The Origins of the Gender Pay Gap: Education and Job Characteristics*

Giulia Bovini† Marta De Philippis‡ Lucia Rizzica§

November 6, 2024

Abstract

This paper analyzes the role that field of study choices and firm and job characteristics play in explaining gender pay gaps among university graduates at the beginning of their careers. We exploit a unique administrative population-wide dataset for Italy. We find that choices of university majors alone explain almost 60% of the early career gender pay gap, since women tend to graduate in fields that give access to lower-paying firms and jobs. Sorting across firms and jobs only accounts for 20% of the gap, indicating that its role is largely mediated by field specialization choices made while still in education.

JEL Codes: J24, J31, J45, J62, J82.
Keywords: gender pay gap, field of study, tertiary education, sorting.

*We thank Antonio Accetturo, Fabrizio Colonna, Federico Cingano, Salvatore Lattanzio, Sauro Mocetti, Matteo Paradisi, Daniele Paserman, Roberto Torrini, Eliana Viviano, and all seminar participants to the Bank of Italy work in progress seminar, the II Istat-Bank of Italy workshop on the Italian production system (TRASPI), the 2023 EEA-ESEM Conference, the IV Bank of Italy Human Capital workshop, the University of Trondheim II Workshop on Education Economics and Policy, and the VII Bank of Italy-CEPR Labour Workshop for beneficial comments. We are indebted to Stefano De Santis for his invaluable help in assembling and making the data available. *The views expressed in the article are those of the authors only and do not involve the responsibility of the Bank of Italy or the Eurosystem.*

†Bank of Italy (giulia.bovini@bancaditalia.it), via Nazionale 91, Rome, Italy

‡Bank of Italy and IZA (marta.dephilippis@bancaditalia.it), via Nazionale 91, Rome, Italy

§Bank of Italy (lucia.rizzica@bancaditalia.it), via Nazionale 91, Rome, Italy

1 Introduction

Despite some convergence, women still earn less than men. A growing literature documents the prominent role behind such a gap of *firms and jobs characteristics*, like firm-specific pay premiums - as women both sort in lower-paying firms and, within firms, receive a smaller share of the surplus (Cardoso et al., 2016; Sorkin, 2017; Bruns, 2019; Casarico and Lattanzio, 2024) - occupations and sectors (Sloane et al., 2021; Lordan and Pischke, 2022; Delfino, 2024), and migration and commuting patterns (Le Barbanchon et al., 2021; Caldwell and Danieli, 2024). Gender gaps, however, manifest much earlier, notably when men and women are still in *education*: girls have overtaken boys in terms of years of schooling and outperform them academically (Blau and Kahn, 2017) but select very different fields of study (Kahn and Ginther, 2018; Sloane et al., 2021). This may itself generate large gender pay gaps. Indeed, in increasingly specialized economies, the *type* of human capital investment is found to matter even more than the *level* of education: earnings disparities across university majors can be as large as the overall college–high school premium (Altonji et al., 2012, 2016).

In this paper, we examine to what extent gender differences in field of study choices explain why women and men sort into different firms and jobs and, consequently, earn different wages. Answering this question is important for the optimal design of both the timing and the target of policies to tackle gender gaps in the labor market.

We focus on early career gender gaps among recent cohorts of university graduates, a population interesting to study for various reasons. Arellano-Bover et al. (2023) show that gender disparities within newer cohorts upon entering the labor market remain large and constitute an important determinant of the aggregate gender earning gap dynamics. Moreover, the literature has documented that initial gaps do not shrink, but rather widen over the life cycle, largely due to the unequal effect of parenthood for men and women (e.g., Kleven et al., 2019, 2020). Finally, the share of youths with a tertiary degree is increasing in most countries[1] and fields of study at university tend to provide students with very specialized skills, making the assessment of their role in earnings gender disparities crucial.[2]

[1]Across OECD countries, the share of 25-34 year-olds with some tertiary education has risen from 24% in 1998 to 47% in 2022.

[2]We focus on university graduates rather than also considering high school graduates also because the curriculum of fields of study at university tends to be more comparable across countries than that of tracks in high school. This increases the external validity of the findings we document in our setting - Italy.

We model the process that generates early career gender gaps as having two steps. First, women and men choose which university major to graduate from. Second, they enter the labor market and match with a job and an employer. Operationally, we start our analysis by examining to what extent gender heterogeneity in field of study choices can explain the *aggregate* gender gap in daily wages. Next, we look at within-field differences in the type of employers men and women work for and of jobs they hold, exploring to what extent they account for *within-field* gender pay gaps. We do this through a series of Oaxaca-Blinder-type decomposition exercises.

The setting of our analysis is Italy. Thanks to a unique, newly assembled, match between several population-wide administrative registries - spanning education records, tax declaration forms, matched employer-employee data, and firm balance-sheet data - for the cohorts graduating in 2011-18 we observe: *(i)* their field of study choices, their university of graduation and their academic performance at university, together with *(ii)* labor market outcomes over 2012-19. We focus our main analysis on graduates who work as private-sector employees because this is the sample for whom the broadest set of information is available: we observe annual gross labor earnings and days of employment, as well as very detailed characteristics of the job (fine-grained occupation and sector, contract type) and the employer (firm identifier, location, workforce composition, size, and balance-sheet variables).[3] However, the part of the analysis that we can replicate for self-employed and public-sector workers delivers similar findings.

We document several important facts. First, the early career gender gap in daily wages is already sizable: it is 21% one year after graduation and widens to 25% five years after.

Second, there are pronounced gender differences in the choice of university majors. While in our cohorts 60% of graduates are females, the share drops significantly in most STEM fields: for instance, only 27% of graduates in ICT and engineering are women. To synthesize this evidence, we build - for each of almost 30 majors - a measure of financial gains based on the median earnings of its native male graduates five years after graduation. We show that women, especially those at the top of the ability distribution, are systematically more likely to choose majors with lower expected average financial returns.

[3]We provide evidence on the magnitude of gender gaps one and five years after graduation. In the paper we mainly report the findings for gaps five years after graduation; in some robustness tests, we show that our results hold also when we look at outcomes 1 year after graduation.

Third, based on an Oaxaca-Blinder decomposition, we show that differences in field of study choices alone explain almost 60% of the average gap in daily wages. Although differences in major choices are wider for high-ability students, their explanatory power for gender gaps is greater at the bottom of the distribution. Other differences in education choices and outcomes, such as the university attended and academic performance, matter very little in comparison. Why do fields of study play such an important role? We show that, even if we consider labor market outcomes of a "representative" graduate in each field of study - i.e. with fixed observable characteristics (native, man) - *(i)* different majors are associated with access to very different jobs and employers, and *(ii)* women are more likely to choose majors whose graduates work in lower-paying, less productive, and closer-to-the-birthplace firms and hold jobs in lower-skilled occupations and with a higher incidence of part-time contracts.

Fourth, despite the prominent role of field of study choices, there exist some gender gaps in daily wages even within majors. Within-field Oaxaca-Blinder decompositions uncover three interesting findings. First, even a very rich set of firm and job characteristics only explains, on average across fields, 40% of the within-field average gender gap. The average explained component is even lower among the high-earners (20%). Second, at the mean and the bottom of the distribution, the characteristic that matters the most is the higher prevalence of part-time contracts among women: if we net out differences in daily hours worked, by looking at full-time-equivalent wages, the average explained component drops to 20%. Notably, the characteristics of the employer and the distance between the workplace and the birthplace - which have been found important for the aggregate gender gap by several studies - play a very little role once one explicitly accounts for gender differences in field of study choices. Third, women tend to graduate in lower numbers from fields where the unexplained component of the gap is larger, as is the case for STEM majors. This suggests that women shy away from fields with a higher uncertainty about the reasons behind gender wage disparities. We speculate that these could be majors where the role played by discrimination, or negotiating abilities, or the cost of not being available to work very long hours is greater.

All in all, a regression that sequentially controls first for education choices, and later also for firm and job characteristics at a higher level of granularity than feasible in the Oaxaca-Blinder decomposition, confirms that the former explain alone over 50% of the early career gender pay gap while the latter account for only 20%. If we were to run a regression that only included firm and job characteristics, we would instead conclude that disparities in such characteristics would account for 70% of the gender gap in daily wages.

