

Social Origins and Field of Study*

Francesco Scervini[†] Laura Volponi[‡]

This Version: February 9, 2026

Abstract

Horizontal gender segregation in education negatively impacts both equity and efficiency in the labour market, contributing to disparities in career opportunities and economic outcomes. This paper examines the influence of social origins on gender segregation in Italian tertiary education over recent decades. Utilizing high-quality data on individuals born between 1941 and 1995, we find that parental education significantly shapes offspring's field of study choices. Specifically, higher parental education increases the likelihood of students making non-traditional field choices – girls selecting traditionally male-dominated fields, such as STEM, and boys opting for traditionally female-dominated fields. This effect is primarily driven by girls entering male-dominated fields. The results remain robust across various field classifications and model specifications, highlighting the enduring role of family background in shaping educational and occupational outcomes. These findings align with previous research conducted in other countries.

Keywords: Higher education; Fields of study; Gender inequality; Equality of opportunities; Intergenerational inequality

JEL Codes: I23 (Higher education), I24 (Education and inequality); J16 (Economics of gender)

*We thank Andrea Parma, Stefania Sabatinelli, Paola Savoldi and all participants in seminar at DASTU – Polytechnic University of Milan, and at the ‘Inequality and Education’ workshop – University of Modena and Reggio Emilia for helpful comments. We also thank Nevena Kulic and Irene Prix for precious comments.

[†]University of Pavia, Department of Social and Political Science, Corso Strada Nuova, 65 – 27100 – Pavia (Italy) and Global Labor Organization (GLO) – francesco.scervini@unipv.it

[‡]University of Torino, Via Verdi, 8 – 10124 – Torino (Italy) and Collegio Carlo Alberto, Piazza Vincenzo Arbarello, 8 – 10122 – Torino (Italy) – laura.volponi@carloalberto.org

1 Introduction

Horizontal gender segregation in education remains a central mechanism through which gender inequality in the labour market is reproduced. Even in contexts where women have caught up with – or overtaken – men in higher education enrolment and completion, men and women continue to sort into different fields of study, with persistent consequences for wages, job quality, and career progression (Charles and Bradley, 2002, 2009; Goldin, Katz, and Kuziemko, 2006; Blau and Kahn, 2017; Card and Payne, 2021). This sorting matters for both equity and efficiency. Male-dominated fields are, on average, more strongly associated with high-paying occupations and broader advancement opportunities, while female-dominated fields are more often linked to sectors with lower pay and fewer opportunities for career growth (England and Li, 2006; Black et al., 2008). As a result, horizontal segregation, driven by women’s disproportionate concentration in lower-paying fields, can generate substantial gender gaps even when women’s average educational attainment is high.

Understanding why students choose gender-typical versus gender-atypical fields is therefore crucial for the design of policies aimed at reducing gender inequality. A growing literature points to the role of gender norms, stereotypes, and information frictions, as well as peer and role-model effects in shaping educational investments and field choices (Eccles et al., 2000; Guiso et al., 2008; Carrell, Page, and West, 2009; Carlana, 2019; Carlana and Corno, 2021; Terrier, 2020; Canaan and Mouganie, 2023). At the same time, field choices are also embedded in broader processes of social stratification and

inter-generational transmission. Family background affects educational decisions through resources, aspirations, and beliefs about the returns and risks associated with different trajectories ([McNally, 2020](#); [Delaney and Devereux, 2019](#)). Yet, despite extensive research on gendered educational choices and on social-origin gradients in education separately, we know comparatively less about how gender and social origins jointly shape entry into gender-atypical fields, and whether these mechanisms differ for daughters and sons.

This paper studies how family background and gender interact in shaping horizontal segregation in higher education in Italy. Italy is a useful setting because educational choices are strongly stratified by social origin within a tracked system, and because fields of study remain highly gender-segregated even though women’s participation in university has increased ([Triventi and Trivellato, 2009](#); [Vergolini and Vlach, 2017](#); [Basiglio, Del Boca, and Pronzato, 2024](#); [Bovini, De Philippis, and Rizzica, 2023](#)). These features make family background particularly relevant for gender-atypical choices, which may involve higher uncertainty and higher social costs. More educated parents may relax financial and informational constraints, help children navigate academic pathways, and support departures from traditional gender norms. However, they may also prioritize status preservation and risk minimization, steering children (especially sons) towards conventional, socially rewarded fields. If the costs of norm-breaking are gender-asymmetric, these mechanisms may differ for daughters and sons. We therefore ask to what extent parental education is associated with choosing a gender-atypical field of study, and whether this association differs between daughters and sons.

To address these questions, we combine two complementary data sources.

We rely on the Italian Labour Force Survey (LFS) to construct cohort-specific measures of the gender composition of fields of study, exploiting its large sample size to classify fields as male-dominated, female-dominated, or neutral over time. We then merge this classification with the Survey on Household Income and Wealth (SHIW) from 1995 to 2020, which contains information on individuals' completed field of study and parental education. Our main analysis focuses on 10,527 tertiary-educated individuals born between 1941 and 1995. We estimate linear probability and logit models for the likelihood of making a gender-atypical choice, where parental education is measured as the highest level of education attained by the most educated parent, and we control for cohort, geography, and other individual and household characteristics. We then examine heterogeneity by gender and cohort, and we perform a battery of robustness checks based on alternative definitions of gendered fields, including time-invariant classifications ([Barone, 2011](#); [Barone and Assirelli, 2020](#)).

The results point to a clear and asymmetric pattern. Higher parental education is associated with a higher probability of choosing a gender-atypical field, but this relationship is driven almost entirely by women: parental education has a significantly positive and increasing effect on daughters' atypical choices, while it is small and statistically weak (and in some specifications negative) for sons. In other words, advantaged family backgrounds appear to facilitate women's entry into traditionally male-dominated fields, particularly STEM, more than they facilitate men's entry into traditionally female-dominated fields. These findings are robust across field-classification thresholds, across time-varying and time-invariant definitions, and across al-

ternative model specifications.

This paper makes three contributions to the literature on inequality, education, and gender. First, it provides new evidence for Italy on how social origins shape horizontal gender segregation in tertiary education, using high-quality micro-data that link individuals' field choices to parental education. Second, by explicitly modelling the interaction between parental education and the child's gender, it documents an important asymmetry: family background promotes gender-atypical choices mainly for women, highlighting a channel through which social origins can influence both gender inequality and equality of opportunity. Third, it offers a transparent framework to classify gendered fields over time and shows that core results are robust to alternative measurement strategies, while also linking educational sorting to subsequent income differentials.

The remainder of the paper is organized as follows. Section 2 describes the Italian institutional context and recent trends in gender segregation in higher education. Section 3 outlines the definitions of gendered fields and gender-atypical choices. Section 4 presents the data, sample construction, and descriptive evidence. Section 5 introduces the identification strategy and discusses the main results, robustness checks, and heterogeneity analyses. Section 6 concludes with implications for policy and future research.

2 Education in Italy

The Italian educational system is structured with a common curriculum for primary and lower secondary education, followed by a strongly differentiated

tracking system at the upper secondary level. At the age of 14, students can choose between academically oriented tracks (*licei*), technically oriented tracks (*istituti tecnici*), and vocational training schools (*istituti professionali*). Academic tracks are the most prestigious and are predominantly chosen by students from higher social classes ([Barone and Assirelli, 2020](#)). This pronounced differentiation in educational pathways at the upper secondary level – and the relatively early age at which these choices are made – has been criticized for perpetuating social inequalities, as it significantly shapes students’ opportunities for accessing higher education ([Triventi and Trivellato, 2009](#); [Vergolini and Vlach, 2017](#)). In terms of gender segregation, Italy has seen a rise in female participation in higher education over recent decades. Women are now more likely to graduate on time and achieve higher grades than men. However, they face greater challenges in the labour market, including lower employment rates and wages compared to their male counterparts ([Almalaurea, 2019a,b](#)). These disparities are largely attributable to gender differences in field-of-study choices. [Triventi \(2010\)](#) found that while Italian higher education has experienced some degree of gender desegregation over the past century, this shift has primarily involved women moving from the humanities to relational-fields, with less movement toward technical and scientific disciplines. Over the past decade, gender segregation in higher education has remained relatively stable, with minimal changes in the gender distribution across fields of study. The persistence of these patterns exacerbates gender inequalities in the Italian labour market, where male-dominated fields tend to offer higher wages and better career prospects on average.

3 Defining gendered fields and atypical choices

A correct conceptualization of gender-typical, gender-atypical or gender-neutral fields is essential for the purpose of our analysis. There are two approaches that have been explored by the literature and can be employed in our study. The first approach is rooted in the historical cultural bias associated with each field, where disciplines are traditionally categorized as feminine or masculine based on socially and historically defined gender roles. The primary advantage of such approach is its normative identification, which is based on the intrinsic characteristics of fields of study and the preferences traditionally attributed to young women. However, it has a significant limitation: it fails to account for potential changes and shifts in such preferences over time. The classification of educational fields by gender can evolve over time as individual and societal perceptions change, either through generational shifts in tastes or through policies aimed at levelling gender biases and increasing women's participation in traditionally male-dominated fields of study. To address this limitation, we also adopt a second definition of gender-typical fields, which is data-driven and based on the actual choices of men and women when selecting specific educational paths.

Historically, the categorization of academic fields in the social sciences has followed a gendered division rooted in a humanities/STEM dichotomy, where women were seen as naturally inclined toward the former and men toward the latter. While this framework has occasionally been questioned, it has more recently faced criticism for being too limited to account for the full diversity of educational programs. Many fields defy this binary classification, making it

insufficient for understanding contemporary patterns in educational choices.

To address these limitations, recent approaches have integrated this traditional framework – still relevant despite its shortcomings – with a more nuanced categorization based on a care/technical divide. This distinction focuses less on curricular content and more on the occupational prospects associated with different fields of study. Care-oriented fields, such as nursing, psychology, teaching, and social work, are classified as feminine, while technical fields, such as economics and engineering, are categorized as masculine. Neutral fields include disciplines like natural sciences, pharmacy, medicine, architecture, dentistry, and law.

