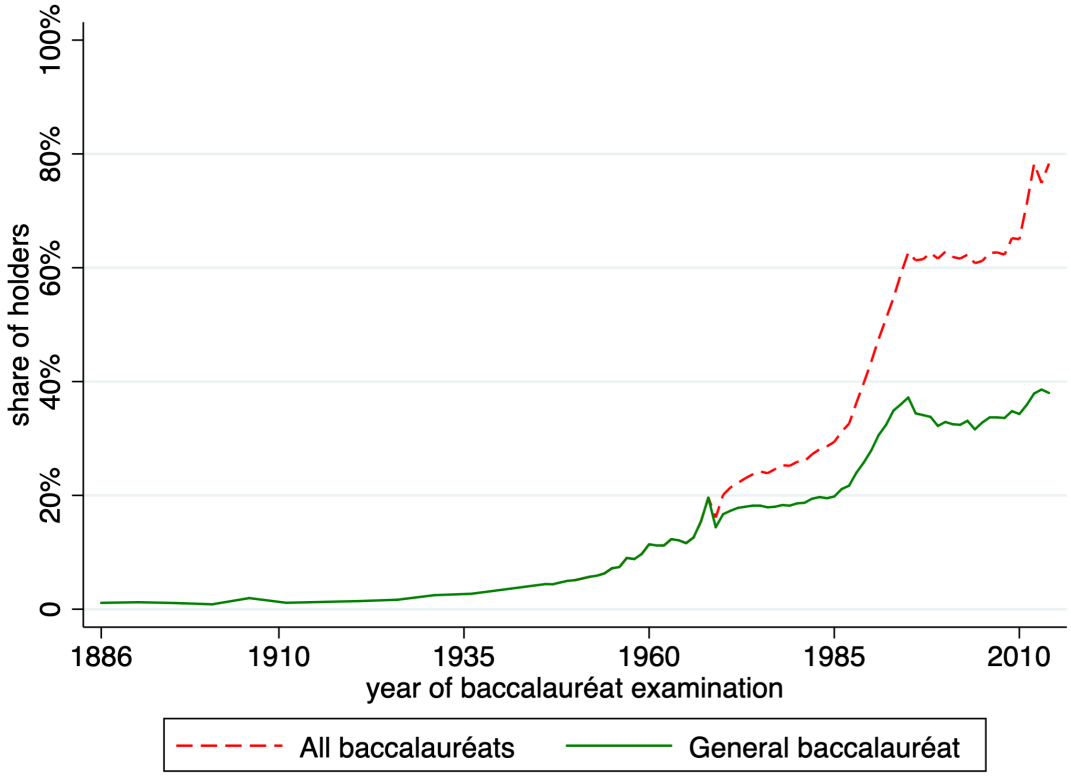
Notes: This heat matrix reports, for the 1971-1995 cohort, the relative admission rate of children of graduates to any given college in the sample (different columns) according to the father's GE (different lines). The darker the cell, the higher the RAR. 95% confidence intervals are provided between brackets below each estimate. Complementary results for previous cohorts are reported in Appendix Tables B.6a, b, c, d.

Online appendix

For Online Publication

A. Supplementary figures

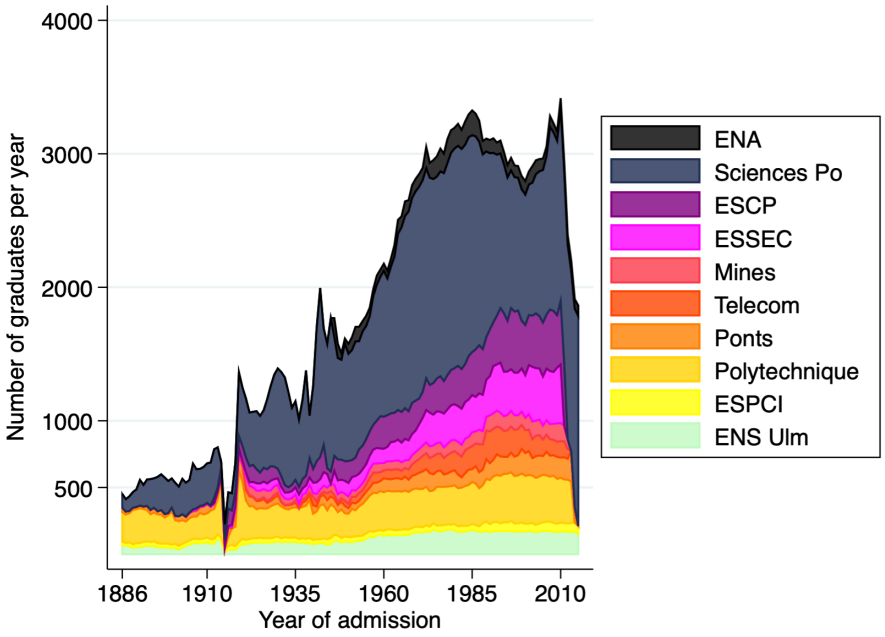
Figure A.1: Historical evolution of the share of *baccalauréat* holders, potential applicants to the *Grandes Écoles*, in the French population.



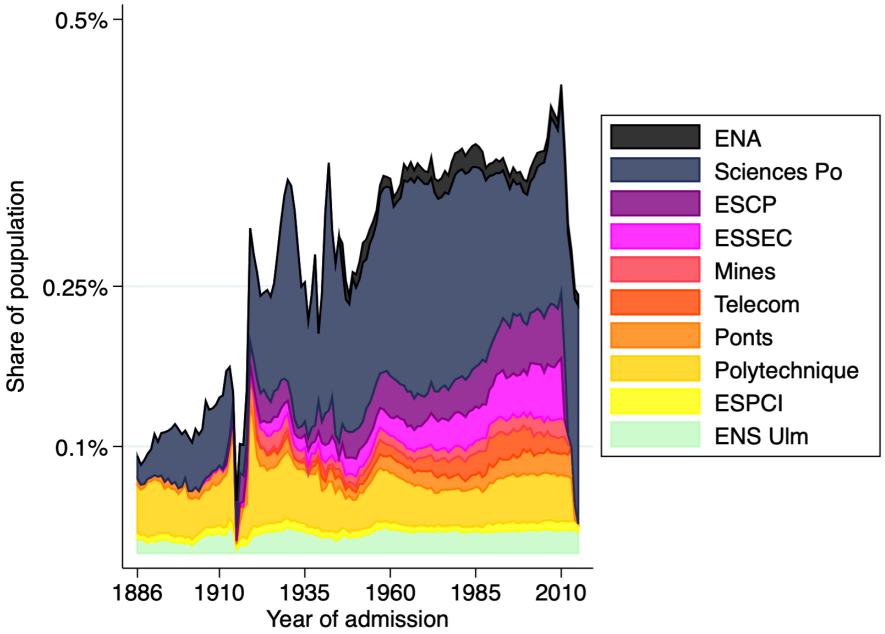
Notes: The figure reports the share of *baccalauréat* holders for each year of examination. It provides a measure of the evolution of the population legally entitled to apply to the preparatory classes for the *Grandes Écoles*: the *baccalauréat* was an entry requirement over most of the period. Up to 1949, I exploit data from the Ministry of National Education: *L'évolution du nombre des bacheliers (1851-1979)*. From 1950 to 1969, data come from the Ministry of Higher Education and Research: *Les évolutions de l'enseignement supérieur depuis 50 ans : croissance et diversification*. From 1970 onwards, I use data from `data.gouv.fr`: *La proportion de bacheliers dans une génération*. The drop in 1970 may therefore be explained by the change of data source.

Figure A.2: Evolution of size of *Grande École* year groups (1886-2015).

(a) Number of graduates.