While our main and most comprehensive analysis is based on private-sector employees only, the analysis we can replicate for self-employed and public-sector workers delivers similar findings. We document that the raw gender pay gap among self-employed is very similar to that of private-sector employees; it is smaller, but still sizable, among public-sector workers. Furthermore, when we perform the Oaxaca-Blinder decomposition on the role of educational factors for all workers, we still find that fields of study account for 50% of the annual earnings gap five years after graduation. We cannot assess the role of firms and job characteristics for all workers because such information is only available for private-sector employees.

These findings carry relevant policy implications. They suggest that the toolkit of policy-makers seeking to tackle labor market gender gaps should extend beyond policies promoting a more equal division of childcare. Disparities materialize before men and women start a family: the child penalty greatly magnifies a gap already present at the onset of the career, even among highly educated individuals in younger cohorts. Based on our analysis, actions targeting field-of-study choices of adolescents who are still in education might be particularly effective.

The remainder of the paper is organized as follows. Section 2 places our contribution in the existing literature. In section 3 we introduce the setting, the data, and the sample. Section 4 presents some descriptive evidence on the early career gender gaps in daily wages, in field of study choices, and in firm and job characteristics within and across fields. In section 5 we present the Oaxaca-Blinder decomposition analysis to assess the role of education and employment characteristics in explaining aggregate and within-field gender pay gaps. Section 6 discusses the implications of our findings for understanding the sources of gender wage gaps. Section 7 presents some alternative decompositions, in terms of samples, outcomes and time horizon as robustness checks. Section 8 concludes.

2 Related literature

By providing new facets on early career gender gaps for the highly skilled in Italy, we contribute to the literature on the pre-market and market determinants of gender pay disparities.

While the body of work on this topic is vast, comprehensive and detailed analysis on both factors *jointly* are scarcer because of the lack of high-quality data that record at the same time granular information on workers' human capital, their jobs, and their employers. On the one hand, survey data that collect information on fields of study typically only contain coarse information on the employer (e.g., class size and aggregate sector) and the job (e.g. broad occupation, full-time status). Overall, survey-based evidence from the US points to a relevant role of major choices but estimates vary greatly across studies, from 10% to more than 90%.[4] Looking at European countries, Machin and Puhani (2003) show that in the 1990s fields of study choices explained a significant part of the gender wage gap among university graduates in the UK and Germany. Francesconi and Parey (2018) document that university majors were the most important explanatory factor for the early-career gender gaps among university graduates in 1989-2009. Based on administrative data for local-born college graduates in Milan[5] - the financial center of Italy -, Anelli and Peri (2015) document a sizable gender gap in annual gross income 5 to 15 years after graduation (37%): major choices explain a third of it (a half when correcting for hours worked).[6]

On the other hand, administrative matched employer-employee data contain detailed information on a worker's employer – allowing to explore the role of sorting across firms for the gender wage gap - but typically do not include employees' field of highest degree. Based on AKM decomposition exercises (Abowd et al., 1999), in advanced economies differences in firm premiums are found to account for between one-fifth and one-third of the average

[4] As noted in their review of the literature by McDonald and Thornton (2007), who find that fields of study can account for almost the entire gender gap among college graduates' starting salaries, this heterogeneity in estimates can reflect differences in the period considered, the regression specification and the controls used, as well as when earnings are observed (soon after graduation like in our study or later in the career) and the degree of aggregation of fields of study. Early works on the US include Eide (1994), Brown and Corcoran (1997), Loury (1997), Weinberger (1998), Joy (2003), Graham and Smith (2005), Black et al. (2008), Weinberger (2011).

[5] The sample is limited to individuals who graduated from a college-preparatory high school (*Liceo classico* or *Liceo scientifico*) in Milan in 1985-95 and later enrolled in a university in Milan: as they attended the most demanding high-school tracks in one of the richest cities in Italy, these students are positively selected, in terms of background and ability, compared to the average graduate. Our dataset covers the entire country and also college graduates who attended technical and vocational high schools (see Section 3 for more information on the Italian upper secondary school system).

[6] Based on survey data on a sample of the 2007 cohort of Italian graduates observed 4 years after graduation, Piazzalunga (2018) on the other hand reports a small gender gap in hourly wages (around 6%), which is not explained by any observable characteristics, including fields of study. Yet, the wage information is missing for individuals with precarious employment, who are disproportionately females.

gender wage gap among all workers (Cardoso et al., 2016; Sorkin, 2017; Bruns, 2019; Li et al., 2023; Casarico and Lattanzio, 2024).

Thanks to our newly assembled dataset, we can jointly assess the relative role of pre-market and market specialization in great detail. We highlight that for university graduates the former is a strong determinant of the latter and, hence, explains more than half of the aggregate gender gap. We complement and add on two recent contributions that have stressed this. Based on US data on university graduates born in 1950-90, Sloane et al. (2021) document that women sort into majors and occupations associated with lower potential earnings: majors explain a large portion of the gender earnings gap for all cohorts; the role of occupations has, instead, declined over time spurred by gender convergence in occupations within some of the highest-paying majors. We show that women also sort in lower-quality firms and contracts and that sorting across majors largely mediates sorting in the labor market. Based on administrative data from Chile (1977-2000 graduating cohorts working as private-sector employees in 2005-19), Huneeus et al. (2021) show in an AKM decomposition exercise that college majors mediate access to high-paying firms and, hence, explain more than two-thirds of gender disparities in firm premiums. While our data is not amenable to an AKM decomposition, we see it as a valuable contribution to explore the many dimensions along which employers of men and women differ and to assess which dimension of quality matters more, rather than collating them into a firm fixed effect. Furthermore, we are among the few papers that look at how the relative roles of pre-market and market determinants change along the distribution of wages. Finally, by estimating for each field of study the unexplained component of the gender gap, we identify the majors with higher uncertainty about the sources of labour market gender disparities, which turn out to be the male-dominated ones (STEM majors, in particular).

3 Institutional setting, data and sample

2.1 Institutional setting – The Italian education system includes primary (five years), lower-secondary (three years), upper-secondary (five years), and tertiary education. Compulsory education starts when students turn 6 and ends when they turn 16; public schools offer it for free. When enrolling in upper secondary school, students can choose the preferred track

- academic, technical, or vocational - as well as the track-specific specialization (sub-track).[7] To obtain the diploma, at the end of upper secondary school pupils sit a final national exam that is sub-track specific. All 5-year upper secondary school diplomas allow students to pursue tertiary education in any field[8] but some majors have a fixed number of seats and thus enrollment is subject to an entrance exam. Differently from the US, students declare their major at the same time as they enroll in university and all courses are tightly focused on the specialty of the chosen major (Anelli and Peri, 2015).

Since 1999 Italy universities have adopted a 3+2 system, like many other European countries. A Bachelor's (*Laurea*) can be obtained after three years of study (1st cycle degree). Students may then specialize with a two-year Master's (*Laurea Specialistica*, 2nd cycle degree). Some majors, however, still award a five- or six-year degree (*Laurea Magistrale a Ciclo Unico*, single-cycle degree).[9] Enrollment in 1st-cycle and single-cycle degrees is open to all students with a 5-year upper secondary school diploma. Access to 2nd-cycle degrees is open to those with a 1st-cycle or single-cycle degree, provided that the curriculum included certain field-specific exams.

2.2 Data – For the first time, we were able to combine administrative records from multiple population-wide registries to assemble a dataset that contains rich information on education choices, early labor market outcomes, and firm and job characteristics.

The population we study consists of all students who graduated from an Italian university in 2011-18. From education records[10], we obtain detailed information on education choices and academic performance: *(i)* the university name and location (for example, University *La Sapienza*, Rome campus); *(ii)* the type of degree: Bachelor (1st cycle) degree, Master (2nd cycle) degree, and 5- or 6-year single-cycle degree; *(iii)* the major (for most of the analysis we aggregate fields of study in 27 major groups); and *(iv)* final grade and age at graduation. For university graduates who obtained their upper secondary school diploma in 2011-18, records also report the type of high school they attended and their final grade.