One of the most frequently cited studies using this integrated framework is [Barone \(2011\)](#), which classifies fields of study along two main dimensions: humanities vs. scientific and care-oriented vs. technical. Building on this approach, we categorize fields into three groups: masculine, feminine, and neutral. i) Feminine fields include disciplines classified as feminine in at least one of the two dimensions. For example, social sciences are considered feminine because they fall under both humanities and care-oriented fields; ii) Masculine fields are those classified as masculine in at least one dimension. For instance, architecture is considered masculine—it is neutral in the humanities-scientific divide but technical in nature; iii) Neutral fields either have a neutral classification in both dimensions (e.g., economics and law) or balance one feminine and one masculine characteristic. For example, biology and mathematics are scientific but often lead to teaching professions, which are care-oriented.

On the other side, the definition of gender fields based on actual choices

requires some data manipulation. First of all, we need a dataset with information on gender, age, field of study and highest level of education attained for a sufficient number of individuals. The European Union Labour Force Survey (LFS), published by Eurostat, has all these features and, specifically, a sample size much larger than other available datasets. Then, starting from all individuals interviewed in Italy between 2009 and 2013 (approximately 3,200,000), we select individuals who are at least 25 years old and have completed tertiary education, and we classify them by gender, cohort,¹ and field of education. Table 1 shows aggregate statistics for all cohorts, while Figure 1 shows how gender composition varies over time for different disciplines.

Starting from the gender composition, for each cohort we define an education field ‘masculine’ if the share of males in that field is larger (by a certain percentage) than the total share of males in that cohort and ‘feminine’ if the share of males is smaller (by the same percentage). As an illustrative example, consider the cohort 1941-1945 in Figure 1. The total share of graduated males is slightly less than 60%. We define ‘masculine’ those fields with a share of males clearly larger than the average (such as Agriculture, Engineering, Architecture, Economics, Law) and ‘feminine’ the fields clearly lower (Humanities). With respect to fields whose share is close to the average, we define them ‘neutral’ if the difference with respect to the average is

¹For confidentiality reasons, LFS does not release the year of birth, but only the cohort within a 5-year interval, making it difficult to compare successive waves. To overcome this problem, we estimate the birth cohort based on the year of obtaining the degree and subtracting the normal duration of the corresponding education cycle. This way, we have a less accurate estimate of age, but a more accurate estimate of when individuals make decisions about their field of education, which is the relevant moment for our study. Finally, we exclude from the analysis all individuals for whom the estimated year of birth is more than 7 years away from the central value of the reference cohort.

smaller than a certain percentage. Since the threshold is discretionary, we use four different thresholds, namely 2.5%, 5%, 7.5% and 10% from the overall mean. According to the thresholds, Social Sciences and Natural Sciences may be classified as feminine/masculine or neutral, while Medicine is always classified as neutral.

Table 2 displays the education fields and their classification over time according to the definitions above. While there are fields whose classification remains stable over time, such as Engineering or Humanities, other fields change from masculine to neutral and eventually feminine, as Medicine. For this reason, we find it useful the comparison between the two kinds of definitions.

While the LFS is used to classify fields by individual gender due to its large sample size, analysing the intergenerational transmission of education and fields of study requires a different dataset, as the LFS lacks information on parental education. Therefore, we utilize the Bank of Italy’s Survey on Household Income and Wealth (SHIW), which is described in detail in the next section. Given that the LFS includes 36 education fields compared to only 9 in the SHIW, we aggregated the 36 LFS categories into the 9 categories used in the SHIW. Table 3 presents the exact definitions and the proposed matches.

Regardless of the methodology used to define masculine, feminine, or neutral fields, an individual’s choice of study is classified based on gender-field alignment. A choice is considered typical when the individual’s gender matches the field’s gender. Conversely, the choice is atypical if there is a mismatch between the individual’s gender and the field’s gender. For example,

a masculine field like Engineering is typical for males but atypical for females, whereas a feminine field such as Humanities is typical for females and atypical for males. Gender-neutral fields are neither typical nor atypical for all individuals. Of course, the classification of the choice changes according to the definition of the fields. Table 4 clarifies the definition of typical and atypical choices based on individuals' gender and field of study classification, Table 5 provides an example, while Figure 2 shows the evolution of atypical choices over cohorts, according to the different definitions of gender fields.

4 Data, sample and descriptive statistics

The main source of data used in the present analysis is the Survey on Household Income and Wealth (SHIW, [Banca d'Italia, 2022](#)), regularly published by the Bank of Italy since 1960s. The sampling unit is the household and about 8,000 households are randomly sampled in every wave. Together with data on income and wealth, SHIW collects many socio-demographic information on every single household member and some information on household head's and partner's parents. Depending on the average size of the households, about 24,000 individuals are interviewed every wave.

Particularly relevant for our goals are the variables about the highest level of education successfully completed, that are collected for all sampled individuals and also for household head's and partner's parents. Household members with tertiary education are also asked about the field of study at tertiary level since 1995. Therefore, we are able to gather all individuals interviewed since 1995. It is important to highlight that, although SHIW is a

panel, we do not exploit its panel component since the choice of educational field is time-invariant and remains unchanged over the life cycle. Likewise, parental formal education, gender and cohort of birth are constant at individual level. Therefore, it is sufficient for our analysis to exploit a cross section composed by all individuals included in the SHIW samples from 1995 to 2020. In order to avoid the replication of observations, we include in the sample only the most recent observation for every individual.²

Overall, we use twelve waves of SHIW³ for a total of 239,196 observations. From these observations, we drop 1,184 observations (.49% of the total) that are presumably twins, and for which it is not possible to select a single observation over the panel dimension. We then select only the most recent observation for every individual, resulting in 132,914 different individuals. Out of this sample, we select the 13,402 individuals with a tertiary education (10.08%) and we drop individuals under the age of 25 (383 individuals), since they could still be in the formal education. Furthermore, we collect information about field of study, parental education, sex, age, area of residence and size of municipality. Missing information about field of study and parental education drop the final sample to 10,527 observations.

Table 6 displays the descriptive statistics of the working sample. With respect to the field, Humanities are the most frequently chosen, by about 21% of the sample, followed by Economics and Statistics (14.4%). Natural science and Pharmacy, Medicine and Dentistry, Engineering, and Law are chosen by

²This choice only affects time-varying characteristics, such as region and size of municipality. However, given the possible endogeneity of these variables, we replicate the models excluding them and results are virtually unchanged. The choice is then substantially irrelevant.

³SHIW was conducted in 1995, 1998, every next year until 2016 and 2020.

about 10% of students respectively. Social sciences (6%), Architecture (4%) and Agriculture and Veterinary (2%) are the less frequently chosen. There is also a residual category that includes about 10% of the sample. As for parental education, about 38% of individuals have the most educated parent with primary or lower secondary education, a condition very frequent for the older cohorts, 35% upper secondary education and 26% tertiary education, a share much higher than the general population, due to the well known inter-generational transmission of inequality. Descriptive statistics on gender, area and size of municipality are in line with official statistics. Finally, we highlight that the distribution of the sample over cohorts is concentrated in the early and central cohorts, while only 8% of the sample is born in 1986-1990 or 1991-1995. As expected, apart from the field of study, there are no significant differences across genders.

A final remark regards the position of individuals within the household. As described above, the sample unit is the household and personal information are collected for all individuals living in the household, while information on parents are collected only for household head and their partner. As a consequence, there are two different types of individuals in our sample. Most of them are household head (47%) and their partners (25%), while the remaining 28% of the sample includes their children. This may have some consequence in terms of self-selection and sample bias, since adult children living in their parental household may have systematically different characteristics with respect to individuals of the same age who are household heads or their partners.⁴ However, in section 5.1 we test and confirm the robustness

⁴This issue is not so problematic in the Italian context since children leave their parental

of our results to sample selection bias.

5 Identification strategy and results

The empirical model used to estimate the effect of parental education on offspring atypical choices is a quite straightforward linear probability model as follows:

$$A_i = \beta_0 + \beta_1 ParEdu_i + \beta_2 X_i + \varepsilon \quad (1)$$

where A_i is a dummy that takes value 1 if the individual i makes a gender-atypical choice, $ParEdu$ is the highest level of education of the most educated parent (in three levels: lower secondary or less, upper secondary, tertiary), and X includes a set of control variables, such gender, cohort of birth (in 5 years brackets from 1941-1945 to 1991-1995), area of residence (North, Center, South and Islands) and the size of the municipality of residence (dummy for municipalities larger than 40.000 inhabitants).⁵ Since the sample includes also individuals belonging to the same household, standard errors are clustered at household level in order to solve potential biases due to the correlation of errors. A non linear binary logit model is in principle more appropriate in this case, but it delivers the same results, both qualitatively and quantitatively, so we prefer the more immediate linear probability model, as it simplifies the interpretation of interacted regressors, which can be more

house on average at the age of 30.

⁵In principle, the place of residence is potentially endogenous, since it may depend on the field of study. We replicate all the analysis by dropping the two variables referred to the area of residence and size of municipality and results are unchanged. Tables are available upon request.

cumbersome in a non-linear framework.⁶

Results from the linear probability model are reported in Table 7 and are mostly in line with expectations. First of all, we observe a positive and significant effect of tertiary education of parents on the probability of atypical choices of children. This effect is significant at standard levels for three over five models, those with a smaller share of neutral fields and therefore a larger sample size, while the significance is slightly lower for the definition from Barone. The same reasoning applies to gender and cohort of birth: female have a larger probability of taking an atypical choices (on average, more girls choose masculine fields than boys feminine fields) and the probability of atypical choices greatly increases with the cohort of birth. This is mainly due to two reasons: on the one side, the share of graduated parents for individuals born in the less recent cohorts is very low, reducing the heterogeneity of the variable; on the other side, the share of females in tertiary education was very low at that time, making the differences across definitions striking (see section 3 above). For these reasons, we replicate the benchmark model only for individuals born after 1955. Results in Table 8 are much more consistent across definitions and with expectations: tertiary education of parents is always positive and significant, while the cohort effect is generally positive only for the last cohort. Unfortunately, we cannot extend the analysis to individuals younger than 1995, as in 2016 they were still in higher education, but we expect an increase in atypical choices, at least for some fields. However, evidence in Triventi (2010) is consistent with our analysis.