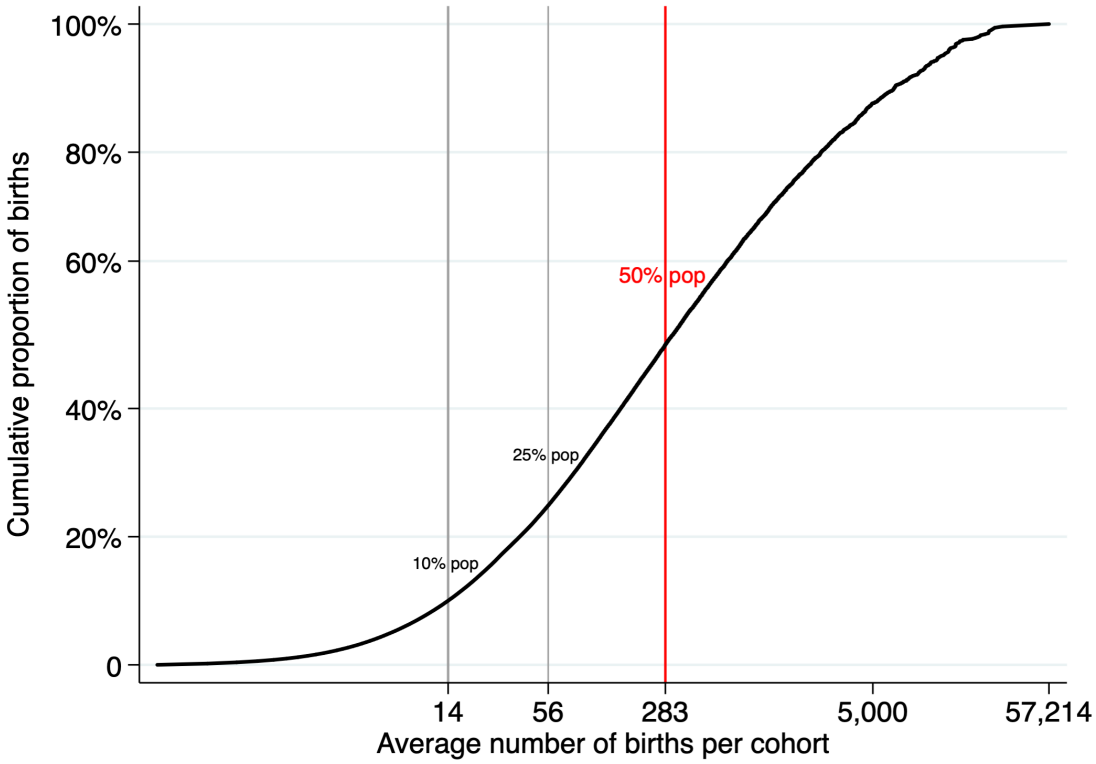


(b) Population share.



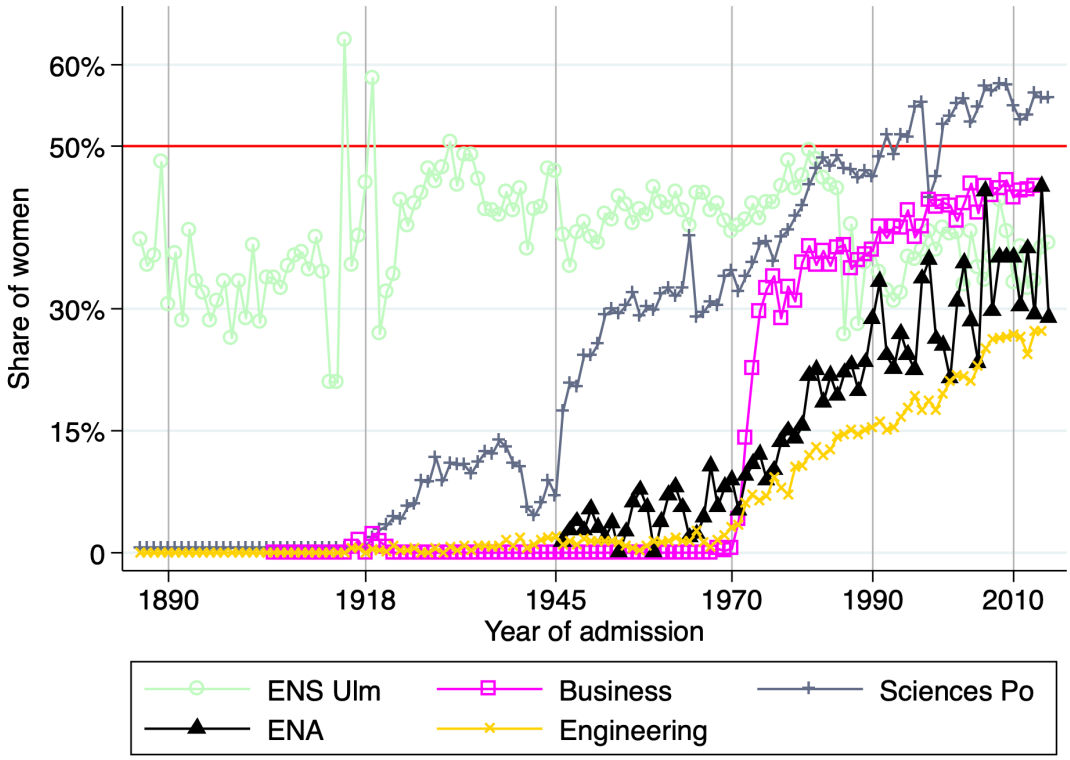
Notes: Panel (a) reports over time the evolution of the annual number of graduates in each college. Panel (b) reports over time the evolution of the share of the population annually admitted to each college. I stack for each year the number of graduates of the different colleges, only including students with "native" surnames—as defined in section 3—, the students considered in the study. For Panel (b), I then divide the number of graduates by the number of births per year at the national level, provided by the French Institute of Statistics (https://www.insee.fr/fr/statistiques/4192361).

Figure A.3: Surnames' frequency in France (1891-1990).



Notes: The figure is based on the number of births by surname per 25-year cohort averaged over the period 1891-1990. The logarithmic scale for the abscissa emphasizes the importance of rare surnames. While the average number of births per cohort for a single surname ranges from 0.5 to 57,214 (*Martin*), the figure shows that surnames with less than 14 births per cohort account for 10% of all births over the period (*10% pop* vertical line). Surnames with at most 56 births per cohort account for 25% of the population (*25% pop* vertical line), whereas half of the population born between 1891 and 1990 had a surname with less than 283 births per cohort (*50% pop* vertical line). This only includes "native" surnames, as defined in section 3. With all surnames included, the 10%, 25%, and 50% cut-offs would correspond to even rarer surnames, with averages of 11, 48, and 246 births per cohort respectively.

Figure A.4: Share of women among enrolled students, by college categories (1886-2015).



Notes: The figure reports by year of admission the share of women students at each college or college category between 1886 and 2015. A small data manipulation was used to improve readability through a narrower scale, as the share of women admitted to *ENS* in 1915 is 100% (here reduced visually to the highest point of 63%). The male side of *ENS* recruited no students that specific year. Up to the integration of the two single-gender schools in 1985, the share of women at *ENS* can be seen as constituting a ratio of relative size between the two.

B. Supplementary Tables

Table B.1: Decomposition by cohort of the contribution of each elite college to the total number of graduates in the baseline sample.

Cohort	Polytech-nique	Ponts	ESPCI	Mines	Télécom	ESCP	ESSEC	ENS Ulm	ENA
1866-1890	67%	10%	10%	0%	1%	0%	1%	13%	0%
1891-1915	45%	10%	6%	7%	2%	15%	7%	15%	0%
1916-1940	31%	7%	5%	7%	5%	19%	12%	13%	9%
1941-1965	22%	8%	3%	6%	9%	19%	17%	13%	11%
1966-1990	20%	9%	4%	7%	10%	23%	20%	10%	4%
1971-1995	21%	10%	4%	7%	9%	22%	19%	11%	4%

Notes: The table reports by cohort the fraction of the total number of graduates in the baseline sample attending each of the 9 *Grandes Écoles*. Naturally, lines sum to more than 100% because some students are counted several times if they attended multiple colleges.

Table B.2: Descriptive statistics for the variables on the probability of a GE-graduate father.