[7]For example, within the academic track, there is a humanities specialization and a scientific studies specialization.

[8]Enrollment rates in university are highest for students who attend the academic track of high school and lowest for students who attend the vocational track.

[9]Such majors include law, primary teacher education, architecture, pharmacy, visual arts, music, medicine, and dentistry.

[10]Maintained by the Ministry of Education and Merit and the Ministry of University and Research.

From annual tax records[11] that cover the period 2011-19, we draw two sets of information on university graduates: *(i)* socio-demographic characteristics: gender, age, place of birth, marital status, as well as parental gross labor and overall income in the year when the student graduates; and *(ii)* annual gross labor income[12] split according to the source (dependent employment or self-employment). For employees, the records report days of employment[13] so that we can compute daily wages. We cannot, however, compute hourly wages as information on hours worked is not available. For these workers, we also match information from three other sources. The first consists of mandatory reporting forms, available for 2012-19, that firms fill when they start or end a contract with a worker.[14] This register is thus particularly suited to track the first main job held by fresh university graduates, which typically starts with a new contract. From this archive, we retrieve the duration (permanent or temporary) and schedule (full-time or part-time) of the contract, the 4-digit occupation (Istat 2011 classification), the 6-digit sector (Nace 2007 classification), the municipality of work, and the employer (anonymized) identifier.[15] The second source is based on mandatory reporting forms, available for 2014-18, that private-sector firms fill out for all employees on their payroll, not just for new hires and separations.[16] We leverage this matched employer-employee dataset to build some firm-level measures of workforce composition, averaged over the entire period: the number of employees, the share of female workers, the average workforce's education level, and the gender wage gap at the mean and the 90th percentile.[17] Third, for private sector firms we also draw information, averaged over 2012-19, from balance sheets,

[11] Collected by the Ministry of Finance.

[12] In tax declaration forms, this encompasses income from employment and from other sources related to employment.

[13] In tax declaration forms, these are the days of employment based on which workers can claim some tax deductions.

[14] *Comunicazioni Obbligatorie* are sent to the Ministry of Labour.

[15] If a worker has multiple contracts in a given year, we select the prevalent employer and contract according to the following two-step procedure. First, we consolidate all contracts between a (worker, firm) pair, and we define the main employer as the one for which the individual has worked the longest in a given year. Second, within all contracts with the main employer, we retain the characteristics of that with the longest duration. In the case of ties in either step, we randomly select the prevalent employer or contract.

[16] *Dichiarazioni Uniemens* are sent to the National Social Security Institute.

[17] These characteristics are computed on the employees who: *(i)* have non-missing information on annual income and months worked; and (*ii*) are between 20 and 60 years old. If a worker has more than one spell with a firm in a given year, we select that with the highest annual income. In the case of ties, we retain that with the longest duration (months worked in the year). If some ties remain, we select one spell randomly.

which include measures of productivity (value added per worker), and firm age.[18]

2.3 Sample – We study early career gender gaps one and five years after graduation to capture the situation in the first job and after some years of labor market experience. For conciseness, we focus on graduates from 2nd-cycle and single-cycle degrees. After completing five years of tertiary education, the most common next step in Italy is to enter the labor market and search for a job: these graduates, hence, are not only the larger group but also those for whom it is most interesting to study the college-to-work transition.[19] Labor market outcomes 1 year after graduation are available for all cohorts (2011-18; around 600.000 women and 415.000 men), while outcomes five years after graduation are available only for the 2011-14 graduating cohorts (hence, the number of observations roughly halves).

To study gender gaps in daily wages we focus on the subset of university graduates who one or five years after graduation *(i)* work as private sector employees, and *(ii)* are not studying. Importantly, we can observe these outcomes only for individuals who remain in Italy after graduating: those who move abroad to either study or work are therefore not included in our analysis.[20] We define private-sector employees as individuals who draw the main source of labor income from dependent employment and hold their prevalent job in a private-sector firm.[21] We restrict our main analysis to these workers in order to be able to jointly evaluate the role of educational choices and also that of employer and job characteristics, which are not available for self-employed and public sector workers. Note, however, that in Section 7, we show that our findings on the role that education choices and outcomes play in the aggregate gender gap are confirmed when we include public employees and the self-employed in the sample of analysis. The restrictions applied leave us with a sample of roughly 360.000 graduates (around 150.000 men and 210.000 women) one year after graduation and roughly 185.000 graduates (around 75.000 men and 110.000 women)

[18]Notice that our records allow us to have balance sheet information for all firms in the private sector, not only for incorporated ones, as usually available.

[19]Based on administrative data on enrollment into Bachelor's, Master's, and single-cycle degrees made available by the Ministry of University and Research: (*i*) among students who first enrolled in university in 2011-18, 86% enrolled in a Bachelor's and the remaining 14% in a single-cycle degree; (*ii*) among students who obtained a Bachelor's in 2013-18, 52% then enrolled in a Master's.

[20]According to Istat (2023), the number of expatriating Italian graduates in the 25-34 age class has increased over time: in this sub-population, the net migration rate over the decade 2012-2021 was negative (-79.000). In our dataset, we cannot distinguish individuals who moved abroad from those who remain in Italy but neither keep studying nor start working.

[21]The prevalent employer and contract are defined according to the procedure described in footnote 15.

five years after graduation. This is the sample based on which we provide descriptive evidence about early career gender gaps in Section 4. Appendix Section B reports more details on the sample selection.

When we analytically decompose the role of fields of study, firms, and jobs based on an Oaxaca-Blinder decomposition (Sections 5.1 and 5.2), we further restrict the sample to graduates for whom the full set of controls across *all* our regression specifications is available ("Oaxaca sample", henceforth). This leaves us with a sample of roughly 250.000 individuals (115.000 males and 135.000 females) one year after graduation and 130.000 individuals (60.000 men and 70.000 women) five years after graduation.

Appendix Section C compares some observable characteristics for the two main samples as well for public-sector employees and the self-employed. Women in the Oaxaca sample appear to be negatively selected compared to the full sample but the magnitude of the gender gap is comparable, suggesting that the findings of the decomposition exercise can be generalized to all private-sector employees.

4 Descriptive evidence

3.1. Early career gender wage gaps – At the start of the career, gender disparities among university graduates working as private-sector employees are sizable: one year after graduation, women's daily wage is already 21% lower than men's; five years after graduation, the gap further widens to 25%. Such gaps are larger than those among public-sector employees (13% and 9% one and five years after graduation, respectively) while being similar to disparities among the self-employed (Appendix Table C.1).[22]

[22] AlmaLaurea – a consortium of almost all Italian universities – publishes every year statistics on the gender gap in earnings among university graduates one and five years after graduation, based on a survey (*Indagine sulla condizione occupazionale dei laureati*). Graduates self-report net monthly earnings (as opposed to gross annual earnings recorded in tax declaration forms), censored below 200 and above 3000 euros. Because the tax system is progressive and very low (high) earners are more prevalent among women (men), these differences in how income is measured can result in a smaller gap. For employed persons who graduated in 2011-2018, the 1-year gap turns out to be very similar to our estimates (18-24%, depending on the year) among second-cycle degree holders while somewhat narrower among the smaller population of single-cycle degree holders (12-16%). The 5-year gaps for 2011-14 graduating cohorts are smaller than ours (roughly 13%) but are computed for the sub-sample of full-time workers who started their current job after graduation. Our estimates are larger than disparities found in the overall population: being computed on men and women at the start of their careers and with the same level of education, they are net of compositional differences which tend to shrink aggregate gender gaps since women are usually more positively

3.2. Educational choices – Among the 2011-18 cohorts of university graduates, 59% are females. Women also obtain their degree with a final grade that - for each group of majors - is as high or even higher than men's (Appendix Figure A.1). However, women and men significantly differ in their field of study choices. Appendix Figure A.2 reports the share of females across the groups of majors. Compared to the overall average, women are under-represented in most STEM fields: they amount to only 27% of graduates in Engineering and ICT and only 46% of graduates in the fields of Mathematics, Statistics, Physics, and Chemistry. On the other hand, they constitute the largest majority of graduates in Arts and Humanities (76%) and Education Sciences (94%).[23] This pattern is not unique to Italy: Eurostat data indicate that, on average in 2022, in EU countries the share of females among all graduates and graduates in Engineering, ICT, and natural sciences was very similar to Italy's.