⁶Results of the benchmark specification from the logit model are in Tables A1 and A2, while robustness checks and heterogeneity analysis for the logit model are available upon request.

Both the literature and the intuition suggest that the boys and girls may have different reasons to choose atypical fields. In particular, relevant for our analysis, the role of parental education may change according to the gender of the student. To investigate this relations, we slightly modify the model in eq.(1) by including the interaction term between parental education and gender. Results represented in Figure 3 and in Table 9 show that the effect of parental education is significantly positive for girls and insignificant, or even negative, for boys. Moreover, in line with both relevant similar studies and theoretical expectations, the effect for girls increases with parental education. We can therefore conclude that the positive effect of parental education is only effective for girls.

As a significant portion of the literature on this topic suggests, the gender of the most educated parent may have an effect on children education due to gender role models and imitation. Research on intergenerational social mobility has shown that the education level of both mothers and fathers significantly influences their children’s outcomes. Girls (boys) may be more likely to follow mothers (fathers) role and suggestions, determining an heterogeneous effect. In order to investigate this mechanism, we replicate the previous model with fathers and mothers separately, instead of the most educated parent. Results in Table 10 show that, after controlling for maternal education, paternal education has virtually the same effect as parental education.⁷ Maternal education seems to have an effect only on girls, and only

⁷Parental and paternal education are very correlated: only in about 9% of the case the most educated parent is the mother, while the opposite is true in about 25% of the sample. 66% of individuals in the sample have both parents with the same educational level (see Table 11)

when mothers have upper secondary education (Table 12). However, it must be noted that, since the sample is composed by relatively aged individuals (born in 1941-1995), the number of graduated girls with graduated mothers is as low as 7.7% of the sample, reducing the number at 1.1% for girls who made atypical choices, and this is likely the reason why we do not find any significant effect of mothers with tertiary education on girls. Including both parents in the same regression with interactions (Table 13) confirms the main result: educated fathers have a positive and significant effect on the probability that girls take atypical fields and a negative and significant effect on the probability that boys choose atypical fields. Indeed, we find a significant positive effect of educated mothers on boys, that may explain this result. Looking at the predicted probabilities of undertaking an atypical choice (Figure 4), we observe that high educated fathers tend to exacerbate the difference between boys and girls, while high educated mothers reduce or even close the gender gap, depending on the definition.

5.1 Robustness checks and heterogeneity

The analysis presented in the previous section compares individuals' choice between typical and atypical fields, therefore excluding from the regressions all individuals who choose a neutral field. An alternative specification could include all individuals, analysing the choice between atypical and non-atypical fields, including then in the same category both typical and neutral fields. Results in Table 14 are qualitatively consistent with the previous analysis, but quantitatively smaller (in absolute value), as expected. Indeed, the

relevance of parental education is “diluted” over a much larger sample, most of which is composed by individuals who took neutral fields. On the other side, their inclusion increases the sample size and makes it consistent across definition of typical and atypical fields.

One of the characteristics of Italian educational system is the early choice of the track of study, as discussed in section 2. Within each track, young students must also choose a field. For instance, within the academically oriented tracks students can choose among classical studies (characterized by the study of ancient Greek and Latin and very few hours of mathematics and natural sciences), scientific studies (more focused on mathematics, physics, chemistry and other natural sciences), human sciences (with a focus on psychology, sociology, education and teaching). Unfortunately, information on upper secondary field is available for graduated individuals only since 2008, therefore reducing the sample by about 40%. We include in the regression a categorical variable identifying students who attended non-academic tracks; academic tracks in field of education and arts; academic tracks in other fields. Including this variable makes results more significant and quantitatively larger (Table 15), reassuring on the validity and robustness of our analyses. Moreover, the model delivers an interesting hint: the probability of undertaking an atypical field is significantly larger for all fields with respect to teaching and arts (omitted category in the regressions), that are typically female-dominated fields.

As detailed in section 4 in the previous analyses we consider jointly two kinds of individuals within the sampled households: household head and their partners, and their children living in the household. However, the

two categories may be characterized by different features. In particular, the generation of children observed in the survey is composed only by individuals who are graduated and are still living with their parents. Although this condition is very frequent in Italy, the sample of children may suffer from a self-selection bias, therefore we replicate all the analysis for the sample of parents only. Also in this case, results are virtually unaffected (Tables 16 and 17).⁸

Lastly, the effect of parental education on atypical field choices may be heterogeneous over cohorts. Figure 5 shows that the effect in our sample is increasing with cohorts according to the definition by Barone, while it is constant in the other definitions. This result is due to the fact that Barone definition is time-invariant, while other indicators change over time according to the gender composition within each field. Therefore, the increasing effect observed for the definition in Barone is absorbed by the atypical choice time-varying indicator in the other definitions. The increase of the standard deviation for the last cohorts is likely due not to an actual loss of significance of the variable of interest, but to the smaller sample size. Indeed, the cohort 1986-1990 includes only about 5% of the sample while the last cohort, 1991-1995, less than 3%. The magnitude of the estimates is however consistent with the previous cohorts.

⁸Predictably, the only difference is the value and significance of cohort dummies. Indeed, the generation of children is on average younger than the generation of parents and this leads to different time-trends for the two groups.

6 Concluding remarks

This paper investigates the effect of parental education on the probability of offspring undertaking an atypical educational choice, defined in several ways, in Italy for cohorts born between 1941 and 1995. Although the results show some heterogeneity, the overall analysis indicates that higher parental education is correlated with a greater likelihood of atypical choices by children. These findings align with theoretical predictions from economic and sociological literature, highlighting how educational field choices contribute to both gender segregation and intergenerational inequality. Specifically, not only does parental education significantly influence the probability of attaining higher education, but it also affects the likelihood of selecting a segregated field of study — one associated with lower wages, more discontinuous career paths, and reduced chances of reaching top positions or breaking the glass ceiling.

These findings raise both ethical and efficiency concerns. A significant portion of these differences among individuals stems from factors beyond personal control and effort, raising equity issues. Furthermore, field segregation generates inefficiencies, as long as talent is evenly distributed across genders while education and occupations remain highly unbalanced.

To address these inequalities and inefficiencies, policies aimed at reducing gender segregation in educational choices are necessary. In recent decades, several Western countries have implemented initiatives targeting high school students and families, such as scholarship programs for under-represented groups, mentorship initiatives, and awareness campaigns promoting STEM

fields among women. However, as reviewed in the introduction, empirical studies indicate that the overall impact of these measures has been limited: while some policies may have short-term effects, the inter-generational nature of role models and cultural shifts requires decades to fully materialize.

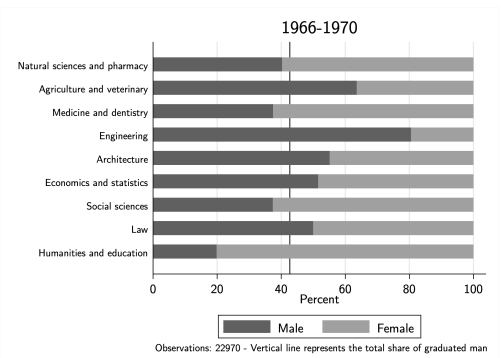
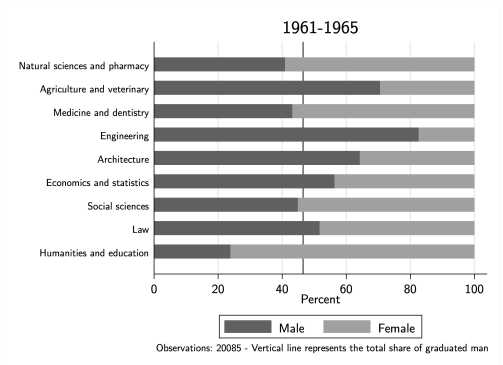
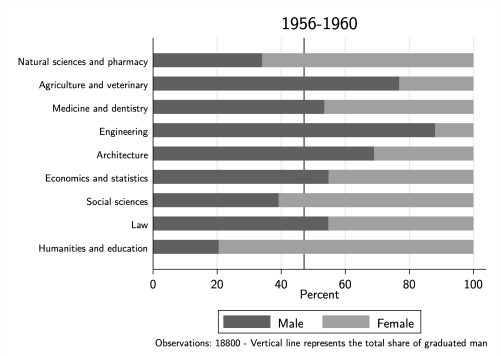
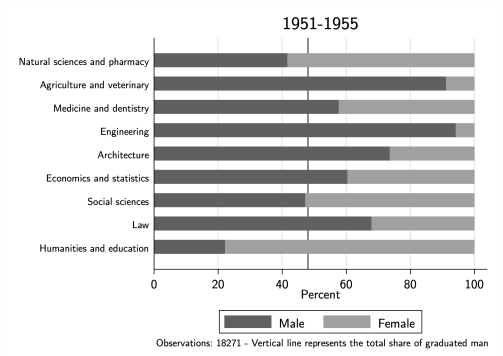
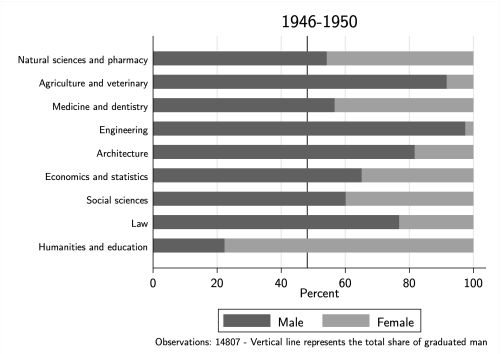
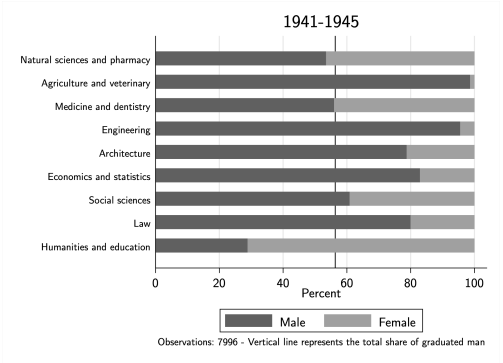
It is important to remark that the findings presented in this paper align with the results of similar studies conducted in significantly different settings from Italy. This consistency suggests that underlying mechanisms or trends may extend beyond specific national characteristics, indicating potential broader relevance, in particular in Western countries. Nonetheless, further research is needed to determine the extent to which these findings hold across diverse socio-economic, cultural, and institutional frameworks, in particular where the gender roles and the value of education differ substantially.