Variable	Mean (among positive)	Number of surnames	Population share	Variable	Mean (among positive)	Number of surnames	Population share
Father_at_9_GE_1891_1915	0.13	5,502	0.14%	Father_at_Sciences_Po_1891_1915	0.21	4,314	0.10%
Father_at_9_GE_1916_1940	0.14	8,602	0.23%	Father_at_Sciences_Po_1916_1940	0.21	7,666	0.18%
Father_at_9_GE_1941_1965	0.13	12,072	0.27%	Father_at_Sciences_Po_1941_1965	0.17	12,073	0.26%
Father_at_9_GE_1966_1990	0.10	16,651	0.29%	Father_at_Sciences_Po_1966_1990	0.12	16,837	0.28%
Father_at_9_GE_1971_1995	0.10	16,972	0.30%	Father_at_Sciences_Po_1971_1995	0.12	16,696	0.28%
Father_at_Polytechnique_1891_1915	0.12	4,012	0.10%	Father_at_ENA_1891_1915	0.00	0	0.00%
Father_at_Polytechnique_1916_1940	0.13	4,530	0.11%	Father_at_ENA_1916_1940	0.00	0	0.00%
Father_at_Polytechnique_1941_1965	0.11	4,625	0.09%	Father_at_ENA_1941_1965	0.10	1,527	0.03%
Father_at_Polytechnique_1966_1990	0.08	5,436	0.08%	Father_at_ENA_1966_1990	0.08	2,553	0.03%
Father_at_Polytechnique_1971_1995	0.08	5,410	0.08%	Father_at_ENA_1971_1995	0.07	2,450	0.03%
Father_at_Ponts_1891_1915	0.12	733	0.01%	Father_at_ENS_Ulm_1891_1915	0.08	909	0.02%
Father_at_Ponts_1916_1940	0.11	1,165	0.02%	Father_at_ENS_Ulm_1916_1940	0.09	1,047	0.02%
Father_at_Ponts_1941_1965	0.11	1,285	0.02%	Father_at_ENS_Ulm_1941_1965	0.10	1,274	0.02%
Father_at_Ponts_1966_1990	0.07	2,226	0.03%	Father_at_ENS_Ulm_1966_1990	0.07	1,980	0.03%
Father_at_Ponts_1971_1995	0.07	2,307	0.03%	Father_at_ENS_Ulm_1971_1995	0.08	2,118	0.03%
Father_at_mines_1891_1915	0.00	0	0.00%	Father_at_ESSEC_1891_1915	0.00	0	0.00%
Father_at_mines_1916_1940	0.14	885	0.02%	Father_at_ESSEC_1916_1940	0.13	767	0.02%
Father_at_mines_1941_1965	0.10	1,211	0.02%	Father_at_ESSEC_1941_1965	0.10	1,947	0.03%
Father_at_mines_1966_1990	0.07	1,704	0.02%	Father_at_ESSEC_1966_1990	0.08	3,622	0.05%
Father_at_mines_1971_1995	0.07	1,832	0.02%	Father_at_ESSEC_1971_1995	0.08	3,764	0.05%
Father_at_Telecom_1891_1915	0.10	50	0.00%	Father_at_ESCP_1891_1915	0.00	0	0.00%
Father_at_Telecom_1916_1940	0.10	263	0.00%	Father_at_ESCP_1916_1940	0.10	1,781	0.04%
Father_at_Telecom_1941_1965	0.09	992	0.02%	Father_at_ESCP_1941_1965	0.10	3,148	0.06%
Father_at_Telecom_1966_1990	0.07	2,328	0.03%	Father_at_ESCP_1966_1990	0.09	3,975	0.05%
Father_at_Telecom_1971_1995	0.07	2,742	0.04%	Father_at_ESCP_1971_1995	0.09	3,913	0.05%
Father_at_espci_1891_1915	0.10	678	0.01%	Father_at_Polyt_ENS_ESPCI_1891_1915	0.12	5,139	0.13%
Father_at_espci_1916_1940	0.11	700	0.01%	Father_at_Polyt_ENS_ESPCI_1916_1940	0.13	5,776	0.15%
Father_at_espci_1941_1965	0.12	707	0.01%	Father_at_Polyt_ENS_ESPCI_1941_1965	0.12	6,073	0.12%
Father_at_espci_1966_1990	0.06	798	0.01%	Father_at_Polyt_ENS_ESPCI_1966_1990	0.08	7,466	0.11%
Father_at_espci_1971_1995	0.06	838	0.01%	Father_at_Polyt_ENS_ESPCI_1971_1995	0.08	7,579	0.12%
Father_at_Engineering_1891_1915	0.13	4,851	0.12%	Father_at_Business_1891_1915	0.17	47	0.00%
Father_at_Engineering_1916_1940	0.14	6,111	0.15%	Father_at_Business_1916_1940	0.11	2,390	0.05%
Father_at_Engineering_1941_1965	0.12	6,648	0.13%	Father_at_Business_1941_1965	0.11	4,636	0.09%
Father_at_Engineering_1966_1990	0.09	8,953	0.14%	Father_at_Business_1966_1990	0.09	6,531	0.09%
Father_at_Engineering_1971_1995	0.08	9,267	0.15%	Father_at_Business_1971_1995	0.09	6,507	0.09%

Notes: The table reports descriptive statistics on explanatory variables. I include the mean value of the variable among non-null observations, the number of surnames with non-null values, as well as the population share concerned by each characteristic. The population share is estimated as detailed in Appendix C.5. The last eleven characters of each variable name indicate the birth cohort. Variables relate to fathers graduating from a college or group of colleges, including the 9 baseline *Grandes Écoles* (9GE) and the colleges with stable student populations (Polyt_ENS_ESPCI).

Table B.3: Complementary results: robustness to the inclusion of *Sciences Po*. Admissions to any of the 10 *Grandes Écoles* of children of graduates from any of the 10 *Grandes Écoles*.

Cohort	Overall admiss. rate	Number of surnames	Number of births*	Popula -tion share*	Number of students*	Share of students*	Group admiss. rate*	Relative admiss. rate**
1891-1915	0.23%	8,806	24,581	0.33%	5,767	24.3%	23.5%	**132** [119-148]
1916-1940	0.29%	14,276	55,174	0.55%	9,765	17.2%	17.7%	**59** [53-66]
1941-1965	0.37%	20,289	92,984	1.03%	11,539	26.1%	12.4%	**53** [48-58]
1966-1990	0.41%	27,341	92,616	1.02%	13,301	35.7%	14.4%	**55** [50-60]
1971-1995	0.38%	27,613	92,525	1.02%	13,920	37.3%	15.0%	**60** [55-66]

Notes: *admiss.* stands for admission. This table reports by cohort the relative admission rates to the 10 *Grandes Écoles* for children of graduates from any of these 10 elite colleges, along with the *number of surnames* with at least one graduate ancestor, *number of births* of those with a GE-graduate father, and the corresponding *population share* within the French "native" population, as defined in the text of the paper. *Share of students* consists of the share of individuals with a GE-graduate father among graduates of the 10 schools. *Group admiss. rate* is the fraction of individuals with a GE-graduate father that is enrolled in any of the 10 GE. The *overall admission rate* to the 10 colleges in the general population is also listed. * These figures are recomputed (details in Appendix C.5) to account both for the share of births of those with versus without a GE-graduate father and the estimated relative admission rate between descendants of GE graduates and the rest of the population. ** 95% confidence intervals are reported between brackets.

Table B.4: Complementary results: robustness of intergenerational mobility estimates.

Cohort	(1) Main result (recalled)	(2) Inclusion of immigrant surnames	(3) 50% pop. with rarer surnames	(4) 30% pop. with rarer surnames	(5) 20% pop. with rarer surnames	(6) surnames ≤100 male births	(7) surnames ≤50 male births	(8) surnames ≤25 male births
1891-1915	**154** [127-187]	**142** [118-171]	**161** [132-195]	**157** [128-191]	**151** [123-185]	**161** [132-195]	**158** [129-192]	**156** [127-190]
1916-1940	**81** [69-96]	**57** [45-72]	**86** [73-102]	**84** [71-100]	**80** [67-95]	**86** [72-102]	**85** [71-100]	**83** [69-98]
1941-1965	**72** [63-83]	**57** [47-69]	**77** [67-88]	**74** [64-85]	**69** [60-80]	**76** [66-87]	**74** [64-85]	**70** [60-80]
1966-1990	**75** [66-86]	**36** [28-46]	**83** [72-94]	**78** [68-89]	**71** [61-81]	**81** [71-92]	**78** [68-89]	**72** [63-83]
1971-1995	**84** [73-96]	**60** [50-72]	**92** [80-104]	**86** [75-98]	**78** [67-89]	**90** [78-102]	**87** [75-99]	**79** [69-91]

Notes: *pop.* stands for population. This table reports by cohort the main measure of social reproduction, i.e., relative admission rates to the baseline 9 *Grandes Écoles* for children of graduates of these 9 schools. 95% confidence intervals are reported between brackets. Column (1) lists the main result from Table 3. Column (2) provides similar estimates for the full sample of students, including bearers of immigrant surnames. Columns (3-8) report results on sub-samples of rare surnames. Column (3) includes rare surnames accounting for 50% of the total population, column (4) restricts to 30% of the population with rarer surnames, and column (5) to 20% of the population. Column (6) restricts to surnames with at most 100 male births in the cohort of interest, while columns (7) and (8) restrict to individuals bearing surnames with at most 50 and 25 male births per cohort respectively. The rarer the surnames, the finer the lineage tracking. Rarer surnames are also associated with higher social status. For instance, the 50% of the population with rarer surnames is 1.135 times more likely to be admitted to a *Grande École* for the cohort born in 1891-1915. Therefore, for each restriction on rare surnames, I multiply the estimates by the over-representation of the sub-sample of rare surnames. Trends and orders of magnitude of estimates from robustness tests are very comparable to the main result. Only for the inclusion of "immigrant" surnames do I find notable discrepancies for the more recent cohorts when the *Grandes Écoles* started to admit more international students. For these surnames, I incompletely track the number of births per cohort and find newly appearing names in schools' registers, without properly relating them to a population size. Mechanically, the measured relative admission rates are slightly lower when these mismeasurements are included.

Table B.5: Complementary results: admissions to *Polytechnique*, *ENS* or *ESPCI* of children of graduates of the same schools.

		Fathers in *Polytechnique* / *ENS* / *ESPCI*		
Cohort	Overall admiss. rate	Popula -tion share*	Share of students*	Relative admiss. rate**
1891-1915	0.08%	0.13%	15.66%	**210** [169-261]
1916-1940	0.07%	0.14%	12.51%	**121** [95-153]
1941-1965	0.07%	0.12%	11.89%	**137** [111-169]
1966-1990	0.08%	0.11%	12.42%	**158** [122-204]
1971-1995	0.08%	0.11%	13.28%	**181** [140-233]

Notes: This table reports by cohort the population share (*pop. share*), share of students (*share of stud.*) and relative admission rates (RAR) to *École Polytechnique*, *ENS Ulm* or *ESPCI* for children of graduates of these same colleges. The *overall admission rate* to these schools is also given, and is very stable across the period. * As noted in Appendix Table B.3, technical details on the computations are presented in Appendix C.5. ** 95% confidence intervals are reported between brackets. 10% confidence intervals are as follows: 1891-1915 : 210 [175;252] ; 1916-1940 : 121 [99;148] ; 1941-1965 : 137 [115;163] ; 1966-1990 : 158 [127;195] ; 1971-1995 : 181 [146;223].