All in all, women select less remunerative fields of study. The scatterplot in Figure 1 shows for each of 27 majors the share of female graduates (on the x-axis) together with the financial return (on the y-axis) measured by the median annual labor earnings of male, native graduates five years after graduation.[24] Following Sloane et al. (2021) we interpret this measure as the potential payoff of a major, since it is computed on a sub-population whose earnings are less likely to be affected by discrimination or tenure gaps due to taking up caregiving responsibilities. First, there is large variability in majors' expected payoffs, in line with what Kirkeboen et al. (2016) document. Second, a striking negative relationship emerges: women are more likely to graduate from majors that offer lower payoffs to their median male native graduate. This, in turn, translates into a 13% gender gap in the expected financial remuneration of field of study choices, a figure in line with Sloane et al. (2021) for the US cohorts of graduates most comparable to ours. Separately for women and men, Appendix Figure A.3 plots the expected economic gains of the chosen majors, defined as in Figure 1, against a measure of ability (the final grade in high school). Men choose higher-payoff majors

selected into the labor force than men (Olivetti and Petrongolo, 2008).

[23] Notice that in Italy, as stated before, students can choose tracks also in upper secondary school, even if these tracks are less specialized than university majors and do not preclude any field of study choice later in the educational career. Also in secondary school, females are less than 50% of graduates in the scientific sub-track of the academic track (as opposed to 81% in the language and art sub-track) and as low as 15% in the ICT sub-track of the technical track (as opposed to 55% in the business and economics track).

[24] The share of female graduates is computed on all cohorts (2011-18). The financial returns of university majors are computed on the sample of male and native graduates of 2011-14 who five years after graduation work and are not still in education (including public-sector employees and self-employed).

at any level of ability but the gender gap widens at the top of the distribution, largely shaped by the lower propensity of women to enroll in STEM majors which are associated with the highest returns for the median graduate. This finding aligns with what Anelli and Peri (2015) and Campbell et al. (2022) document for Italy (limited to the city of Milan) and UK students, respectively.[25]

3.3. Firm and job characteristics: overall gaps – Table 1 displays the characteristics of jobs and firms, by gender, five years after graduation. The likelihood of holding a temporary or a part-time contract is 10 p.p. and 19 p.p. higher among women, respectively. College-educated are mostly found in medium-to-high-skill occupations: however, men are much more likely to work in high-skilled occupations (as professionals and technicians) and women in medium-skill occupations (as clerical staff or sales workers). Men are more likely to be employed in the industrial sector, and women in non-market services (health, education, and NGOs, for instance). Interestingly, women work 38 kilometres closer to their birthplace (men on average work approximately 250 kilometres away from their birthplace).

The characteristics of the employer are also very different between genders. Fresh female graduates are on average more likely to be employed in firms that are younger and smaller, with a far larger share of females in the workforce (57% against 36% among firms that employ male workers) and with a lower gender wage gap both at the mean and, especially, at the top of the distribution. Notably, these firms pay roughly 15% lower average wages and have a 20% lower value added per employee, a canonical measure of firm productivity.

3.4. Firm and job characteristics: across-field gaps – To capture to what extent aggregate gender differences in firm and job characteristics reflect between-fields-of-study gaps, we adopt the same approach used to show heterogeneity in majors' payoffs (Figure 1). For each major, we compute the average value of a given firm or job characteristic for the median male native graduate, five years after graduation. We plot these measures, which capture which jobs and firms a representative graduate has potentially access to, against the share of female graduates in each major (Figure 2). First, cross-field disparities in firm and job characteristics are sizable: as an example, even among native men, humanities graduates

[25] Compared to men, they find that high-achieving women enroll in degrees (defined as the interaction between a faculty and a major) associated with potential earning - computed based on the earnings, five years after graduation, of previous cohorts of graduates - that are roughly eight percentiles lower; the gap is driven by differences in major choices rather than in the university attended.

tend to work in less productive, smaller firms, that are closer to home and pay on average lower wages than STEM graduates. They are also much more likely to hold a part-time contract and to work in a lower-skill occupation. Second, a pattern as striking as in Figure 1 emerges: the share of women is systematically lower in fields whose graduates work in higher quality firms (larger, higher-paying, more productive, father-away-from-home firms), and hold better jobs (full-time contracts and highly skilled occupations). All in all, in our setting majors appear to strongly mediate access to firms and jobs.

3.5. Firm and job characteristics: within-field gaps – Figures 3 and 4 display within-field-of-study gender gaps in daily wages and firm and job characteristics. The first bar plots the aggregate gap, the second bar the average within-field gap, and the following four bars the average within-field gaps separately for different groups of majors.

Average within-field gaps are always much smaller than aggregate ones. This confirms that sorting into different fields of study largely mediates early career gender gaps in wages and in firm and job characteristics. Nonetheless, there are still some differences within majors. For instance, within-major gaps in daily wages are approximately 13% on average. Moreover, women are employed in firms that are 6% less productive than boys graduated from the same major and pay on average 5% less. Finally boys are, on average within-field, almost 4 percentage points more likely to be employed in high-skilled occupations and in firms where the gender pay gap at the top of the distribution (i.e., likely referred to top managers) is 2 p.p. larger.

There is no clear pattern when looking at whether some groups of majors display systematically higher within-field gaps. If anything, disparities in narrow STEM fields are smaller or, in some cases, even reverted in sign (e.g., in the case of average firm size and distance from the birthplace) than in other fields.

To sum up, the descriptive evidence suggests that differences in fields of study choices play an important role in explaining the early career gender pay gap: cross-field disparities in the firms and jobs (male native) graduates have access to are sizable. Yet, even within majors, young females hold somewhat lower-quality jobs. To quantify the relevance of between- and within-fields factors, we turn to the decomposition exercise outlined in Section 5.1.

5 Oaxaca-Blinder decomposition analysis

In this section, we present the results of our two-step Oaxaca-Blinder decomposition analysis. We show which factors can explain, in statistical terms, the average differences in earnings between male and female graduates. Under the assumption that there are no innate gender-specific comparative advantages for certain majors or jobs, this type of analysis informs on how the early-career gender gap would change if men and women made more similar education and employment choices.

5.1 Empirical strategy

To study the joint role of fields of study, firms, and jobs for the early-career gender gap, we adopt a two-step approach. First, we show how much of the *aggregate* gender gap in daily wages can be attributed to differences in education choices. Second, *conditional on education choices*, we analyze how the residual gender gap is shaped by differences in the type of jobs and employers young women and men match with and how much remains unexplained. This two-step approach captures the sequential process where men and women first select a major and then, once they graduate, look for a job.

Let w^g_{it} be the (log) daily wage earned in year t by individual i who belongs to one of two groups g (women and men). In the first step of the decomposition exercise, equation 1 models w^g_{it} as a function of a set of $\mathbf{\Omega}$ observable characteristics (X) that capture education choices.

$$w^g_{it} = \beta^g_0 + \Sigma_\omega \beta^g_\omega X^g_{it\omega} + \eta^g_{it} \tag{1}$$

In particular, X includes a set of dummies for 27 majors, a set of dummies for the various university identifiers, and a dummy that takes the value 1 if the university is located in a region different from that of birth. Furthermore, it includes two measures of academic performance: the final grade at university and the age at graduation. Finally, it contains some basic demographic and background controls: cohort and region-of-birth fixed effects, marital status, and dummies for the income ventile of the parents.