Future research should focus on assessing the long-term impact of these policies to design more effective interventions. In particular, more attention should be given to understanding which types of programs yield the highest returns in different socioeconomic contexts and how policy design can account for the persistence of cultural norms.

7 Figures and Tables

Table 1: Frequency distribution of Italian graduates by field of study and gender based on data provided by LFS (respondents from waves 2009-2013 aged more than 25 and born after 1940), by estimated birth cohort (1941-1988). Source: European Union Labour Force Survey.

	Sex	
	Male	Female
Mathematics, physics, chemistry, biology, pharmacy	10,804 42.12%	14,845 57.88%
Agricultural and veterinary science	21,089 74.60%	7,179 25.40%
Medicine, dentistry, health care professions	19,917 41.41%	29,366 59.59%
Engineering	168,063 84.71%	30,324 15.29%
Architecture and urban planning	49,594 80.71%	11,856 19.29%
Economics and statistics	103,657 39.11%	161,362 60.89%
Political science, sociology	15,259 37.62%	25,305 62.38%
Law	14,181 51.66%	13,269 48.34%
Arts, philosophy, languages, education, psychology	21,135 23.41%	69,155 76.59%
Total	423,699 53.88%	362,661 46.12%



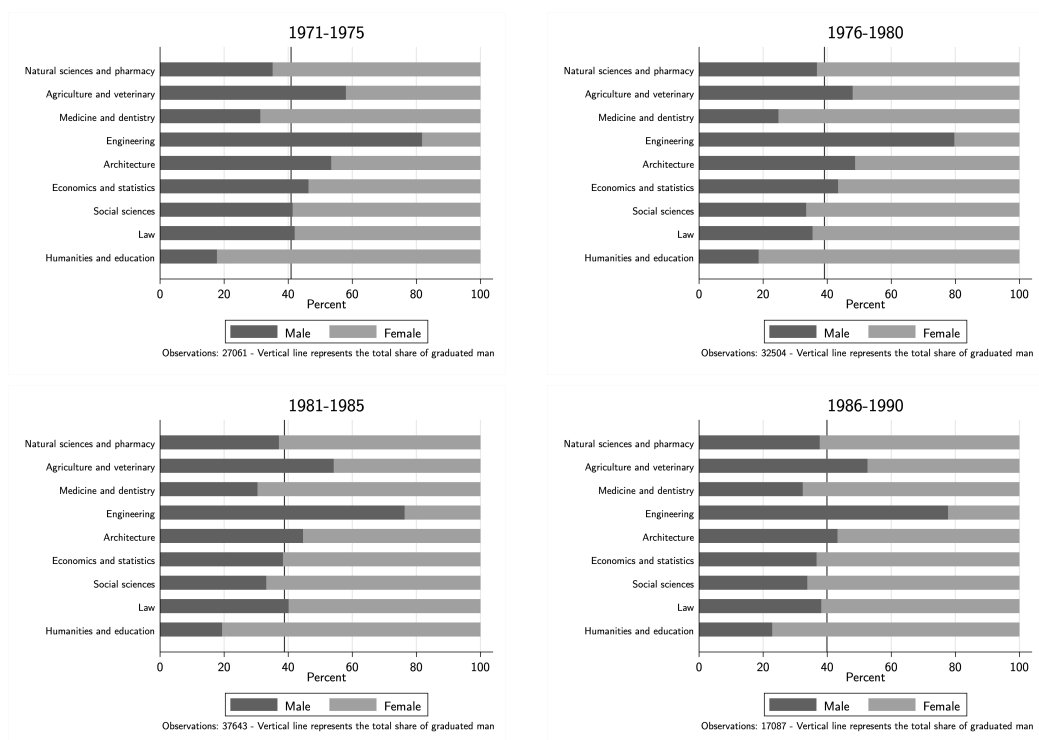


Figure 1: Bar graphs: frequency distribution (percent) of Italian graduates by field of study and gender based on data provided by LFS (respondents from waves 2009-2013 aged more than 25 and born after 1940), by estimated birth cohort (1941-1988). Source: Eurostat Labour Force Survey.

Table 2: Classification of fields by cohort and definition.

	Cohort										
	41-45	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95
	Mathematics, physics, chemistry, biology, pharmacy										
Barone	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
2.5%	Fem.	Masc.	Fem.	Fem.	Fem.	Neut.	Fem.	Neut.	Neut.	Neut.	Neut.
5%	Neut.	Masc.	Fem.	Fem.	Fem.	Neut.	Fem.	Neut.	Neut.	Neut.	Neut.
7.5%	Neut.	Neut.	Neut.	Fem.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
10%	Neut.	Neut.	Neut.	Fem.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
	Agricultural and veterinary science										
Barone	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
2.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
7.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
10%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Masc.	Masc.	Masc.
	Medicine, dentistry, health care professions										
Barone	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
2.5%	Neut.	Masc.	Masc.	Masc.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
5%	Masc.	Masc.	Masc.	Masc.	Neut.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
7.5%	Neut.	Masc.	Masc.	Neut.	Neut.	Neut.	Fem.	Fem.	Fem.	Fem.	Fem.
10%	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Fem.	Neut.	Neut.	Neut.
	Engineering										
Barone	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
2.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
7.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
10%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
	Architecture and urban planning										
Barone	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
2.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Neut.
7.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.
10%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.	Neut.
	Economics and statistics										
Barone	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.
2.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Fem.	Fem.
5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.	Neut.
7.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.	Neut.	Neut.
10%	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
	Political science, sociology										
Barone	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
2.5%	Masc.	Masc.	Neut.	Fem.	Neut.	Fem.	Neut.	Fem.	Fem.	Fem.	Fem.
5%	Neut.	Masc.	Neut.	Fem.	Neut.	Fem.	Neut.	Fem.	Fem.	Fem.	Fem.
7.5%	Neut.	Masc.	Neut.	Fem.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
10%	Neut.	Masc.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
	Law										
Barone	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
2.5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Fem.	Neut.	Neut.	Neut.
5%	Masc.	Masc.	Masc.	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.	Neut.	Neut.
7.5%	Masc.	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
10%	Masc.	Masc.	Masc.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
	Arts, philosophy, languages, education, psychology										
Barone	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
2.5%	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
5%	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
7.5%	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
10%	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.

Table 3: Labour Force Survey (LFS) and Survey on Household Income and Wealth (SHIW) complete fields of study definitions.

LFS fields of study complete definitions	SHIW fields of study complete definitions
Science, mathematics and computing	Mathematics, physics, chemistry, biology, pharmacy
Life science	
Physical science	
Physics	
Chemistry	
Meteorology and weather forecasting science	
Mathematics and statistics	
Mathematics	
Agriculture and veterinary	Agricultural and veterinary sciences
Agriculture, forestry and fishery	
Health and welfare	Medicine, dentistry, health care professions
Health and medicine	
Nutrition and dietetics	
Computer science	Engineering
Computer use	
Engineering, manufacturing and construction	
Architecture and building	Architecture and urban planning
Architecture	
Economics	Economics and statistics
Business and administration	
Statistics and actuarial science	
Social sciences, business and law	Political science, sociology
Social services	
Law	Law
Teacher training and education science	Arts, philosophy, languages, education, psychology
Teacher training	
Humanities, languages and arts	
Arts	
Humanities	
Foreign languages	
General programmes	Other
Services	
Personal services	
Sports	
Transport services	
Unknown	
Not applicable	

Table 4: Definition of individual choice according to gender and field of study.

Field of study	Gender of individual	
	Male	Female
Masculine	Typical	Atypical
Feminine	Atypical	Typical
Neutral	Not in the main sample	

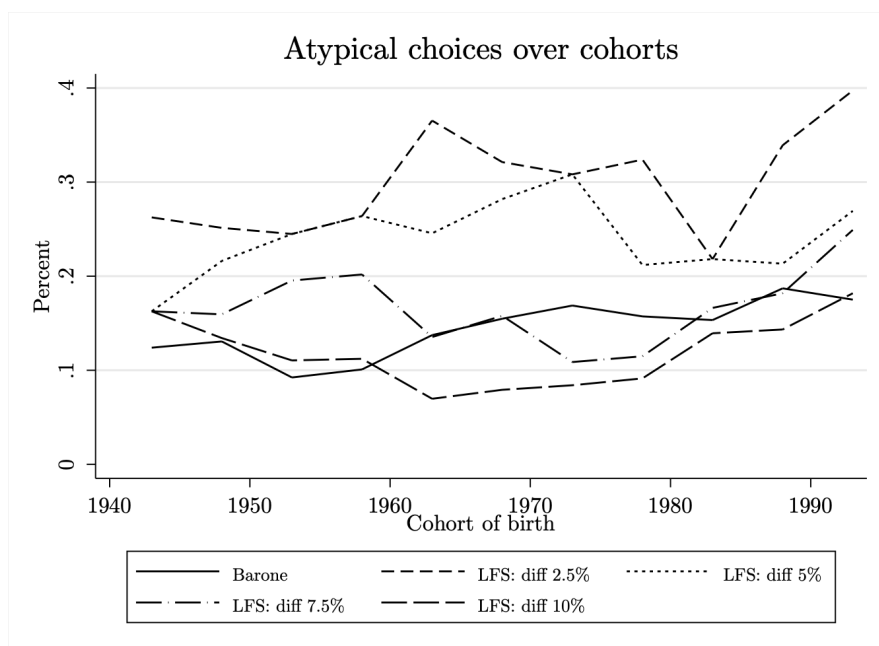


Figure 2: Line graph: share of atypical choices by definition of fields' gender and birth cohort. Data sources: individuals from the Survey on Household Income and Wealth (SHIW), matched with field definitions from the Labour Force Survey (LFS).

Table 5: Definition of individual choice according to gender and field of study: example.