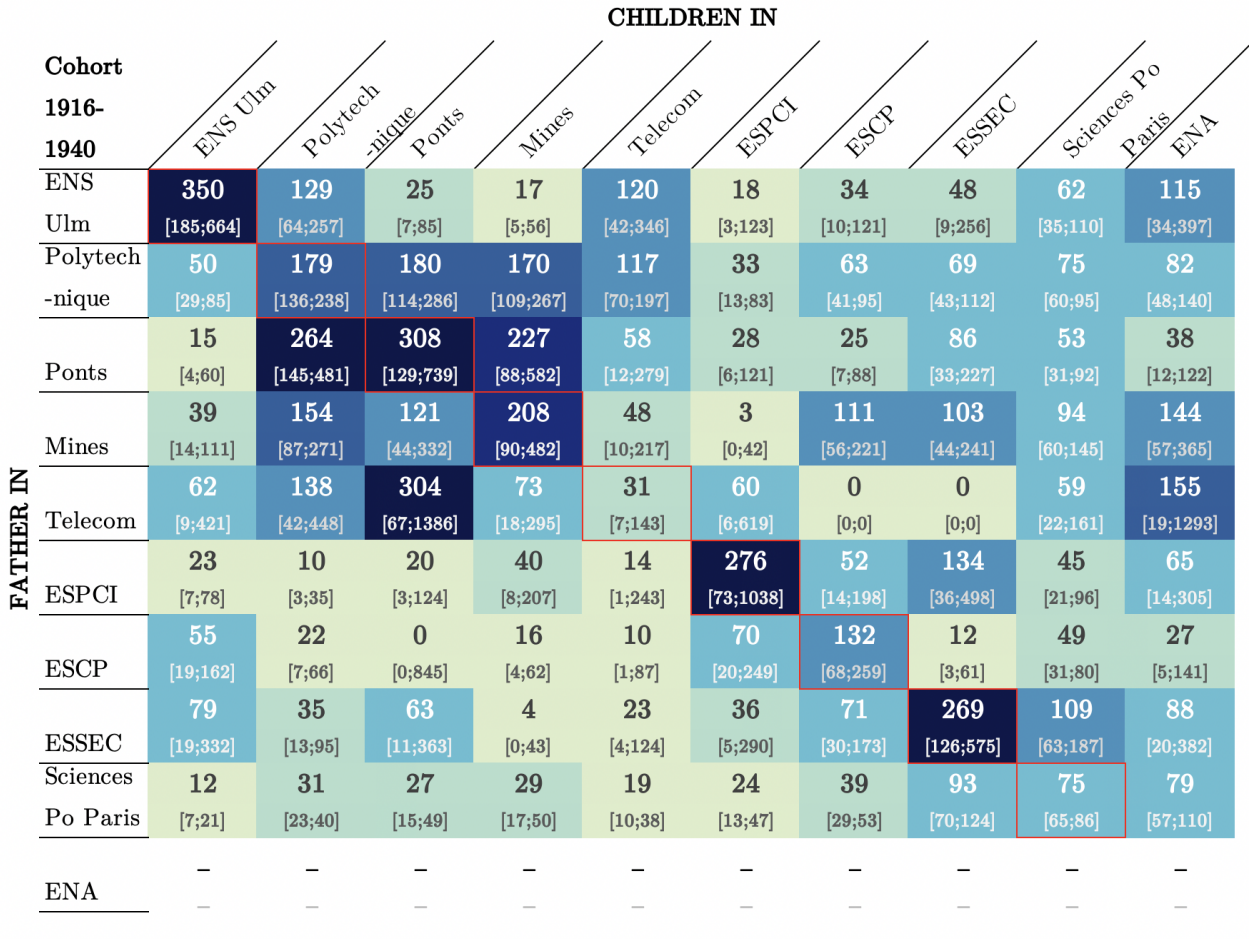
Table B.6: Complementary results: college of origin – college of destination matrices of intergenerational mobility.

(a) 1891-1915 cohort.

Cohort 1891-1915	CHILDREN IN: ENS Ulm	Polytech-nique	Ponts	Mines	Telecom	ESPCI	ESCP	ESSEC	Sciences Po Paris	ENA
FATHER IN: ENS Ulm	458 [197;1064]	79 [38;164]	38 [13;108]	327 [117;918]	430 [65;2859]	47 [13;169]	6 [0;88]	0 [0;0]	135 [73;251]	–
Polytech-nique	56 [29;106]	309 [240;398]	267 [170;421]	396 [244;645]	264 [48;1455]	174 [78;390]	17 [6;48]	81 [41;159]	191 [152;240]	–
Ponts	6 [1;29]	180 [95;342]	367 [139;969]	434 [125;1505]	0 [0;0]	19 [4;82]	4 [0;36]	102 [22;486]	214 [129;356]	–
Mines	–	–	–	–	–	–	–	–	–	–
Telecom	–	–	–	–	–	–	–	–	–	–
ESPCI	2 [0;244]	46 [12;176]	39 [9;161]	0 [0;0]	2 [0;1585]	1166 [518;2627]	107 [25;448]	5 [0;108]	48 [17;138]	–
ESCP	–	–	–	–	–	–	–	–	–	–
ESSEC	–	–	–	–	–	–	–	–	–	–
Sciences Po Paris	12 [4;37]	59 [44;80]	67 [36;125]	145 [85;246]	4 [0;25]	21 [9;50]	20 [9;45]	165 [102;269]	245 [211;285]	–
ENA	–	–	–	–	–	–	–	–	–	–

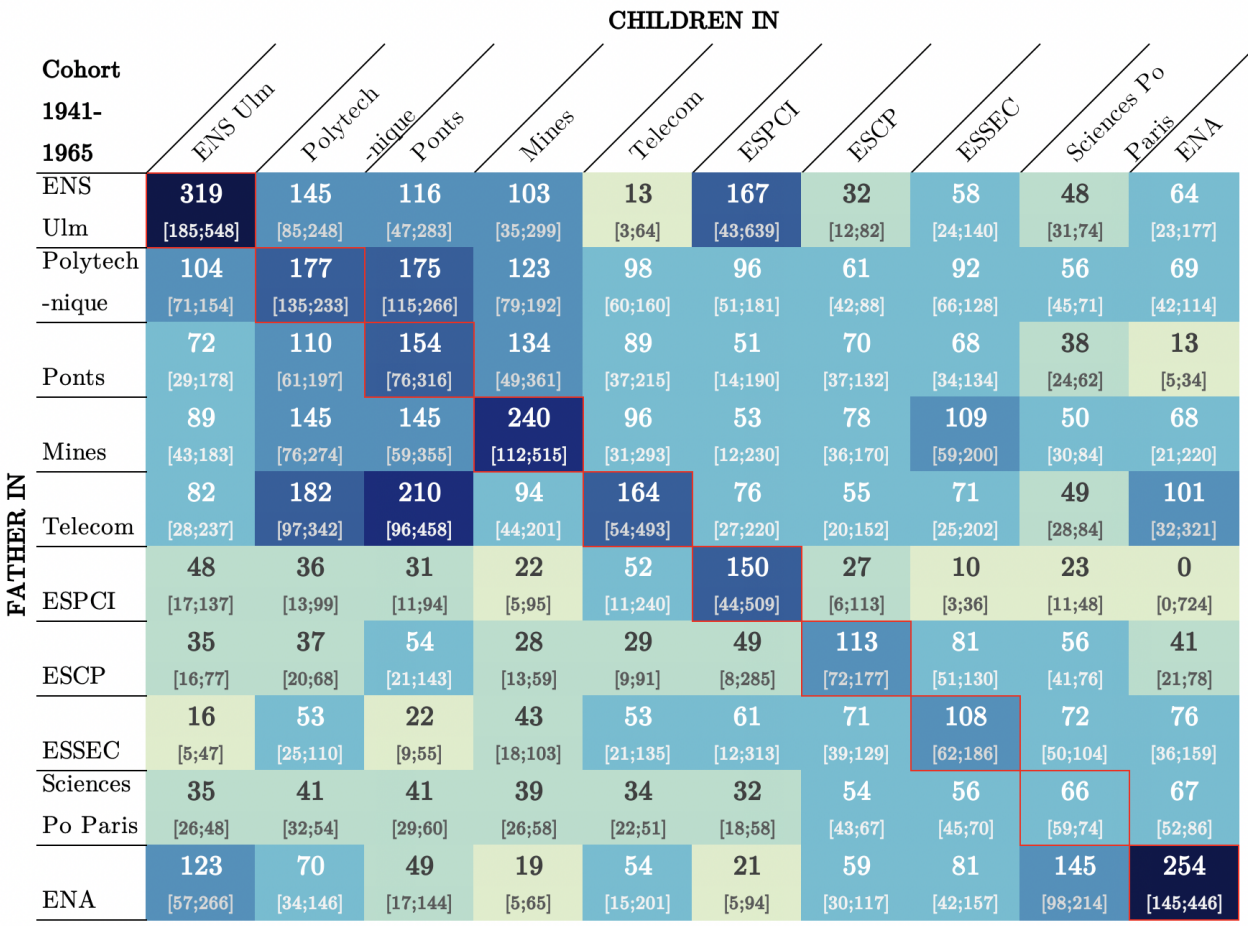
Notes: This heat matrix reports, for the 1891-1915 cohort, the relative admission rate of children of GE graduates to any given college in the sample (different columns) according to the father's GE (different lines). The darker the cell, the higher the RAR. 95% confidence intervals are provided between brackets below each point estimate. The association with paternal schooling is not available for *Mines*, *ESCP*, *ESSEC*, *ENA*, as there was no graduate from these schools in the first ancestors' cohort (born between 1866-1890). There were also too few students at *Télécom* to provide relevant estimations. About noone born between 1891 and 1915 was admitted to *ENA*, so the last column is empty.