Starting from equation 1, we perform a standard Oaxaca-Blinder decomposition to distinguish the explained and the unexplained components of the gap in daily wage between the two groups:

$$w_{it}^{men} - w_{it}^{women} = \underbrace{\Sigma_{\omega}\hat{\beta}_{\omega}^{men}(\overline{X}_{\omega}^{men} - \overline{X}_{\omega}^{women})}_{\text{Explained } (E)} + \underbrace{(\hat{\beta}_{0}^{men} - \hat{\beta}_{0}^{women}) + \Sigma_{\omega}\overline{X}_{\omega}^{women}(\hat{\beta}_{\omega}^{men} - \hat{\beta}_{\omega}^{women})}_{\text{Unexplained } (U)} \quad (2)$$

The explained component shows which part of the gender wage gap stems from differences in average characteristics between men and women, assuming that the returns to each characteristic for women are the same as those for men; the unexplained component accounts for the remaining part of the gap.

In the second step of the decomposition exercise, we focus on the within-field gender gap in daily wages. We do so because, in the first step, we find that, among education controls, fields of study are by far the most important explanatory factor of aggregate gender wage gaps (Section 5.2). Let $\boldsymbol{J}$ denote the set of fields and j denote the $j-th$ field in the set. For each field j[26], we can relate the within-field gender gap in daily wages to a set of $\boldsymbol{\Omega}$ firm and job characteristics included in Z:

$$w_{it}^{g,j} = \alpha_{0}^{g,j} + \Sigma_{\omega}\alpha_{\omega}^{g,j}Z_{it\omega}^{g,j} + \epsilon_{it}^{g,j} \quad (3)$$

where $w_{it}^{g,j}$ is the daily wage earned in year t by individual i who belongs to group g and graduated from field j. In this specification, Z is a set of worker and firm-level controls consisting of (i) the same socio-demographic characteristics and control for academic performance as in specification 1; and (ii) a very rich set of firm and job characteristics (whether part-time contract, whether temporary contract, 2-digit sector fixed effects, 2-digit occupation fixed effects, birthplace-workplace distance, firm size, value added per worker, average workforce education, share of female workforce, firm gender wage gap at the mean and the 90th percentile). To assess the contribution of each job and firm characteristic, we perform the following set of field-specific Oaxaca decompositions:

[26] We select fields for which we have at least 1000 observations, implying that we end up working with 17 university majors.

$$w_{it}^{men,j} - w_{it}^{women,j} = \underbrace{\Sigma_\omega \hat{\alpha}_\omega^{men,j}(\overline{Z}_\omega^{men,j} - \overline{Z}_\omega^{women,j})}_{\text{Explained } (E_\omega)} + \underbrace{(\hat{\alpha}_0^{men,j} - \hat{\alpha}_0^{women,j}) + \Sigma_\omega \overline{Z}_\omega^{women,j}(\hat{\alpha}_\omega^{men,j} - \hat{\alpha}_\omega^{women,j})}_{\text{Unexplained } (U_\omega)} \tag{4}$$

We summarize the results of the $\boldsymbol{J}$ decompositions by computing, for any given firm or job characteristic ω, the *average* contribution to the *average* within-field gender gap, weighting each field by their number of graduates.

$$E_\omega = \Sigma_j \frac{N_j}{N} E_\omega^j = \Sigma_j \frac{N_j}{N} \hat{\alpha}_\omega^{men,j}(\overline{Z}_\omega^{men,j} - \overline{Z}_\omega^{women,j}) \tag{5}$$

where E_ω is the average contribution explained by characteristics ω to the within-field gender wage gap, E_ω^j is the contribution explained by characteristics ω to the gender wage gap in fields j, N_j is the number of students graduated in field j, N is the overall number of university students, $\hat{\alpha}_\omega^{boys,j}$ is the coefficient estimated for boys for characteristics ω in equation 3 and $\overline{Z}_\omega^{boys,j}$ and $\overline{Z}_\omega^{women,j}$ are the averages of characteristic ω computed for men and women graduated from field j, respectively.

We also decompose the gender gap at the right and left tails of the wage distribution. In particular, we estimate the same regressions as described before also using as dependent variables the probability of being a top earner (top 10% of daily wage distribution) or a bottom earner (bottom 10%).

5.2 Results

*5.1. Step 1: The role of educational choices for the **aggregate** gender wage gap* – Figure 5 displays the decomposition of the gender gap in (log) daily wages five years after graduation, based on equation 2. Field of study choices play by far the most significant role: they explain alone almost 60% of the average gap. Interestingly, while gender differences in university majors are largest among higher-ability students (in terms of disparities of potential earnings; see Section 4), their role is more important for disparities at the bottom than at the top of the wage distribution. This suggests that other - more difficult to observe - factors matter more for gender gaps in the probability of reaching the top of the wage distribution.

Differences in the choice of which university to attend and whether to study out of the region of birth play a very limited role that - if anything - is more visible when focusing on top earners. Disparities in academic performance would imply, instead, a smaller gap than that observed (especially at the top), due to women's advantage in this dimension.

All in all, the full set of controls that capture socio-demographic characteristics and education choices and outcomes accounts for 60% of the gap in average (log) daily wages; the unexplained component is smaller at the bottom and larger at the top of the wage distribution.

5.2. Step 2: The role of firms and jobs for the ***within-field*** *gender wage gap* – As shown in Section 4, young women earn less than men even within each field of study, although gaps are way smaller than in the aggregate due to the large role of cross-field disparities. Figure 6 presents the results of the Oaxaca-Blinder decomposition of the within-field gender pay gaps five years after graduation, as described in equation 5. Several interesting findings emerge.

First, our very rich set of firm and job controls explains, on average across all fields, only slightly more than 40% of the within-field gender gap in daily wages. The characteristic that plays the largest role is the higher prevalence of part-time contracts among women. Other job and firm characteristics, like occupation, sector, and firm-level gender wage gaps matter very little. Notably, also firm productivity and workplace-birthplace distance, which are found to be very relevant for aggregate gaps in analysis without information on fields of study (Cardoso et al., 2016; Bruns, 2019; Le Barbanchon et al., 2021; Casarico and Lattanzio, 2024; Caldwell and Danieli, 2024), matter very little for the within-major gender wage gap.

Second, firm and job characteristics, and in particular the probability of working part-time, matter more in the left than in the right tail of the distribution: the unexplained component accounts on average for approximately 40% of the gender gap among bottom earners, while it rises to 80% among top earners. This residual (unexplained) gender gap is probably related to factors that cannot be easily observed like, for instance, negotiation skills, availability to work very long hours and overtime, or the likelihood to be asked to perform non-remunerative tasks, or employer discrimination against women (see Bowles et al., 2007; Goldin, 2014; Babcock et al., 2017; Gonzalez-Rozada and Yeyati, 2018).

Third, there is some heterogeneity in the role of firm and job characteristics across groups of majors. In line with the evidence in Section 4 that gaps in observable measures of firm and job quality are smaller or even reverted in narrow STEM fields, the unexplained component of gender gaps in daily wages is largest for STEM majors (Figure 7). In particular, in narrow

STEM fields, it is much smaller the contribution of the propensity to work part-time and of firm productivity.

Fourth, Figure 8 shows a negative and significant relationship between the share of women graduating in a given major and the unexplained component of the within-field average gender wage gap. This evidence suggests that women tend to choose majors that entail lower uncertainty about the reasons behind gender wage gaps.

6 Discussion: the mediating role of fields of study for the gender wage gap

Our decomposition exercises show that a form of pre-market specialization - the choice of the university major - explains more than half of the aggregate gender gap. Because different majors are associated with different labor markets, gender differences in firms and jobs characteristics are much milder within field of study than in the aggregate. Also, we show that on average they explain only slightly more than 40% of the gap within field.

In order to assess the relative explanatory power of educational and labor market controls, in Table 2 we resort to a standard regression approach that allows adding controls at a level of granularity higher than that feasible in the Oaxaca-Blinder decomposition. Column (1) reports the raw (unconditional) gap in log daily wages five years after graduation for the considered sample, which is equal to -0.299. Column (2) displays the residual gap after controlling for all variables that capture education choices and academic achievement: it shrinks by more than 50%, in line with what emerges from Section 5.2. Column (3) additionally controls for job attributes and employer characteristics. The gap reduces to -0.082: hence, firm and job attributes account for only 20% of the gap, once we control for educational choices. Finally, column (4) displays the residual gender gap in a regression that controls for job and firm characteristics alone (together with socio-demographic controls), as if we did not have any information on education choices. The resulting residual gender gap is 0.087, thus suggesting that in a standard setting where we do not observe workers' fields of study, we would conclude that firm and job attributes account for 70% of the gender gap in daily wages.