		Cohort of Birth										
		41-45	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95
Medicine, dentistry, health care professions: gender labels												
Barone	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.
2.5%	Neut.	Masc.	Masc.	Masc.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
5%	Neut.	Masc.	Masc.	Masc.	Neut.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
7.5%	Neut.	Masc.	Masc.	Neut.	Neut.	Neut.	Fem.	Fem.	Fem.	Fem.	Fem.	Fem.
10%	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Neut.	Fem.	Neut.	Neut.	Neut.	Neut.
Medicine, dentistry, health care professions: Male Student's educational choice												
Barone	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral
2.5%	Neutral	Typical	Typical	Typical	Atypical	Atypical	Atypical	Atypical	Atypical	Atypical	Atypical	Atypical
5%	Neutral	Typical	Typical	Typical	Neutral	Atypical	Atypical	Atypical	Atypical	Atypical	Atypical	Atypical
7.5%	Neutral	Typical	Typical	Neutral	Neutral	Neutral	Atypical	Atypical	Atypical	Atypical	Atypical	Atypical
10%	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Atypical	Neutral	Neutral	Neutral	Neutral
Medicine, dentistry, health care professions: Female Student's educational choice												
Barone	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral
2.5%	Neutral	Atypical	Atypical	Atypical	Typical	Typical	Typical	Typical	Typical	Typical	Typical	Typical
5%	Neutral	Atypical	Atypical	Atypical	Neutral	Typical	Typical	Typical	Typical	Typical	Typical	Typical
7.5%	Neutral	Atypical	Atypical	Neutral	Neutral	Neutral	Typical	Typical	Typical	Typical	Typical	Typical
10%	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Typical	Neutral	Neutral	Neutral	Neutral

Table 6: Descriptive statistics of the working sample by gender

	Male		Gender Female		Total	
	No.	%	No.	%	No.	%
Field of study						
Natural sciences & Pharmacy	488	9.6	660	12.1	1,148	10.9
Agriculture & Vet	155	3.1	65	1.2	220	2.1
Medicine & Dentistry	586	11.5	472	8.7	1,058	10.1
Engineering	870	17.1	137	2.5	1,007	9.6
Architecture	237	4.7	196	3.6	433	4.1
Economics & Stats	924	18.2	590	10.8	1,514	14.4
Social sciences	284	5.6	351	6.4	635	6.0
Law	653	12.9	588	10.8	1,241	11.8
Humanities	432	8.5	1,789	32.8	2,221	21.1
Other	452	8.9	598	11.0	1,050	10.0
Total	5,081	100.0	5,446	100.0	10,527	100.0
Highest parental education						
Parent lower sec. or less	1,957	38.5	2,075	38.1	4,032	38.3
Parent upper sec.	1,726	34.0	2,012	36.9	3,738	35.5
Parent tertiary	1,398	27.5	1,359	25.0	2,757	26.2
Total	5,081	100.0	5,446	100.0	10,527	100.0
Geographical area						
North	2,404	47.3	2,462	45.2	4,866	46.2
Center	1,091	21.5	1,233	22.6	2,324	22.1
South and islands	1,586	31.2	1,751	32.2	3,337	31.7
Total	5,081	100.0	5,446	100.0	10,527	100.0
Degree of urbanization						
Up to 40k	1,511	29.7	1,782	32.7	3,293	31.3
More than 40k	3,570	70.3	3,664	67.3	7,234	68.7
Total	5,081	100.0	5,446	100.0	10,527	100.0
Cohort of birth						
1941-1945	353	6.9	201	3.7	554	5.3
1946-1950	559	11.0	428	7.9	987	9.4
1951-1955	576	11.3	501	9.2	1,077	10.2
1956-1960	559	11.0	542	10.0	1,101	10.5
1961-1965	644	12.7	687	12.6	1,331	12.6
1966-1970	679	13.4	772	14.2	1,451	13.8
1971-1975	578	11.4	774	14.2	1,352	12.8
1976-1980	408	8.0	558	10.2	966	9.2
1981-1985	366	7.2	506	9.3	872	8.3
1986-1990	224	4.4	323	5.9	547	5.2
1991-1995	135	2.7	154	2.8	289	2.7
Total	5,081	100.0	5,446	100.0	10,527	100.0
Household composition						
Head of family	3,131	61.6	1,831	33.6	4,962	47.1
Partner	546	10.7	2,079	38.2	2,625	24.9
Child	1,404	27.6	1,536	28.2	2,940	27.9
Total	5,081	100.0	5,446	100.0	10,527	100.0

Table 7: OLS benchmark model

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Parent upper sec.	0.003 (0.014)	0.008 (0.011)	0.016 (0.012)	0.012 (0.013)	0.006 (0.014)
Parent tertiary	0.022 (0.016)	0.031** (0.013)	0.035*** (0.013)	0.033** (0.014)	0.020 (0.016)
Female	0.000 (0.012)	0.092*** (0.010)	0.143*** (0.011)	0.110*** (0.012)	-0.002 (0.013)
1946-1950	0.027 (0.028)	-0.030 (0.024)	0.015 (0.024)	-0.013 (0.025)	-0.009 (0.025)
1951-1955	-0.009 (0.028)	-0.038 (0.024)	0.026 (0.024)	0.010 (0.024)	-0.022 (0.025)
1956-1960	0.002 (0.028)	-0.006 (0.025)	0.057** (0.025)	0.034 (0.025)	0.008 (0.027)
1961-1965	0.066** (0.028)	0.092*** (0.024)	0.108*** (0.025)	0.047* (0.026)	-0.005 (0.028)
1966-1970	0.079*** (0.027)	0.056** (0.024)	0.113*** (0.024)	0.068*** (0.025)	0.023 (0.027)
1971-1975	0.114*** (0.028)	0.070*** (0.025)	0.128*** (0.025)	0.029 (0.028)	0.031 (0.028)
1976-1980	0.084*** (0.030)	0.056** (0.026)	0.082*** (0.028)	0.021 (0.030)	0.028 (0.030)
1981-1985	0.066** (0.030)	0.017 (0.028)	0.073*** (0.028)	0.047 (0.029)	0.069** (0.030)
1986-1990	0.142*** (0.035)	0.106*** (0.032)	0.109*** (0.035)	0.112*** (0.036)	0.132*** (0.037)
1991-1995	0.146*** (0.045)	0.174*** (0.039)	0.185*** (0.044)	0.198*** (0.045)	0.199*** (0.049)
Center	0.003 (0.015)	0.010 (0.012)	0.007 (0.012)	0.010 (0.014)	0.005 (0.016)
South and islands	-0.021 (0.014)	-0.002 (0.011)	0.002 (0.011)	-0.007 (0.013)	-0.018 (0.014)
More than 40k	-0.009 (0.013)	0.005 (0.011)	0.008 (0.011)	-0.004 (0.012)	-0.006 (0.014)
Constant	0.213*** (0.027)	0.236*** (0.023)	0.143*** (0.024)	0.167*** (0.024)	0.219*** (0.025)
R ²	0.012	0.024	0.037	0.027	0.013
Observations	5597	9284	8271	6247	4782

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Table 8: OLS model with restricted sample: born after 1955

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Parent upper sec.	0.022 (0.016)	0.021 (0.014)	0.026* (0.014)	0.033** (0.016)	0.033* (0.019)
Parent tertiary	0.054*** (0.018)	0.049*** (0.015)	0.049*** (0.016)	0.061*** (0.018)	0.058*** (0.021)
Female	0.090*** (0.015)	0.094*** (0.012)	0.146*** (0.013)	0.090*** (0.015)	-0.008 (0.017)
1961-1965	0.064*** (0.025)	0.097*** (0.019)	0.050** (0.020)	0.011 (0.022)	-0.015 (0.026)
1966-1970	0.076*** (0.024)	0.061*** (0.019)	0.054*** (0.019)	0.032 (0.022)	0.011 (0.026)
1971-1975	0.109*** (0.025)	0.075*** (0.020)	0.070*** (0.020)	-0.006 (0.024)	0.020 (0.027)
1976-1980	0.073*** (0.027)	0.060*** (0.022)	0.023 (0.024)	-0.016 (0.027)	0.014 (0.028)
1981-1985	0.054** (0.027)	0.020 (0.024)	0.013 (0.024)	0.009 (0.026)	0.054* (0.029)
1986-1990	0.128*** (0.032)	0.108*** (0.028)	0.049 (0.031)	0.073** (0.033)	0.116*** (0.036)
1991-1995	0.135*** (0.043)	0.174*** (0.036)	0.123*** (0.041)	0.153*** (0.043)	0.176*** (0.049)
Center	0.005 (0.017)	0.001 (0.015)	0.004 (0.015)	0.005 (0.018)	0.004 (0.021)
South and islands	-0.025 (0.016)	-0.004 (0.013)	0.001 (0.014)	-0.006 (0.016)	-0.017 (0.018)
More than 40k	-0.003 (0.015)	0.014 (0.013)	0.015 (0.013)	-0.001 (0.015)	0.001 (0.017)
Constant	0.152*** (0.024)	0.217*** (0.019)	0.187*** (0.020)	0.197*** (0.022)	0.210*** (0.027)
R ²	0.019	0.018	0.029	0.019	0.016
Observations	4156	6757	5892	4142	3050

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Table 9: OLS model with interactions

	(1) Barone b/se	(2) 2.5% b/se	(3) 5% b/se	(4) 7.5% b/se	(5) 10% b/se
Parent upper sec.	-0.050** (0.021)	-0.023 (0.016)	-0.016 (0.016)	-0.027 (0.017)	-0.056*** (0.021)
Parent tertiary	-0.047** (0.023)	0.016 (0.017)	0.005 (0.017)	0.001 (0.019)	-0.036 (0.023)
Female	-0.068*** (0.019)	0.063*** (0.016)	0.106*** (0.016)	0.067*** (0.018)	-0.069*** (0.019)
Parent upper sec. X Female	0.098*** (0.027)	0.058*** (0.023)	0.061*** (0.023)	0.076*** (0.026)	0.116*** (0.028)
Parent tertiary X Female	0.133*** (0.032)	0.031 (0.025)	0.060** (0.026)	0.063** (0.029)	0.108*** (0.032)
Constant	0.246*** (0.028)	0.250*** (0.024)	0.160*** (0.024)	0.186*** (0.025)	0.251*** (0.026)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.016	0.025	0.038	0.028	0.017
Observations	5597	9284	8271	6247	4782
Effect of up.sec. ed. on males	-0.050** (0.021)	-0.023 (0.016)	-0.016 (0.016)	-0.027 (0.017)	-0.056*** (0.021)
Effect of up.sec. ed. on females	0.047** (0.018)	0.036** (0.016)	0.045*** (0.017)	0.050** (0.019)	0.059*** (0.019)
Effect of tert. ed. on males	-0.047** (0.023)	0.016 (0.017)	0.005 (0.017)	0.001 (0.019)	-0.036 (0.023)
Effect of tert. ed. on females	0.086*** (0.022)	0.046** (0.019)	0.064*** (0.020)	0.065*** (0.022)	0.072*** (0.023)