(b) 1916-1940 cohort.



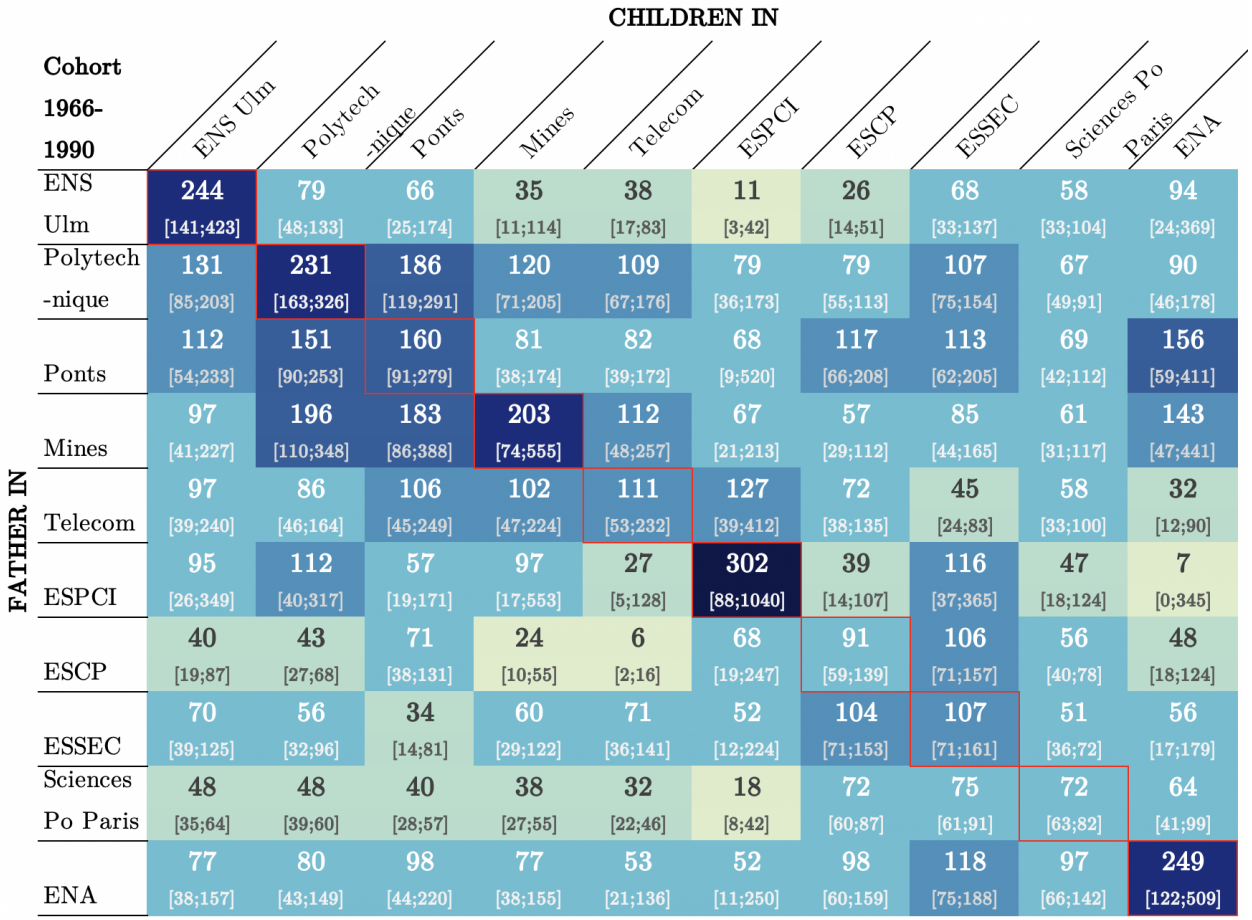
Notes: This heat matrix reports on the 1916-1940 cohort. The reading is similar to Appendix Table B.6a. About noone born between 1891 and 1915 was admitted to *ENA*, so the last line is empty.

(c) 1941-1965 cohort.



Notes: This heat matrix reports on the 1941-1965 cohort. The reading is similar to Appendix Table B.6a.

(d) Cohort 1966-1990.



Notes: This heat matrix reports on the 1966-1990 cohort. The reading is similar to Appendix Table B.6a.

Table B.7: Complementary results: admissions of sons *vs* daughters of graduates:

(a) to a given college, from which the father graduated.

Grande Ecole		1891-1915	1916-1940	1941-1965	1966-1990	1971-1995
Sciences Po Paris	sons	**204** [179-234]	**72** [63-83]	**62** [55-71]	**76** [65-89]	**87** [75-103]
	daughters	**194** [121-313]	**77** [62-96]	**71** [61-82]	**67** [57-78]	**66** [56-78]
ENA	sons	**-** -	**-** -	**265** [147-477]	**265** [114-616]	**290** [124-678]
	daughters	**-** -	**-** -	**170** [63-458]	**213** [73-621]	**416** [144-1204]
ENS Ulm	sons	**639** [251-1624]	**495** [249-985]	**355** [203-620]	**228** [117-441]	**211** [108-413]
	daughters	**170** [44-658]	**135** [50-365]	**279** [129-607]	**274** [129-583]	**366** [189-709]
ESPCI	sons	**1190** [544-2602]	**305** [75-1242]	**34*** [8-153]	**435** [122-1551]	**555** [145-2123]
	daughters	**-** -	**50** [6-435]	**834*** [215-3234]	**37** [6-242]	**5** [0-7054]
Polytechnique	sons	**266** [210-338]	**173** [132-226]	**161** [123-212]	**219** [153-314]	**272** [187-395]
	daughters	**-** -	**-** -	**476** [200-1130]	**246** [150-403]	**361** [229-568]
Ponts	sons	**348** [134-900]	**304** [128-719]	**135** [63-291]	**132** [70-246]	**195** [96-396]
	daughters	**-** -	**-** -	**410** [87-1924]	**273** [122-613]	**287** [130-630]
Télécom	sons	**-** -	**31** [7-143]	**171** [56-521]	**108** [47-244]	**158** [64-388]
	daughters	**-** -	**-** -	**59** [19-189]	**121** [37-396]	**120** [39-366]
Mines Paris	sons	**-** -	**206** [89-474]	**245** [112-532]	**180** [51-636]	**292** [82-1035]
	daughters	**-** -	**-** -	**146** [43-493]	**275** [102-740]	**418** [155-1128]
ESSEC	sons	**-** -	**263** [125-555]	**112** [61-204]	**125** [78-203]	**93** [58-148]
	daughters	**-** -	**-** -	**72** [25-202]	**84** [45-154]	**123** [70-217]
ESCP	sons	**-** -	**130** [67-252]	**113** [69-185]	**88** [53-146]	**87** [54-141]
	daughters	**-** -	**-** -	**108** [47-247]	**93** [51-169]	**86** [49-149]

Notes: This table reports by cohort (columns) the relative admission rate to each *Grande École* of sons versus daughters (lines) with fathers who graduated from this same college. I relate admission rates of sons of graduates to those of sons of non-graduates and compare this to the admission rates of daughters of graduates relative to daughters of non-graduates. 95% confidence intervals are provided between brackets to the right of each point estimate in bold. * The stars identify significant differences between sons and daughters, only pertinent to *ESPCI* graduates born in 1941-1965; however, this result should be interpreted with caution due to the very small number of students at *ESPCI* (1,016 for this cohort, only 160 of them women).

(b) to a given college, the father having graduated from any of the baseline 9 GE.