Figure 9 makes this point visually, not only for the average wage gap but also for the gaps at the top and the bottom of the distribution: the second and third bars show the role of jobs and firms when including or excluding education controls, respectively. The figures clearly show that field of study choices largely mediate disparities in firm and job characteristics along the entire distribution of wage gaps.

Our results can be compared to the most recent literature that aims to explain the sources of gender gaps through AKM models. First, the finding that job and employer characteristics explain about 20% of the raw average gender gap is roughly in line with the lower end of existing estimates of the contribution of firm premiums, which in advanced countries ranges between one-fifth and one-third (Cardoso et al., 2016; Sorkin, 2017; Bruns, 2019; Li et al., 2023; Casarico and Lattanzio, 2024). Given that we are looking at the very first years of the career, when sorting across employers as a result of job-to-job mobility is still at the beginning, it is reasonable to find that firm characteristics play a smaller role. Second, we somehow open the black box of individual fixed effects in AKM models. As women have overtaken men in terms of education levels, our results suggest that for younger cohorts individual fixed effects mainly capture differences in the type of education, specifically fields of study.

7 Robustness

In this Section, we show that our results are robust across different samples, outcomes, and time horizons.

Panel (a) of Appendix Figure A.4 shows the role of field of study choices for the aggregate gender gap when looking at all workers (i.e., including also public employees and self-employed individuals). The outcome is the gap in annual labor earnings since we do not have information on days worked for the self-employed. Also in this larger sample, the share of the early career gender gap explained by majors is very sizable (approximately 50%), and it is larger at the bottom than at the top of the distribution. In panel (b), we focus again on private-sector employees but decompose the gender gap in daily wages one year after graduation. The results are very similar to those obtained for gaps five years after graduation. In panel (c), we look at two other outcomes: full-time-equivalent (FTE) wages and daily wages for the sample of full-time employees only. Again, the findings align with the main ones.

Appendix Figure A.5 presents alternative decompositions of within-field wage gaps. The sample remains that of private-sector employees because we have very limited information about firm and job characteristics for public-sector employees and the self-employed. Panel (a) shows that results are very similar when looking at the gender gap in daily wages 1 year after graduation. Also at this shorter horizon, the unexplained component is larger at the top of the distribution. The role of differences in the probability of holding a part-time contract is slightly smaller at the beginning of the career, suggesting that the gendered take-up in part-time jobs is a phenomenon that builds up over time. Panel (b) reports the decomposition of the gender gap in FTE wages and in daily wages of full-time employees only: the role of differences in the probability of holding a part-time job becomes very small and even negative. Consequently, the share of the within-field gender wage gap explained by the very rich set of observable job and firm characteristics available to us becomes even smaller (about 20%), leaving the vast majority of the within-field gap unexplained.

8 Conclusions

This paper documents four important facts on early career gender gaps. First, they are substantial already right after men and women complete their education, even among very highly educated individuals. Second, while women outperform boys academically they select very different fields of study: they are more likely to choose majors with lower expected payoffs, especially if at the top of the ability distribution. Third, by means of an Oaxaca-Blinder decomposition analysis, we show that the choice of the field of study is the most important explanatory factor of the early career gender gap: it accounts for almost 60% of the gap in (log) daily wages five years after graduation from university, and its explanatory power is larger at the bottom than at the top of the wage distribution. This happens because majors are found to strongly mediate access to high-quality firms and jobs. Fourth, even if smaller than aggregate gaps, there are also significant within-field gender differences: holding the major fixed, women are more likely from the start to have lower-paying contracts (part-time, in lower-skilled occupations) and work for firms which are smaller, with a higher share of female workers, closer to home, and less productive. However, these differences in firm and job characteristics within majors explain only 40% of the average within-field gender pay gap among university graduates. When studying the size of the unexplained component within each field, we find that women are more likely to graduate from majors where it is smaller, i.e. where there is lower uncertainty about the sources of gender disparities.

Overall, these results indicate that field of study choices are the most important factor in explaining the gender pay gap at the beginning of university graduates' careers. The observed gender differences in men's and women's sorting across firms and jobs are mediated by pre-market specialization choices made while still in education.

This finding bears important policy implications. Interventions at the time when women and men choose their university major could be especially effective in mitigating early career pay gaps. For policymakers, then, it becomes crucial to understand better the mechanisms driving pre-market education investments: for instance, do women choose certain majors in anticipation of future family demands? Based on individual preferences? Conditioned by stereotypes or social norms?

There is a consensus that gender differences in fields of study largely reflect heterogeneity in preferences (Zafar, 2013; Ceci et al., 2014; Wiswall and Zafar, 2014) and can affect education choices as early as during adolescence, as in many countries boys and girls can choose which subjects to take in high school. However, a growing literature argues that such preferences are not innate but partly influenced by contextual factors such as social norms and gender stereotypes (e.g. Guiso et al., 2008, Pope and Sydnor, 2010, Nollenberger et al., 2016, Kahn and Ginther, 2018, Carlana, 2019, Bertrand, 2020, Brenøe, 2022, Miserocchi, 2024). Given how consequential choosing a university major is, ensuring that such factors do not restrict the choice set of girls is particularly relevant. Little is known, however, about which policies can effectively challenge these stereotypes and the burden of social norms Early exposure to STEM content, including interaction with female role models, has for example been found successful in a number of settings (e.g.,Carlana and Fort, 2022; Breda et al., 2023) but further research is needed on this topic.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Altonji, J., P. Arcidiancono, and A. Maurel (2016). *The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects*, Volume 5. Handbook of the Economics of Education.

Altonji, J., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics 4*(1), 185–223.

Anelli, M. and G. Peri (2015). Gender gap in italy: The role of college majors. *Unexplored Dimensions of Discrimination*, 79.

Arellano-Bover, J., N. Bianchi, S. Lattanzio, and M. Paradisi (2023). Gender convergence in the labor market: are women climbing up or men falling down? . Mimeo, Bank of Italy.

Babcock, L., M. P. Recalde, L. Vesterlund, and L. Weingart (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review 107*(3), 714–747.

Bertrand, M. (2020). Gender in the Twenty-First Century. *AEA Papers and Proceedings 110*, 1–24.

Black, D. A., A. M. Haviland, S. G. Sanders, and L. J. Taylor (2008). Gender wage disparities among the highly educated. *Journal of human resources 43*(3), 630–659.

Blau, F. D. and L. M. Kahn (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature 55*(3), 789–865.

Bowles, H. R., L. Babcock, and L. Lai (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes 103*(1), 84–103.

Breda, T., J. Grenet, M. Monnet, and C. Van Effenterre (2023). How effective are female role models in steering girls towards stem? evidence from french high schools. *The Economic Journal 133*(653), 1773–1809.

Brenøe, A. A. (2022). Brothers Increase Women's Gender Conformity. *Journal of Population Economics 35*(4), 1859–1896.

Brown, C. and M. Corcoran (1997). Sex-based differences in school content and the male-female wage gap. *Journal of Labor Economics 15*(3), 431–465.

Bruns, B. (2019, April). Changes in workplace heterogeneity and how they widen the gender wage gap. *American Economic Journal: Applied Economics 11*(2), 74–113.

Caldwell, S. and O. Danieli (2024). Outside Options in the Labour Market. *The Review of Economic Studies*, rdae006.

Campbell, S., L. Macmillan, R. Murphy, and G. Wyness (2022). Matching in the Dark? Inequalities in Student to Degree Match. *Journal of Labor Economics 40*(4), 807–850.

Cardoso, A. R., P. Guimaraes, and P. Portugal (2016). What drives the gender wage gap? a look at the role of firm and job-title heterogeneity. *Oxford Economic Papers 68*(2), 506–524.

Carlana, M. (2019). Implicit Stereotypes: Evidence from Teachers' Gender Bias. *The Quarterly Journal of Economics 134*(3), 1163–1224.