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

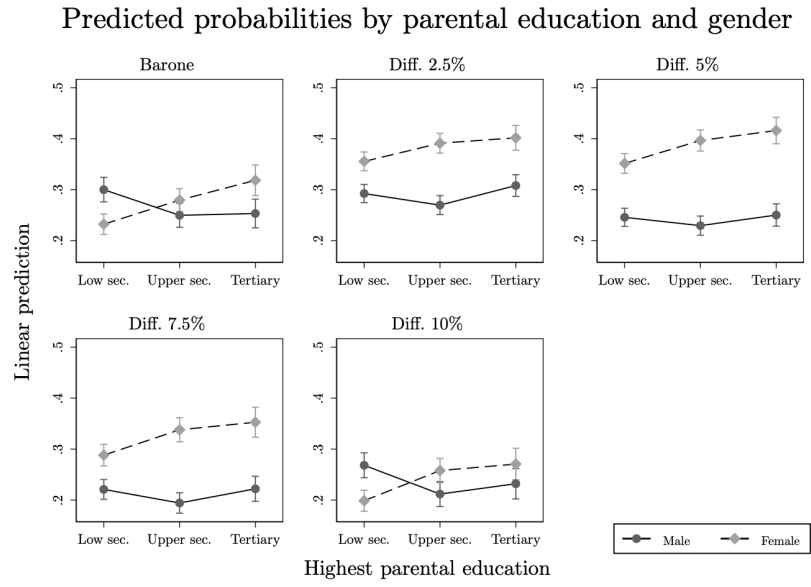


Figure 3: Predicted probabilities of atypical choice, by gender and parental highest education. Confidence intervals at 90% level.

Table 10: OLS model with interactions (father only)

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Father upper sec.	-0.049** (0.022)	-0.032* (0.017)	-0.024 (0.017)	-0.030 (0.019)	-0.053** (0.023)
Father tertiary	-0.080*** (0.027)	-0.009 (0.020)	-0.031 (0.021)	-0.034 (0.023)	-0.060** (0.027)
Female	-0.059*** (0.018)	0.065*** (0.015)	0.106*** (0.016)	0.070*** (0.017)	-0.065*** (0.019)
Father upper sec. X Female	0.080*** (0.028)	0.065*** (0.023)	0.063*** (0.024)	0.069*** (0.026)	0.108*** (0.029)
Father tertiary X Female	0.176*** (0.034)	0.060** (0.027)	0.105*** (0.028)	0.105*** (0.031)	0.141*** (0.034)
Mother upper sec.	0.004 (0.016)	0.004 (0.014)	0.010 (0.014)	0.013 (0.015)	0.005 (0.017)
Mother tertiary	0.022 (0.024)	0.019 (0.020)	0.025 (0.020)	0.022 (0.023)	0.007 (0.025)
Constant	0.248*** (0.028)	0.254*** (0.024)	0.163*** (0.024)	0.189*** (0.025)	0.252*** (0.026)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.018	0.027	0.042	0.031	0.018
Observations	5326	8826	7864	5951	4553
Effect of up.sec. ed. on males	-0.049** (0.022)	-0.032* (0.017)	-0.024 (0.017)	-0.030 (0.019)	-0.053** (0.023)
Effect of up.sec. ed. on females	0.030 (0.020)	0.032* (0.018)	0.039** (0.019)	0.039* (0.021)	0.055*** (0.021)
Effect of tert. ed. on males	-0.080*** (0.027)	-0.009 (0.020)	-0.031 (0.021)	-0.034 (0.023)	-0.060** (0.027)
Effect of tert. ed. on females	0.096*** (0.027)	0.051** (0.023)	0.074*** (0.024)	0.071*** (0.027)	0.080*** (0.028)

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Table 11: Cross tabulation of parental education

Fathers' education	Mother's education			Total
	Lower sec. or less	Upper sec.	Tertiary	
Lower sec. or less	3,815	480	72	4,367
	38.09	4.79	0.72	43.60
Upper secondary	1,187	1,870	343	3,400
	11.85	18.67	3.42	33.95
Tertiary	274	982	992	2,248
	2.74	9.81	9.91	22.45
Total	5,276	3,332	1,407	10,015
	52.68	33.27	14.05	100

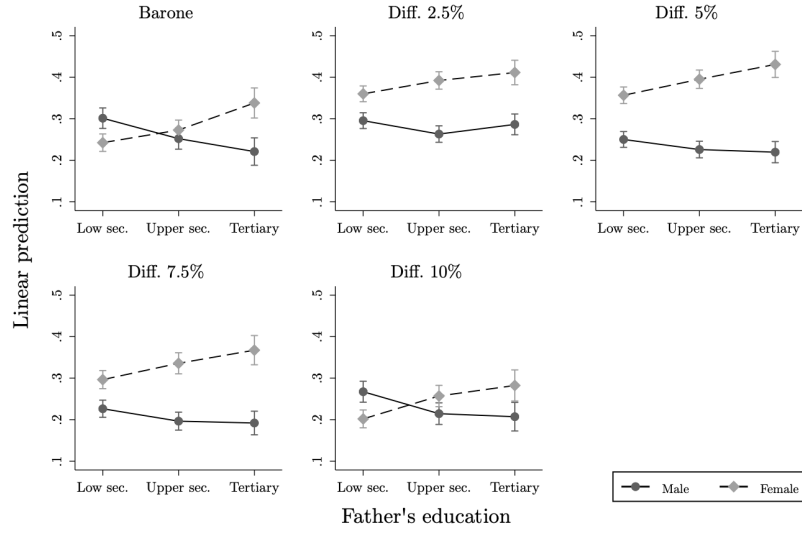
Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Table 12: OLS model with interactions (mother only)

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Mother upper sec.	-0.062*** (0.022)	-0.014 (0.017)	-0.014 (0.017)	-0.023 (0.018)	-0.046** (0.022)
Mother tertiary	-0.037 (0.032)	0.045* (0.025)	0.036 (0.026)	0.036 (0.029)	-0.021 (0.034)
Female	-0.050*** (0.017)	0.096*** (0.014)	0.139*** (0.014)	0.097*** (0.015)	-0.036** (0.017)
Mother upper sec. X Female	0.122*** (0.027)	0.033 (0.022)	0.044* (0.023)	0.069*** (0.025)	0.095*** (0.028)
Mother tertiary X Female	0.109*** (0.040)	-0.050 (0.032)	-0.021 (0.033)	-0.029 (0.038)	0.054 (0.043)
Father upper sec.	-0.005 (0.016)	0.002 (0.013)	0.009 (0.014)	0.007 (0.015)	0.006 (0.016)
Father tertiary	0.012 (0.021)	0.022 (0.017)	0.022 (0.018)	0.018 (0.019)	0.013 (0.021)
Constant	0.243*** (0.027)	0.240*** (0.024)	0.148*** (0.024)	0.176*** (0.025)	0.237*** (0.025)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.017	0.027	0.040	0.030	0.016
Observations	5326	8826	7864	5951	4553
Effect of up.sec. ed. on males	-0.062*** (0.022)	-0.014 (0.017)	-0.014 (0.017)	-0.023 (0.018)	-0.046** (0.022)
Effect of up.sec. ed. on females	0.062*** (0.019)	0.028* (0.016)	0.041** (0.017)	0.054*** (0.019)	0.056*** (0.020)
Effect of tert. ed. on males	-0.037 (0.032)	0.045* (0.025)	0.036 (0.026)	0.036 (0.029)	-0.021 (0.034)
Effect of tert. ed. on females	0.072** (0.031)	-0.005 (0.025)	0.015 (0.027)	0.008 (0.030)	0.032 (0.032)

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Predicted probabilities by paternal education and gender



Predicted probabilities by maternal education and gender

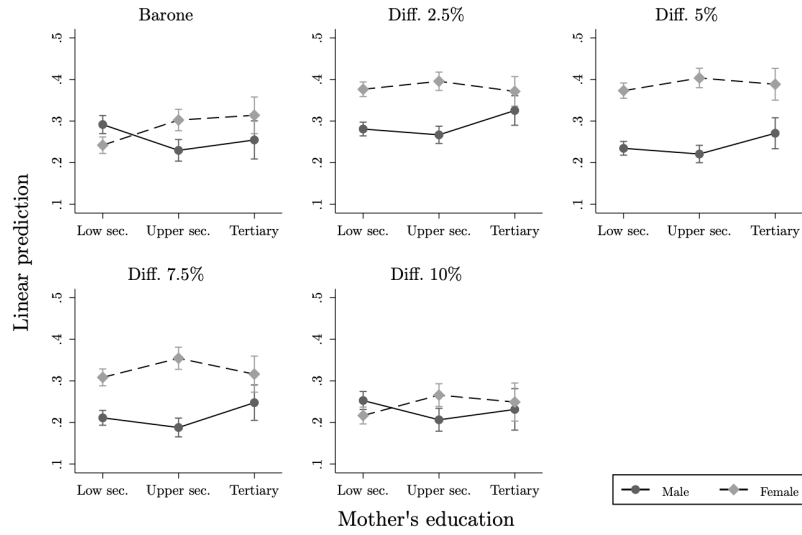


Figure 4: Predicted probabilities of atypical choice, by gender and paternal (upper panel) and maternal (lower panel) highest education. Confidence intervals at 90% level.