Grande Ecole		1891-1915	1916-1940	1941-1965	1966-1990	1971-1995
Sciences Po Paris	sons	**136*** [111-166]	**62** [52-73]	**55** [47-64]	**53** [44-64]	**63** [52-76]
	daughters	**363*** [224-588]	**83** [64-108]	**64** [54-76]	**59** [48-71]	**63** [53-76]
ENA	sons	**-** -	**68** [45-103]	**64** [47-86]	**84** [52-134]	**84** [48-148]
	daughters	**-** -	**100** [12-827]	**87** [49-153]	**73** [36-147]	**84** [37-190]
ENS Ulm	sons	**123*** [70-217]	**80** [53-121]	**83** [60-114]	**77** [57-104]	**72** [53-97]
	daughters	**24*** [9-59]	**35** [18-68]	**85** [59-121]	**98** [68-142]	**130** [91-185]
ESPCI	sons	**258** [148-449]	**43** [22-85]	**61** [36-102]	**60** [31-115]	**47** [25-87]
	daughters	**-** -	**35** [7-179]	**153** [59-397]	**84** [44-161]	**100** [53-190]
Polytechnique	sons	**207** [166-259]	**121** [96-153]	**96** [79-117]	**99** [79-124]	**114** [90-143]
	daughters	**-** -	**-** -	**180** [86-377]	**147** [103-210]	**162** [114-230]
Ponts	sons	**214** [143-321]	**122** [83-181]	**89** [65-121]	**87** [63-119]	**98** [70-135]
	daughters	**-** -	**-** -	**228** [101-511]	**135** [81-225]	**169** [102-281]
Télécom	sons	**214** [51-906]	**72** [46-114]	**65** [46-92]	**52** [38-72]	**64** [45-91]
	daughters	**-** -	**-** -	**53** [17-167]	**77** [41-143]	**96** [47-196]
Mines Paris	sons	**296** [192-456]	**111** [76-163]	**82** [58-116]	**75** [51-110]	**98** [65-146]
	daughters	**-** -	**-** -	**71** [35-148]	**71** [40-126]	**93** [52-169]
ESSEC	sons	**65** [35-122]	**77** [55-108]	**73** [58-92]	**95** [75-120]	**90** [69-117]
	daughters	**-** -	**-** -	**82** [53-127]	**86** [68-110]	**102** [79-132]
ESCP	sons	**23** [10-52]	**68** [51-91]	**66** [51-85]	**84** [67-105]	**91** [71-116]
	daughters	**-** -	**-** -	**49** [29-80]	**71** [55-92]	**72** [56-93]

Notes: This table reports on admissions to each *Grande École* of sons versus daughters, whose fathers graduated from any of the baseline 9 *Grandes Écoles*. The reading is similar to Appendix Table B.7a. Significant differences are observed only for sons *vs* daughters of graduates born in 1891-1915 in admission to *Sciences Po Paris* and to *ENS Ulm*.

C. Supplementary Information

C.1 Supplementary contextual details on the *Grandes Écoles* sample

This appendix section provides additional information, complementary to section 2 on the institutional context and to section 3 on the dataset. I precede the presentation of each college's particular features with more general aspects that are common to several colleges.

The sample consists of the GE that are particularly involved in training the French elite—with the notable exception of *HEC Paris* and *École Centrale Paris*, for which I did not manage to collect data. The oldest college in the sample is *École nationale des Ponts et chaussées*, which dates back to 1747, and the most recently founded is *École Nationale d'Administration* (1945). As discussed in the paper, the 10 colleges work together as a system of elite training, but each has its own particular features. I broadly classified the colleges into three categories: schools of administration and research, engineering schools, and business schools. While the two latter categories are self-evident, the first could be challenged. If *Sciences Po Paris* and *ENA* are strongly linked—the former serving as a preparatory school for the latter[34]—, *ENS Ulm* remains unique in many aspects. Tables 4 and B.6 suggest that, in terms of intergenerational dynamics, *ENS* has more of an affinity with engineering schools than with *Sciences Po* or *ENA*.

As outlined in section 2, a fundamental characteristic of the *Grandes Écoles* is the admission process through highly competitive examinations, called *concours*. They take the form of written tests as a first screening, followed by oral examinations and interviews for those eligible (*admissible*). The exams follow two years of a dedicated post-secondary school preparatory program—*classes préparatoires aux Grandes Écoles*.

Tuition fees used to be the exception until recent decades, being negligible at the beginning of the period of study in most institutions except business schools. There are no fees at *ENA*, and they remain very limited at *ENS Ulm*. Studying at *Sciences Po Paris* was also inexpensive until the late 1980s, after which there were continuous fee increases, although combined with substantial grant opportunities: depending on household resources, annual fees today range from 0 to 18,000€ with an average of around 6,000€. The five engineering

[34] 53% of *ENA* students in the sample had studied at *Sciences Po*.

schools of the sample are public institutions with limited tuition fees—historically almost free and costing approximately 2,500€ per year in recent years—, except for ESPCI, where studies remain fully subsidized. Like almost all business schools in France, the two in the sample are private institutions, and have always had tuition fees.[35] They currently average around 15,000€ annually. Reductions and grants may however be provided conditional on household resources. Another special case is that of students of *ENA*, *ENS*, and *École Polytechnique*, who have the status of civil-servant trainees. This entitles them to a monthly payment during their education, usually slightly over the minimum wage, translating in recent years into about 16,000€ annually.

Finally, a crucial structural change across the 20$^{\text{th}}$ century in tertiary education was the slow adoption of the practice of admitting women. This is highlighted for the French *Grandes Écoles* by Appendix Figure A.4, which reports for relevant college categories the share of female students by year of admission. Only *École Normale Supérieure* admitted women across the whole period—though in separate single-gender schools until 1985. Several surges in the admission of women can be observed, the first during World War 1, when many replaced men in the labor force, and when women also sporadically attended some *Grandes Écoles*. Subsequently, women were admitted to *Sciences Po Paris—École libre des sciences politiques* as it was known until 1945—, although the *baccalauréat* was required for women and not for men. The Second World War saw a second surge, whose tangible effects remained, however, limited, except at *Sciences Po Paris*. There, a differing admission examination according to gender was introduced in 1941, and women's admissions consequently collapsed during World War 2. 1945 was a pivotal year, with French women finally gaining voting rights. Now taking the same examination as men, their share at *Sciences Po* began to expand. 1945 was also the year *ENA* was founded, and women were admitted right away, although under different conditions (for instance not for all majors): they constituted less than 10% of *ENA*'s student body until 1970.

The last and most general surge came in the early 1970s, when women were finally granted legal access to all *Grandes Écoles*. The share of women in business schools increased very rapidly, reaching slightly below 50% in the early 2010s. The numbers of women at *Sciences Po* rose sharply once more in 1971 and kept increasing, reaching around 55% since 2000. Women made much slower progress into engineering schools, where they had barely

[35] Universities have, since the 1950s, been a public competitor for business schools, with the almost free curriculum offered by the *Institut d'Administration des Entreprises* (*IAE*).

reached 30% by 2015; this was true to a lesser extent at *ENA*, where only 3 graduating classes had more than 36% of women. Interestingly, the number of women at *École Normale Supérieure* suddenly declined in 1985, the year that saw the merger of the female *École Normale Supérieure de jeunes filles* and the male ENS Ulm—Ferrand et al. (1999) describe the admission of women to *ENS* during the period 1985-1990—. This persistent under-representation of women in most *Grandes Écoles*, even in the early 21st century, contrasts with the reversed gender gap in access to higher education observed in most developed countries, including France (Buchmann and DiPrete, 2006, 2013).

The following paragraphs provide contextual details specific to each college in the sample.

Grandes Écoles for administration and research.

Sciences Po Paris. *Sciences Po*—originally the *École libre des sciences politiques*—was founded in 1872 to train a new political elite, as the one in place was blamed for dragging France into an unwinnable war against Prussia (Suleiman, 1978). Its founder Émile Boutmy initially designed the school as a liberal private institution as opposed to the traditional model of the *Grandes Écoles*. Yet the college shares many characteristics with the other GE. In a momentum including the foundation of *École Nationale d'Administration*, *Sciences Po* was partially nationalized in 1945 and divided into two distinct institutions, operating side by side since. The *Institut d'études politiques de l'université de Paris* is a public institution in charge of education. The *Fondation nationale des sciences politiques* is a private institution that manages administrative and financial matters. Since 2001, admission is partially through a dedicated affirmative action process, targeting pupils from educational priority areas. The college appears as a pioneer in the movement towards more equality of opportunities. Students are trained in many different disciplines, including political science, humanities, law, sociology, economics, and history. Professional training has progressively emerged, notably in journalism, management, urbanism, and communication. Women were admitted to *Sciences Po* in 1919 for the first time.

École Nationale d'Administration (ENA). After several prior attempts, notably one by the *Front Populaire*, *École Nationale d'Administration*—a public college—was founded in 1945 to train senior civil servants, including women from the start. It has a particular status, being attended at a slightly older age, usually after studies in another *Grande École*. Students are civil-servant trainees and receive payment during their education. Until 1978, *ENA* was hosted in a Parisian building owned by the *Fondation nationale des sciences*

politiques, literally only separated by one garden from the facilities of the political science school. This illustrates the proximity of the two institutions, with *Sciences Po* designing specific preparatory programs for the *ENA concours*. In 1991, *ENA* was relocated to Strasbourg. About 100 students are trained each year. They are ranked at the end of studies, and the 15 best-ranked students (called *la botte*) may directly choose their assignment in the public service, and especially their *grands corps*.[36]

École Normale Supérieure (ENS Ulm). *École Normale Supérieure* was founded in 1794. This public institution has been located in rue d'Ulm in Paris since 1841, hence the usual reference to *École Normale Supérieure de la rue d'Ulm* to distinguish it from the other *ENS* in Cachan, Lyon, or Fontenay. Its mission is to provide an academic curriculum of excellence in science or humanities and train researchers and teachers. Since 1948, students have had the status of civil-servant trainee and are expected to spend at least 10 years serving the State after graduation—although this is not fully enforced. They receive a monthly payment during their years of education, counted as part of the 10-year service. In 1985, the school merged with *École Normale Supérieure de jeunes filles*, dedicated to training female teachers. The latter school was founded in 1881 and was located in Sèvres until the German occupation of World War 2, moving to Paris in 1948.