Carlana, M. and M. Fort (2022). Hacking gender stereotypes: Girls' participation in coding clubs. In *AEA Papers and Proceedings*, Volume 112, pp. 583–587. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Casarico, A. and S. Lattanzio (2024). What firms do: Gender inequality in linked employer-employee data. *Journal of Labor Economics 42*(2), 325–355.

Ceci, S., D. Ginther, S. Kahn, and W. Williams (2014). Women in Academic Science: A Changing Landscape. *Education Economics 15*(3), 75–141.

Delfino, A. (2024). Breaking gender barriers: Experimental evidence on men in pink-collar jobs.

Eide, E. (1994). College major choice and changes in the gender wage gap. *Contemporary Economic Policy 12*(2), 55–64.

Francesconi, M. and M. Parey (2018). Early gender gaps among university graduates. *European economic review 109*, 63–82.

Goldin, C. (2014, April). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Gonzalez-Rozada, M. and E. L. Yeyati (2018). Do women ask for lower salaries? the supply side of the gender pay gap. Department of Economics Working Papers 2018-02, Universidad Torcuato Di Tella.

Graham, J. W. and S. A. Smith (2005). Gender differences in employment and earnings in science and engineering in the us. *Economics of education review 24*(3), 341–354.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales (2008, 01). Culture, Gender, and Math. *Science 320*.

Huneeus, F., C. Miller, C. A. Neilson, and S. Zimmerman (2021). Firm Sorting, College Major, and the Gender Earnings Gap. Working Paper Series 648, Industrial Relations Section.

Istat (2023). Migrazioni Interne e Internazionali della Popolazione Residente. Anno 2021. Technical report, Istat.

Joy, L. (2003). Salaries of recent male and female college graduates: Educational and labor market effects. *ILR Review 56*(4), 606–621.

Kahn, S. and D. Ginther (2018). Women and Science, Technology, Engineering, and Mathematics (STEM). In L. M. A. Susan L. Averett and S. D. Hoffman (Eds.), *The Oxford Handbook of Women and the Economy.*

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics 131*(3), 1057–1111.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimuller (2019). Child penalties across countries: Evidence and explanations. In *AEA Papers and Proceedings*, Volume 109, pp. 122–26.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2020, November). Do family policies reduce gender inequality? evidence from 60 years of policy experimentation. Working Paper 28082, National Bureau of Economic Research.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics 136*(1), 381–426.

Li, J., B. Dostie, and G. Simard-Duplain (2023). Firm pay policies and the gender earnings gap: the mediating role of marital and family status. *ILR Review 76*(1), 160–188.

Lordan, G. and J.-S. Pischke (2022). Does rosie like riveting? male and female occupational choices. *Economica 89*(353), 110–130.

Loury, L. D. (1997). The gender earnings gap among college-educated workers. *ILR Review 50*(4), 580–593.

Machin, S. and P. A. Puhani (2003). Subject of degree and the gender wage differential: evidence from the uk and germany. *Economics Letters 79*(3), 393–400.

McDonald, J. A. and R. J. Thornton (2007). Do new male and female college graduates receive unequal pay? *Journal of Human Resources 42*(1), 32–48.

Miserocchi, F. (2024). Discrimination through biased memory. *Working Paper*.

Nollenberger, N., N. Rodríguez-Planas, and A. Sevilla (2016, May). The Math Gender Gap: The Role of Culture. *American Economic Review 106*(5), 257–61.

Olivetti, C. and B. Petrongolo (2008). Unequal pay or unequal employment? A cross-country analysis of gender gap. *Journal of Labor Economics* (26), 621–65.

Piazzalunga, D. (2018). The gender wage gap among college graduates in italy. *Italian Economic Journal 4*(1), 33–90.

Pope, D. G. and J. R. Sydnor (2010, June). Geographic Variation in the Gender Differences in Test Scores. *Journal of Economic Perspectives 24*(2), 95–108.

Sloane, C. M., E. G. Hurst, and D. A. Black (2021, November). College majors, occupations, and the gender wage gap. *Journal of Economic Perspectives 35*(4), 223–48.

Sorkin, I. (2017). The role of firms in gender earnings inequality: Evidence from the united states. *American Economic Review 107*(5), 384–387.

Weinberger, C. J. (1998). Race and gender wage gaps in the market for recent college graduates. *Industrial Relations: A Journal of Economy and Society 37*(1), 67–84.

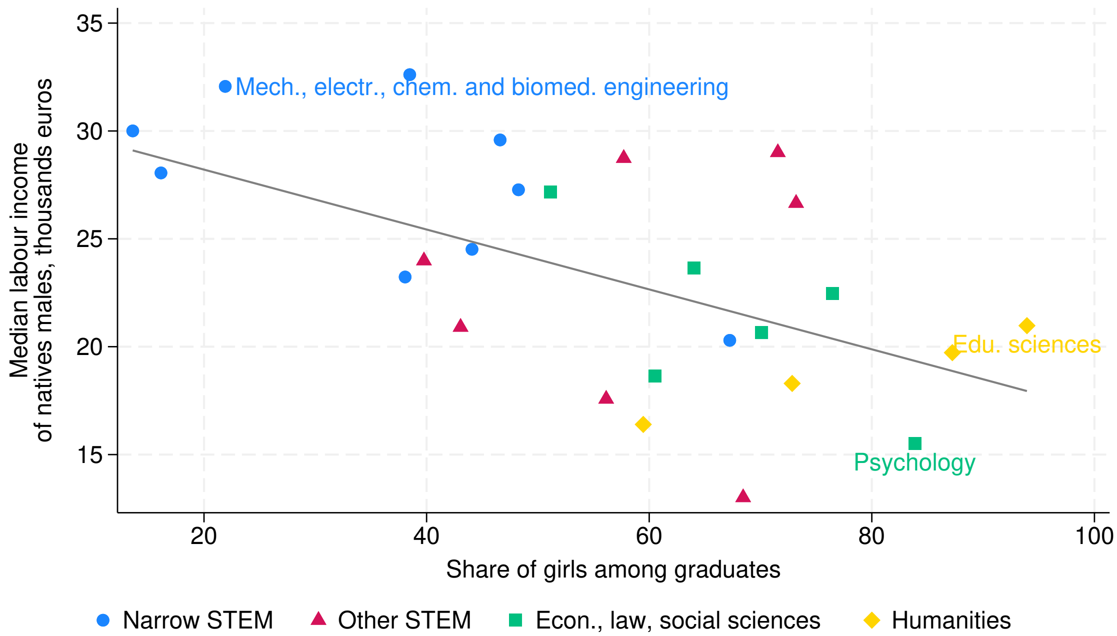
Weinberger, C. J. (2011). In search of the glass ceiling: Gender and earnings growth among us college graduates in the 1990s. *ILR Review 64*(5), 949–980.

Wiswall, M. and B. Zafar (2014, 12). Determinants of College Major Choice: Identification using an Information Experiment. *The Review of Economic Studies 82*(2), 791–824.

Zafar, B. (2013). College Major Choice and the Gender Gap. *Journal of Human Resources 48*(3), 545–595.

Figures

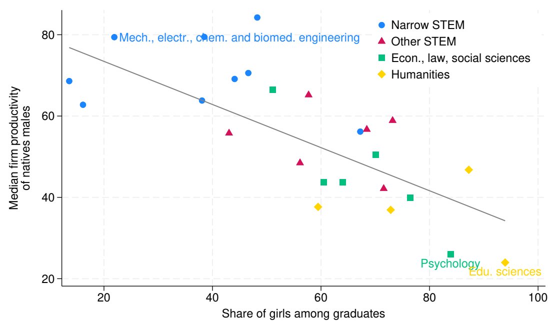
Figure 1: Financial returns for the median male native graduate 5 years after graduation vs. share of females graduating from a major



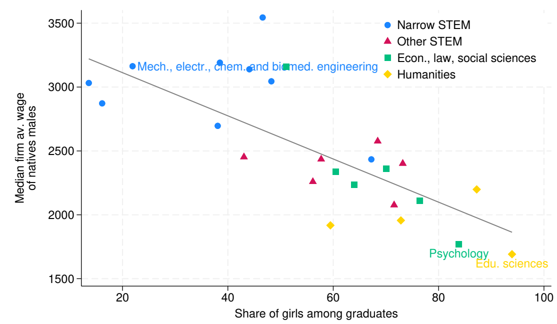
Notes: For each major, the scatterplot displays: on the x-axis, the share of females among 2nd cycle and single-cycle graduates in 2011-18; on the y-axis, the financial return of the major, as measured by the median labour income of its male, native graduates in 2011-14 who 5 years after graduation work and no longer study. Narrow STEM fields include: engineering, chemistry, mathematics, physics, biology, and ICT. Other STEM fields include: agriculture, veterinary, architecture, pharmacy, medicine, dentistry, and nursing. Econ., law, and social sciences include: law, psychology, economics and management, political science, sociology, and communication. Humanities include: literature, arts, philosophy, languages, and education.