Table 13: OLS model with interactions (both parents)

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Father upper sec.	-0.031 (0.024)	-0.045** (0.018)	-0.036** (0.018)	-0.034* (0.020)	-0.048** (0.024)
Father tertiary	-0.062** (0.030)	-0.041* (0.023)	-0.061*** (0.023)	-0.058** (0.025)	-0.061** (0.030)
Female	-0.067*** (0.019)	0.069*** (0.015)	0.111*** (0.016)	0.072*** (0.017)	-0.067*** (0.019)
Father upper sec. X Female	0.046 (0.032)	0.087*** (0.026)	0.085*** (0.027)	0.077** (0.030)	0.099*** (0.033)
Father tertiary X Female	0.143*** (0.042)	0.124*** (0.034)	0.165*** (0.035)	0.152*** (0.039)	0.143*** (0.042)
Mother upper sec.	-0.035 (0.024)	0.019 (0.018)	0.024 (0.018)	0.011 (0.020)	-0.010 (0.024)
Mother tertiary	0.011 (0.035)	0.087*** (0.027)	0.092*** (0.028)	0.087*** (0.031)	0.025 (0.037)
Mother upper sec. X Female	0.073** (0.033)	-0.027 (0.026)	-0.026 (0.027)	0.005 (0.030)	0.028 (0.034)
Mother tertiary X Female	0.018 (0.048)	-0.132*** (0.038)	-0.129*** (0.040)	-0.128*** (0.045)	-0.036 (0.049)
Constant	0.252*** (0.028)	0.253*** (0.024)	0.162*** (0.024)	0.189*** (0.025)	0.253*** (0.026)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.019	0.028	0.043	0.033	0.019
Observations	5326	8826	7864	5951	4553
Effect of fath. up.sec. ed. on males	-0.031 (0.024)	-0.045** (0.018)	-0.036** (0.018)	-0.034* (0.020)	-0.048** (0.024)
Effect of moth. up.sec. ed. on males	-0.035 (0.024)	0.019 (0.018)	0.024 (0.018)	0.011 (0.020)	-0.010 (0.024)
Effect of fath. up.sec. ed. on females	0.015 (0.021)	0.043** (0.019)	0.049** (0.020)	0.043* (0.022)	0.051** (0.022)
Effect of moth. up.sec. ed. on females	0.038* (0.022)	-0.008 (0.019)	-0.002 (0.021)	0.015 (0.023)	0.018 (0.024)
Effect of fath. tert. ed. on males	-0.062** (0.030)	-0.041* (0.023)	-0.061*** (0.023)	-0.058** (0.025)	-0.061** (0.030)
Effect of moth. tert. ed. on males	0.011 (0.035)	0.087*** (0.027)	0.092*** (0.028)	0.087*** (0.031)	0.025 (0.037)
Effect of fath. tert. ed. on females	0.081*** (0.029)	0.083*** (0.025)	0.104*** (0.027)	0.095*** (0.029)	0.082*** (0.030)
Effect of moth. tert. ed. on females	0.029 (0.032)	-0.045 (0.028)	-0.037 (0.029)	-0.041 (0.033)	-0.010 (0.033)

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Table 14: OLS model with full sample: including neutral fields

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Parent upper sec.	0.001 (0.008)	0.009 (0.010)	0.017* (0.010)	0.010 (0.008)	0.006 (0.007)
Parent tertiary	-0.005 (0.009)	0.024** (0.011)	0.023** (0.011)	0.015* (0.009)	0.002 (0.008)
Female	0.000 (0.007)	0.080*** (0.009)	0.111*** (0.009)	0.063*** (0.007)	0.010* (0.006)
Constant	0.139*** (0.017)	0.225*** (0.022)	0.111*** (0.019)	0.143*** (0.018)	0.169*** (0.018)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.009	0.019	0.024	0.016	0.011
Observations	10527	10527	10527	10527	10527

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Table 15: OLS model with high school tracking

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Parent upper sec.	0.018 (0.017)	0.019 (0.015)	0.029* (0.015)	0.026 (0.016)	0.016 (0.018)
Parent tertiary	0.044** (0.020)	0.045*** (0.016)	0.051*** (0.017)	0.045** (0.019)	0.028 (0.022)
Female	0.053*** (0.018)	0.091*** (0.014)	0.149*** (0.015)	0.121*** (0.017)	0.023 (0.019)
Non-academic track	0.174*** (0.030)	0.088*** (0.025)	0.102*** (0.026)	0.089*** (0.029)	0.009 (0.031)
Academic track (other fields)	0.106*** (0.025)	0.057*** (0.022)	0.066*** (0.023)	0.050** (0.024)	0.012 (0.025)
Other	0.154 (0.099)	0.179*** (0.068)	0.193** (0.082)	0.264** (0.104)	0.164 (0.123)
Constant	0.007 (0.043)	0.139*** (0.037)	0.014 (0.037)	0.043 (0.039)	0.138*** (0.040)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.025	0.027	0.041	0.036	0.021
Observations	3616	5868	5187	3988	3008

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses. Sample restricted to individuals graduated after 2007 because of high school tracking data availability.

Table 16: OLS model with restricted sample: generation of parents only

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Parent upper sec.	-0.008 (0.016)	0.010 (0.013)	0.014 (0.013)	0.006 (0.014)	0.001 (0.016)
Parent tertiary	0.011 (0.018)	0.025* (0.014)	0.029** (0.015)	0.018 (0.016)	0.001 (0.018)
Female	-0.026* (0.014)	0.116*** (0.012)	0.159*** (0.012)	0.125*** (0.013)	-0.031** (0.015)
Constant	0.230*** (0.028)	0.234*** (0.024)	0.143*** (0.025)	0.167*** (0.025)	0.237*** (0.026)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.015	0.029	0.043	0.027	0.009
Observations	4077	6885	6206	4778	3589

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

Table 17: OLS model with restricted sample: generation of children only

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
	b/se	b/se	b/se	b/se	b/se
Parent upper sec.	0.041 (0.030)	-0.002 (0.025)	0.015 (0.026)	0.032 (0.030)	0.034 (0.032)
Parent tertiary	0.068** (0.033)	0.037 (0.027)	0.042 (0.029)	0.072** (0.034)	0.084** (0.037)
Female	0.077*** (0.025)	0.029 (0.021)	0.099*** (0.022)	0.069*** (0.026)	0.096*** (0.027)
Constant	-0.005 (0.036)	-0.052* (0.027)	-0.053* (0.029)	-0.048 (0.038)	-0.053 (0.042)
Area	Yes	Yes	Yes	Yes	Yes
Municip. size	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
R ²	0.018	0.013	0.018	0.025	0.032
Observations	1520	2399	2065	1469	1193

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Clustered standard errors in parentheses.

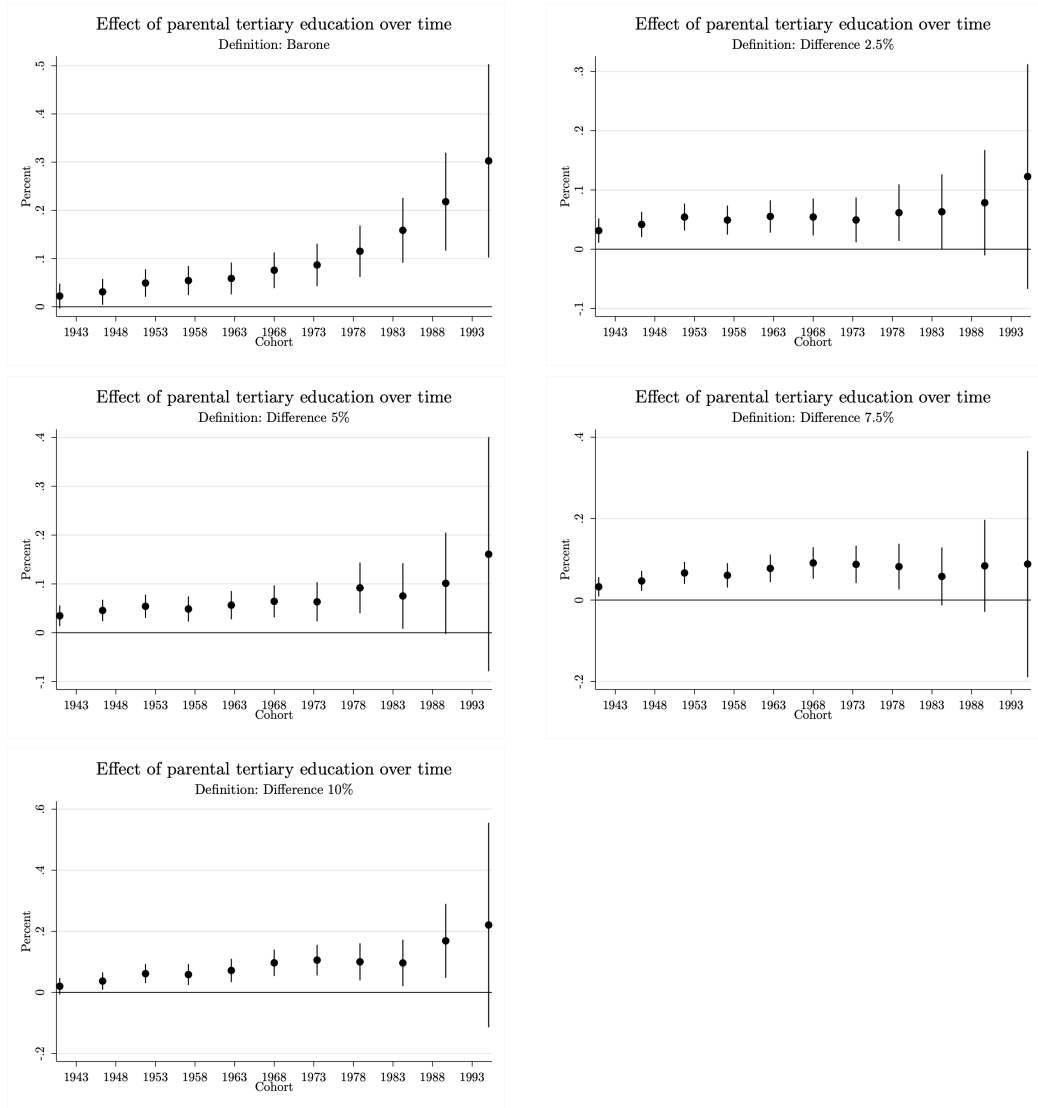


Figure 5: Estimated coefficients for the variable of interest (at least one parent with tertiary education) from the benchmark OLS model in eq.(1) by cohort. Confidence intervals at 90% level.