Engineering schools.

École supérieure de physique et de chimie industrielles de la ville de Paris (ESPCI Paris). The municipality of Paris founded *ESPCI Paris* in 1882 and has remained the supervisory institution since. The school is sometimes called *l'école des Prix Nobel* because, although graduating classes are very small, six Nobel-prize laureates worked there: Marie and Pierre Curie, Frédéric and Irène Joliot-Curie, Pierre-Gilles de Gennes, and Georges Charpak. Although Marie Curie produced her research with her husband in the facilities of the college, the first female students were admitted only in 1919, still much ahead of other engineering schools. Students of *ESPCI* receive a general science education both in physics and in chemistry—as well as in biology since 1994—, before choosing a specialization in the last phase of the program. This pluridisciplinary approach was always a special feature of the college. There was never any tuition fee in this public college.

École Polytechnique. *École Polytechnique* is among the most prestigious colleges in

[36] See Suleiman (1978) for a comprehensive study of the *grands corps*, which are official civil-servant groups with corresponding status, positions, and salaries.

the world, and is usually referred to simply as "X", an allusion to the mathematical symbol and the crossed cannon barrels of its military logo. The school was founded in 1794 and still has the military status granted by Naopléon Bonaparte in 1804. It was instituted as a prerequisite for entry to *École des Ponts* or *École des Mines de Paris*, which served as *écoles d'application* (schools of applied engineering). Although the latter colleges restored direct accessibility (see below), *Polytechnique* has always provided a more general and "*polytechnician*" curriculum, and its students still often spend one year of specialization in in *Ponts*, *Mines*, *Télécom*, or another *école d'application*. The initial aim of *École Polytechnique* was more the dissemination of science, and its graduates generally joined the public sector. After World War 2, the college added a stated objective of training the industrial elite. Studying at *Polytechnique* is, for instance, the most promising route to becoming either an administrator of the French Institute of Statistics (INSEE) or CEO of one of the major French companies, as mentioned in section 2. Located in Paris until 1976, the institution moved to a campus in Palaiseau, in the Parisian suburbs. Women have been admitted since 1972.

École nationale des Ponts et chaussées. *École nationale des Ponts et chaussées* was founded in 1747. As its name suggests, it was designed to train engineers for the construction and development of bridges (*ponts*) and roads (*chaussées*), and more generally for town and country planning. Between 1795 and 1848, the college only admitted students after their studies at *Polytechnique*, and provided practical training. Since then, the college has continued to take in *Polytechnique* students for a one-year specialization, but has also reestablished its own full engineering track. Like *École Polytechnique*, *Ponts*' training of engineers became more oriented towards the private sector after World War 2. Women have been admitted since 1962.

Télécom Paris. The college was founded as the *École supérieure de télégraphie* in 1878 when the French government set up its Post and Telegraph administration. It was located in Paris until 2019, when it moved to Palaiseau, near *École Polytechnique*. Studies focus on communications and networks, with increasing importance given to computer sciences. New fields of study have been added since the last decades of the 20$^{\text{th}}$ century, even including a dedicated program in economics. Women have been admitted since 1963.

École des Mines de Paris. This is one of the oldest *Grandes Écoles*, founded in 1783. Its original mission was to train directors for the booming mining industry. The primary fields of study have inevitably evolved to include energy and raw materials. The college's

facilities are in Paris, and women have been admitted since 1969.

Business schools.

École supérieure des sciences économiques et commerciales (ESSEC). *ESSEC* was founded in 1907 by Jesuits. In its early years, the college underwent several crises: it had to close temporarily during World War 1, due to an insufficient number of students, and was impacted by the crisis of the 1930s, when fewer could afford its relatively high tuition fees. Until the 1960s, law constituted a major share of the curriculum, which also comprised trade, languages, accounting, and political economy. The college was in Paris, under the supervision of the Parisian Catholic Institute. In the early 1970s, *ESSEC* gained some degree of autonomy and moved to Cergy, one of the *villes nouvelles* (new towns), in the Parisian suburb. After new financial difficulties, the college was saved in 1980 by the Chamber of Commerce of Versailles, which became its new supervisor. The admission concours was instituted only in the 1940s, marking out the college slightly from the others in the sample until that period. Indeed, it was not open to those in public preparatory classes until 1951. Women have been admitted since 1969.

École Supérieure de Commerce de Paris (ESCP). Founded in 1819 by two merchants, and often associated with early patronage by Jean-Baptiste Say, *ESCP* is considered the doyenne of worldwide business schools. The school was bought by the Chamber of Commerce of Paris in 1869, at a time when regional chambers of commerce were founding their own business schools, e.g., in 1872 for Lyon, Marseille, and Lille. Studies were highly oriented towards trade, including merchant shipping or the hospitality trade. The college remained located in Paris, with the addition of new European campuses in recent decades. Indeed, in 1973, the Chamber of Commerce of Paris also founded the *European School of Management*, known by its French acronym *EAP*. This school—which had campuses in France, Germany, United Kingdom, and Spain—merged in 1999 with *ESCP*, reinforcing the international orientation of *ESCP*. Women have been admitted to *ESCP* since 1972.

Sources: This set of information predominantly relies on the institutional presentations available on the GE websites. I added information taken from Suleiman (1978) for several colleges, Belhoste (2002) and Picon et al. (1994) for *École Polytechnique*, as well as Passant (2020) for *ESCP*.

C.2 Conditions identifying graduates with multiple curricula

With data at the curriculum level, it is important to identify students with multiple degrees, so that they are taken into account only once when examining admissions to a pool of colleges—like the baseline 9 GE. This appendix section details the identification procedure.

I identify distinct curricula as being followed by a single individual if one of the four following conditions applies.

1. Observations share the same non-missing first name, spouse's name, and patronym. In addition, there is at most a 9-year gap between admission to two curricula (or 24 years between admission to any other school and a later admission to *ENA*).

2. Observations share the same non-missing spouse's name or patronym, as well as the same first name and 2 middle names (first, second, and third given names are non-missing and similar). There is at most a 9-year gap between admission to two curricula (or 24 years between admission to any other school and a later admission to *ENA*).

3. Observations share the same non-missing spouse's name or patronym, as well as the same set of first name and one middle name (first and second given names are non-missing and similar). There is at most an 8-year gap between admission to two curricula (or 19 between admission to any other school and a later admission to *ENA*).

4. Observations share the same non-missing spouse's name or patronym. They also share at least two names among first name and middle names (there are two common given names among the list of first, second, third, and sometimes fourth given names). In addition, at least one of the following conditions (a), (b), (c), or (d) applies.

   (a) There is at most an 8-year gap between admissions (or 19 years for a later admission to *ENA*). There are less than 10,000 births over 1891-1990 for the surname, which is common to the distinct curricula—be it the spouse's name or the patronym.

   (b) There is at most a 4-year gap between admissions, or 9 years for a later admission to *ENA*. There are less than 20,000 births over 1891-1990 for the surname, which is common to the distinct curricula—be it the spouse's name or the patronym.

(c) There is at most a 3-year gap between admissions, or 6 years for a later admission to *ENA*. There are less than 50,000 births over 1891-1990 for the surname, which is common to the distinct curricula—be it the spouse's name or the patronym.

(d) There is at most a 1-year gap between admissions, or 4 years for a later admission to *ENA*.

In addition to the above-mentioned criteria, I ensure that when spouse's names are similar, patronyms are not distinct. Vice versa when patronyms are similar, spouse's names cannot be distinct. I also ensure that genders are not different, only an issue with gender-neutral first names.

In addition, I screened ad-hoc most matches, with particular attention to those with 3 or more identified curricula, those with uncommon sequences of schools, and those with highly occurring surnames (i.e., more than 12,000 births per cohort). I discarded matching errors due to homonyms by comparing biographies and curricula, birth dates, maiden names, or middle names. To this end, I used *LinkedIn*, *Wikipedia* and *Who's who in France* entries, lesbiographies.com, viadeo.journaldunet.com, and lemoniteur.fr websites, biographies published by the newspaper *Les Échos*, as well as institutional biographies available from firms' or institutions' websites.

I also applied bigram and token fuzzy matching of observations to increase the quality of the matching. For each curriculum, I defined a string of characters containing the patronym, spouse's name if applicable, and first names. By visually screening higher scores and comparing complementary observables, I was able to identify potential misspellings, as well as different forms of names in the distinct school registers (e.g., the political representative *Laurent Wauquiez* also appears as *Laurent Wauquiez-Motte*). I consequently matched these curricula manually.

C.3 Conditions identifying "foreign" surnames

Taking advantage of the national birth censuses to approximate the number of potential applicants per surname to the *Grandes Écoles*, the analysis has to be restricted to those surnames providing the clearest observations of number of births. I call them "native" surnames and they are defined as opposed to "foreign" ones. The categorization operates purely at the surname level, regardless of individuals' nationality or migration history, which are not observed.

I identify "foreign" surnames in two ways. First, I use the evolution of births by surname in the national census. Second, I compare the frequency of surnames among students to their frequency in the French birth records.

Using the birth census, I define as "foreign" the 490,565 surnames with only one birth in the censuses over the period 1891-1990. Out of the 786,531 remaining surnames, those for which no birth appears over the first two generations (1891-1940) are classified as foreign. I also consider surnames to be of foreign origin if the birth rate is 10 times higher in the last cohort (1966-1990), as compared to the mean of the first two cohorts (1891-1940), or where the birth rate is 10 times higher than in the previous cohort (e.g., in 1941-1965 compared to 1916-1940).

Finally, I compute by surname S two coefficients of variation of the number of births per cohort. A surname showing a wide-ranging number of births between cohorts is considered to be associated with immigration in a specific generation followed by children born in France over the next generations. I compute $CV^s_{1891-1990}$ for the four generations between 1891 and 1990 and $CV^s_{1891-1966}$ over the first three generations, i.e., 1891-1966, targeting specifically the early immigration of the 20^{th} century.[37] Surnames with an average number of births per cohort μ^s_t above 30 and a coefficient of variation above 0.6 over period t are classified as immigrants. These choices are based on visual inspection at different potential thresholds.

Additionally, I use the *Grandes Écoles* data to classify a surname as foreign if there are more students than there are births in France bearing this surname in any given cohort.

The conditions mean that I consider as "native" surnames for which the immigration phase occurred at latest in the first cohort, between 1891 and 1915. In fact, surnames of

[37] $CV^s_t = \frac{\mu^s_t}{\sigma^s_t}$ where μ^s_t stands for the average number of births of bearers of surname s over the timeframe t—here either 3 or 4 cohorts—and σ^s_t for the standard deviation.

foreign origin stemming from immigration before the period of study are considered native. Therefore, I literally study a stable set of surnames over the period, more than a "native" set of surnames *per se*. Above all, this ensures that the censuses of births in France provide a proper picture of potential applicants to the *Grandes Écoles* for each cohort.

C.4 Multigenerational explanatory variables

Beyond the the probability of having a father who graduated from an elite college presented in section 4.1, this appendix section defines the probability that the paternal grandfather, great-grandfather and great-great-grandfather studied at a *Grande École*.

The probability of the paternal grandfather having studied at a *Grande École* is defined as the probability of being linked to a given father among those bearing the same surname ($\frac{1}{N_{c-1,S}/2}$), multiplied by the probability that a given student in the GE in the grand-paternal cohort is the father of this identified father ($\frac{St^M_{GE,c-2,S}}{N_{c-2,S}/2}$). The probabilities that a paternal grandfather, great-grandfather and great-great-grandfather studied in a given GE are then respectively:

$$X^{Gen-2}_{GE,c,S} = \frac{1}{N_{c-1,S}/2} \times \frac{St^M_{GE,c-2,S}}{N_{c-2,S}/2} = 4 \times \frac{St^M_{GE,c-2,S}}{N_{c-1,S} \times N_{c-2,S}}$$

$$X^{Gen-3}_{GE,c,S} = \frac{8 \times St^M_{GE,c-3,S}}{N_{c-1,S} \times N_{c-2,S} \times N_{c-3,S}}$$

$$X^{Gen-4}_{GE,c,S} = \frac{16 \times St^M_{GE,c-3,S}}{N_{c-1,S} \times N_{c-2,S} \times N_{c-3,S} \times N_{c-4,S}}$$

While $N_{1866-1890}$ is the only missing information, I assume the distribution of births per surname in 1866-1890 to be similar to that in 1891-1915.

C.5 Computation of re-estimated number of students according to continuous explanatory variables

Because the explanatory variables $X_{GE,c,S}$ capture the probability of a dummy variable (for having a GE-graduate father), the number of births and the number of students associated with this characteristic need to be re-estimated. This appendix section formalizes how. These re-estimated figures do not constitute the main results of the paper, but they help interpret the importance of the findings, for instance by approximating the proportion of the *Grandes Écoles*' student body who are descendants of GE-graduates.

The number of births in cohort c of those with a GE-graduate father (identified with probability $X_{GE,c,S}$) bearing a given surname S is $\hat{N}_{c,X_{GE,c,S}} = X_{GE,c,S}.N_{c,S}$ (with $N_{c,S}$ being the total number of births with surname S in cohort c). The number of births per surname of those with the given characteristic ($\hat{N}_{c,X_{GE,c,S}}$) is a proportion of $N_{c,S}$. Indeed, not all bearers of the surname share this characteristic, and $0 < X_{GE,c,S} < 1$.

At the surname level, I can also approximate by cohort the number of GE students with a GE-graduate father in cohort c ($\hat{St}_{GE,c,X_{GE,c,S}}$), using the definition from section 4.2 of the relative admission rate (RAR). This is also a share of $St_{GE,c,S}$, i.e., the number of GE students bearing surname S born in cohort c. The reader can refer to section 4 for additional reminders of notations. The assumption here is that the estimated RAR of those with a GE-graduate father also applies at the surname level. In other words, if children of GE-graduates are x times more likely than the rest of the population to be admitted to GE, the bearers of any surname whose father was a GE-graduate are also x times more likely to be admitted to GE than a bearer of the same surname whose father was not a GE-graduate.

I detail below the computation used to re-estimate the number of students with a GE-graduate father in cohort c, starting from the definition of the RAR:

$$RAR_{GE,c,X_{GE,c,S}} = \frac{AR_{GE,c,X_{GE,c,S}}}{AR_{GE,c,X'_{GE,c,S}}}$$

$$\Leftrightarrow RAR_{GE,c,X_{GE,c,S}} = \frac{\frac{St_{GE,c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}}}{\frac{St_{GE,c,X'_{GE,c,S}}}{N_{c,X'_{GE,c,S}}}}$$

$$\Leftrightarrow \frac{St_{GE,c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}} = RAR_{GE,c,X_{GE,c,S}} . \frac{St_{GE,c,X'_{GE,c,S}}}{N_{c,X'_{GE,c,S}}}$$

$$\Leftrightarrow St_{GE,c,X_{GE,c,S}} = RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X'_{GE,c,S}}} . St_{GE,c,X'_{GE,c,S}}$$

$$\Leftrightarrow St_{GE,c,X_{GE,c,S}} = RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X'_{GE,c,S}}} . (St_{GE,c,S} - St_{GE,c,X_{GE,c,S}})$$

$$\Leftrightarrow St_{GE,c,X_{GE,c,S}} = RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}} . St_{GE,c,S} - RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}} St_{GE,c,S}$$

$$\Leftrightarrow St_{GE,c,X_{GE,c,S}}(1 + RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}}) = RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}} . St_{GE,c,S}$$

$$\Leftrightarrow St_{GE,c,X_{GE,c,S}} = \frac{RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}}}{1 + RAR_{GE,c,X_{GE,c,S}} \cdot \frac{N_{c,X_{GE,c,S}}}{N_{c,X_{GE,c,S}}}} . St_{GE,c,S}$$

$$\Leftrightarrow St_{GE,c,X_{GE,c,S}} = \frac{RAR_{GE,c,X_{GE,c,S}} \cdot \frac{X_{GE,c,S} . N_{c,S}}{(1 - X_{GE,c,S}) . N_{c,S}}}{1 + RAR_{GE,c,X_{GE,c,S}} \cdot \frac{X_{GE,c,S} . N_{c,S}}{(1 - X_{GE,c,S}) . N_{c,S}}} . St_{GE,c,S}$$

$$\Leftrightarrow St_{GE,c,X_{GE,c,S}} = \frac{RAR_{GE,c,X_{GE,c,S}} \cdot \frac{X_{GE,c,S}}{(1 - X_{GE,c,S})}}{1 + RAR_{GE,c,X_{GE,c,S}} \cdot \frac{X_{GE,c,S}}{(1 - X_{GE,c,S})}} . St_{GE,c,S}$$

Therefore, I use the following two formulas to compute the re-estimated number of births and the number of students at the surname level among children of GE graduates:

$$\hat{N}_{c,X_{GE,c,S}} = X_{GE,c,S} . N_{c,S}$$

$$\hat{St}_{GE,c,X_{GE,c,S}} = \frac{RAR_{GE,c,X_{GE,c,S}} \cdot \frac{X_{GE,c,S}}{(1 - X_{GE,c,S})}}{1 + RAR_{GE,c,X_{GE,c,S}} \cdot \frac{X_{GE,c,S}}{(1 - X_{GE,c,S})}} . St_{GE,c,S}$$