References

- Almalaurea. 2019a. *Graduates' employment status*. Bologna, Italy: Almalaurea.
- . 2019b. *Graduates' profile*. Bologna, Italy: Almalaurea.
- Banca d'Italia. 2022. "Survey on Household Income and Wealth." Data retrieved from Survey on Household Income and Wealth, <https://www.bancaditalia.it/pubblicazioni/indagine-famiglie/>.
- Barone, C. 2011. "Some Things Never Change: Gender Segregation in Higher Education across Eight Nations and Three Decades." *Sociology of Education* 84 (2):157–176.
- Barone, C. and G. Assirelli. 2020. "Gender segregation in higher education: an empirical test of seven explanations." *Higher Education* 79 (1):55–78.
- Basiglio, S., D. Del Boca, and C. Pronzato. 2024. "The Impact of the 'Coding Girls' Program on High School Students' Skills, Awareness and Aspirations." *CESifo Economic Studies* 70:271–283.
- Black, Dan A, Amelia M Haviland, Seth G Sanders, and Lowell J Taylor. 2008. "Gender wage disparities among the highly educated." *Journal of human resources* 43 (3):630–659.
- Blau, Francine D and Lawrence M Kahn. 2017. "The gender wage gap: Extent, trends, and explanations." *Journal of economic literature* 55 (3):789–865.

- Bovini, G, M De Philippis, and L Rizzica. 2023. “The origins of gender wage gaps: the role of school to work transition.” *Temi di Discussione (Working Papers) forthcoming, Bank of Italy* .
- Canaan, Serena and Pierre Mouganie. 2023. “The impact of advisor gender on female students’ STEM enrollment and persistence.” *Journal of Human Resources* 58 (2):593–632.
- Card, David and A Abigail Payne. 2021. “High school choices and the gender gap in STEM.” *Economic Inquiry* 59 (1):9–28.
- Carlana, M. 2019. “Implicit Stereotypes: Evidence from Teachers’ Gender Bias.” *The Quarterly Journal of Economics* 134 (3):1163–1224.
- Carlana, M. and L. Corno. 2021. “Parents and peers: Gender stereotypes in the field of study.” *Centre for Economic Policy Research Discussion Paper* DP16582.
- Carrell, Scott E, Marianne E Page, and James E West. 2009. “Sex and Science: How Professor Gender Perpetuates the Gender Gap. NBER Working Paper No. 14959.” *National Bureau of Economic Research* .
- Charles, M. and K. Bradley. 2009. “Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries.” *American Journal of Sociology* 114 (4):924–976.
- Charles, M. and M. Bradley. 2002. “Equal but Separate? A Cross-national Study of Sex Segregation in Higher Education.” *American Sociological Review* 67:573–99.

- Delaney, Judith M and Paul J Devereux. 2019. "Understanding gender differences in STEM: Evidence from college applications." *Economics of Education Review* 72:219–238.
- Eccles, J.S., C. Freedman-Doan, P. Frome, J. Jacobs, and K. Suk Yoon. 2000. "Gender-role socialization in the family: A longitudinal approach." In *The developmental social psychology of gender*, edited by T. Eckes and H.M. Trautner. Mahwah, United States: Lawrence Erlbaum Associates Publishers, 333–360.
- England, P. and S. Li. 2006. "Desegregation Stalled: The Changing Gender Composition of College Majors, 1971-2002." *Gender & Society* 20 (5):657–677.
- Goldin, Claudia, Lawrence F Katz, and Ilyana Kuziemko. 2006. "The homecoming of American college women: The reversal of the college gender gap." *Journal of Economic perspectives* 20 (4):133–156.
- Guiso, Luigi, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. 2008. "Culture, gender, and math." *Science* 320 (5880):1164–1165.
- McNally, Sandra. 2020. "Gender differences in tertiary education: what explains STEM participation?" Tech. rep., IZA Policy Paper.
- Terrier, Camille. 2020. "Boys lag behind: How teachers' gender biases affect student achievement." *Economics of Education Review* 77:101981.
- Triventi, M. and P. Trivellato. 2009. "Participation, performance and in-

equality in Italian higher education in the 20th century: Evidence from the Italian Longitudinal Household Survey.” *Higher Education* 57:681–702.

Triventi, Moris. 2010. “Something changes, something not. Long-term trends in gender segregation of fields of study in Italy.” *Italian Journal of Sociology of education* 2010 (5 (2)):47–80.

Vergolini, Loris and Eleonora Vlach. 2017. “Family background and educational path of Italian graduates.” *Higher Education* 73:245–259.

A Additional tables

Table A1: Logit benchmark model: estimated coefficients

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
Parent upper sec.	0.015 (0.071)	0.037 (0.053)	0.077 (0.057)	0.066 (0.068)	0.037 (0.081)
Parent tertiary	0.113 (0.080)	0.144** (0.057)	0.169*** (0.062)	0.171** (0.075)	0.112 (0.090)
Female	0.003 (0.064)	0.420*** (0.047)	0.681*** (0.052)	0.575*** (0.062)	-0.009 (0.072)
1946-1950	0.156 (0.167)	-0.149 (0.122)	0.098 (0.142)	-0.076 (0.148)	-0.054 (0.150)
1951-1955	-0.054 (0.175)	-0.189 (0.123)	0.159 (0.141)	0.062 (0.142)	-0.138 (0.155)
1956-1960	0.014 (0.172)	-0.025 (0.122)	0.321** (0.140)	0.196 (0.145)	0.047 (0.158)
1961-1965	0.367** (0.160)	0.417*** (0.116)	0.561*** (0.138)	0.259* (0.146)	-0.031 (0.164)
1966-1970	0.430*** (0.156)	0.262** (0.115)	0.584*** (0.135)	0.366*** (0.141)	0.128 (0.156)
1971-1975	0.592*** (0.157)	0.322*** (0.118)	0.649*** (0.136)	0.166 (0.156)	0.175 (0.157)
1976-1980	0.452*** (0.166)	0.260** (0.124)	0.440*** (0.150)	0.124 (0.165)	0.155 (0.166)
1981-1985	0.365** (0.168)	0.087 (0.134)	0.396*** (0.151)	0.256 (0.156)	0.370** (0.160)
1986-1990	0.719*** (0.181)	0.473*** (0.143)	0.562*** (0.173)	0.559*** (0.177)	0.658*** (0.180)
1991-1995	0.731*** (0.217)	0.752*** (0.170)	0.888*** (0.202)	0.927*** (0.205)	0.937*** (0.219)
Center	0.015 (0.074)	0.045 (0.056)	0.036 (0.060)	0.051 (0.071)	0.027 (0.086)
South and islands	-0.109 (0.072)	-0.008 (0.051)	0.009 (0.055)	-0.040 (0.068)	-0.099 (0.080)
More than 40k	-0.046 (0.065)	0.023 (0.049)	0.037 (0.053)	-0.021 (0.064)	-0.031 (0.075)
Constant	-1.313*** (0.155)	-1.153*** (0.115)	-1.673*** (0.138)	-1.570*** (0.143)	-1.275*** (0.143)
Pseudo R ²	0.011	0.019	0.030	0.023	0.011
Observations	5597	9284	8271	6247	4782

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Table shows logit coefficients and relative clustered standard errors in parentheses.

Table A2: Logit benchmark model: marginal effects

	(1)	(2)	(3)	(4)	(5)
	Barone	2.5%	5%	7.5%	10%
Parent upper sec.	0.003 (0.014)	0.008 (0.011)	0.016 (0.012)	0.012 (0.013)	0.007 (0.014)
Parent tertiary	0.022 (0.016)	0.032** (0.013)	0.035*** (0.013)	0.033** (0.014)	0.020 (0.016)
Female	0.001 (0.012)	0.092*** (0.010)	0.143*** (0.011)	0.110*** (0.012)	-0.002 (0.013)
1946-1950	0.027 (0.028)	-0.030 (0.025)	0.018 (0.025)	-0.013 (0.025)	-0.009 (0.025)
1951-1955	-0.009 (0.029)	-0.038 (0.025)	0.029 (0.025)	0.011 (0.025)	-0.022 (0.025)
1956-1960	0.002 (0.028)	-0.005 (0.025)	0.061** (0.026)	0.036 (0.026)	0.008 (0.027)
1961-1965	0.067** (0.028)	0.093*** (0.025)	0.112*** (0.026)	0.048* (0.027)	-0.005 (0.028)
1966-1970	0.079*** (0.027)	0.057** (0.024)	0.117*** (0.025)	0.070*** (0.026)	0.023 (0.027)
1971-1975	0.114*** (0.028)	0.071*** (0.025)	0.131*** (0.025)	0.030 (0.028)	0.031 (0.028)
1976-1980	0.084*** (0.030)	0.057** (0.026)	0.085*** (0.028)	0.022 (0.030)	0.027 (0.029)
1981-1985	0.066** (0.030)	0.018 (0.028)	0.076*** (0.028)	0.048 (0.029)	0.069** (0.030)
1986-1990	0.142*** (0.035)	0.106*** (0.032)	0.112*** (0.035)	0.111*** (0.036)	0.132*** (0.037)
1991-1995	0.145*** (0.045)	0.174*** (0.040)	0.187*** (0.044)	0.196*** (0.045)	0.197*** (0.049)
Center	0.003 (0.015)	0.010 (0.012)	0.008 (0.012)	0.010 (0.014)	0.005 (0.016)
South and islands	-0.021 (0.014)	-0.002 (0.011)	0.002 (0.011)	-0.008 (0.013)	-0.017 (0.014)
More than 40k	-0.009 (0.013)	0.005 (0.011)	0.008 (0.011)	-0.004 (0.012)	-0.006 (0.014)
Pseudo R ²	0.011	0.019	0.030	0.023	0.011
Observations	5597	9284	8271	6247	4782

Note: * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. Table shows marginal effects and relative clustered standard errors in parentheses.