Figure 2: Labor markets male native graduates have access to 5 years after graduation vs. share of females in a major

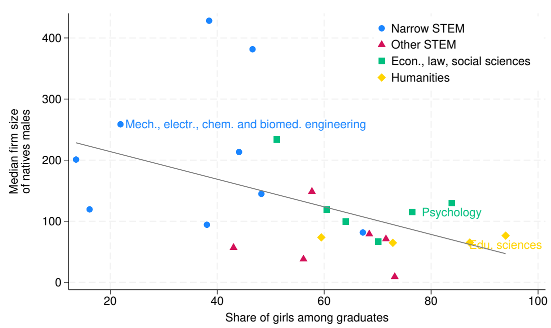
(a) Firm productivity



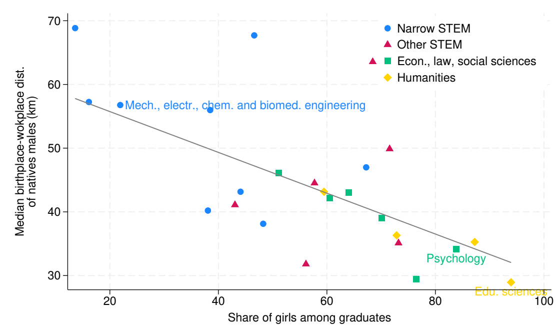
(b) Firm average wage



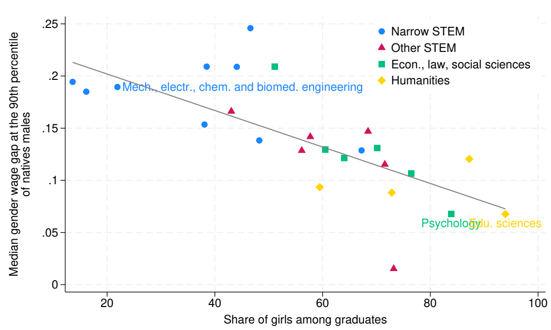
(c) Firm size



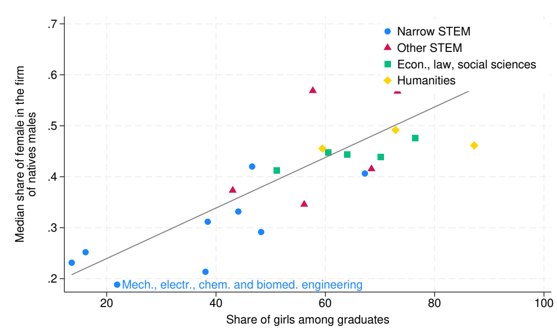
(d) Birthplace-workplace distance



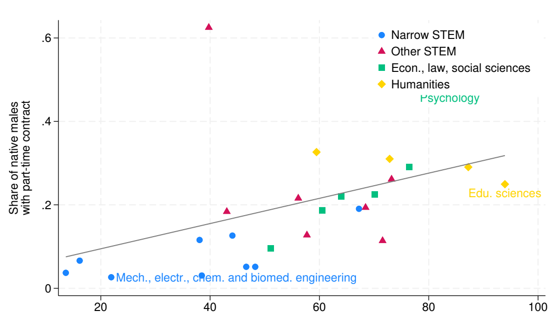
(e) Within-firm gender wage gap (90th pct)



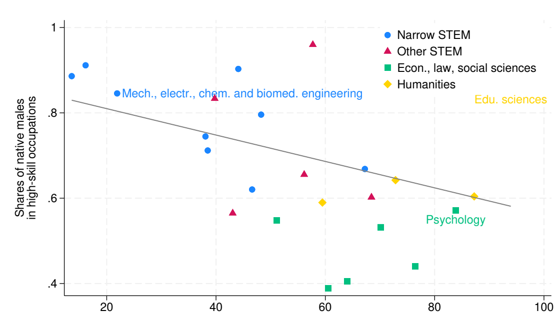
(f) % of women in the firm



(g) % with a part-time contract

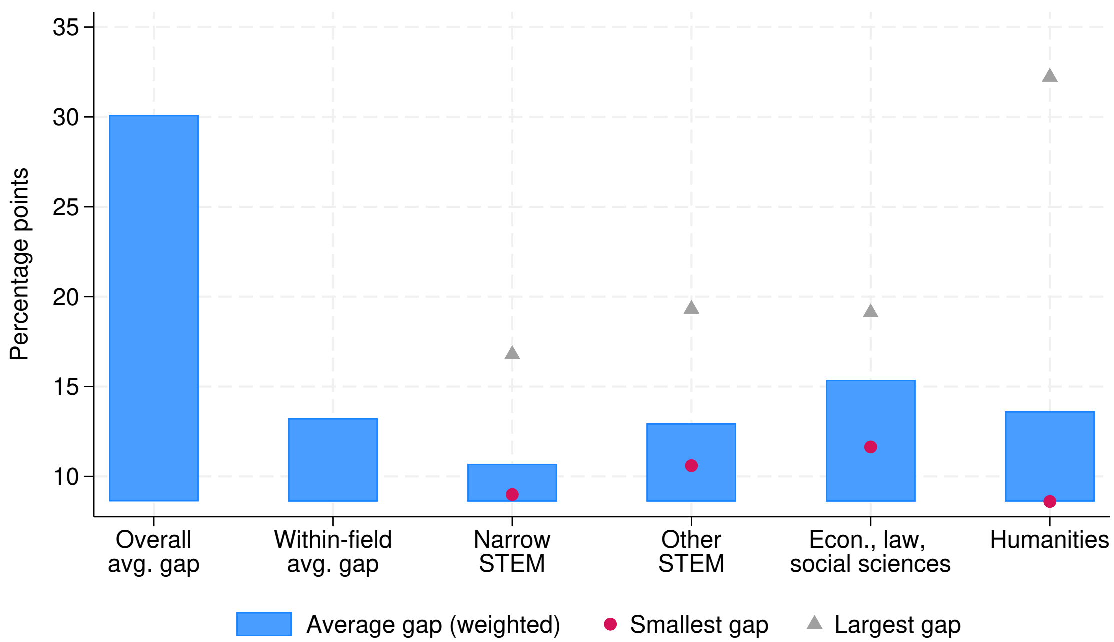


(h) % in high skilled occupations



Notes: For each major, the scatterplots display: on the x-axis, the share of females among 2nd cycle and single-cycle graduates in 2011-18; on the y-axis, the characteristic of the firm and the job where the median male native graduate who 5 years after graduation works and no longer studies is employed.

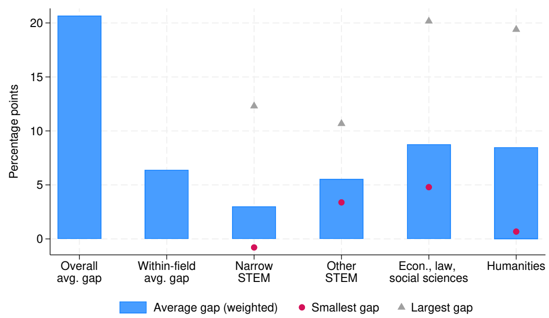
Figure 3: Unconditional gender gap in average daily wages, 5 years after graduation, overall and by group of fields



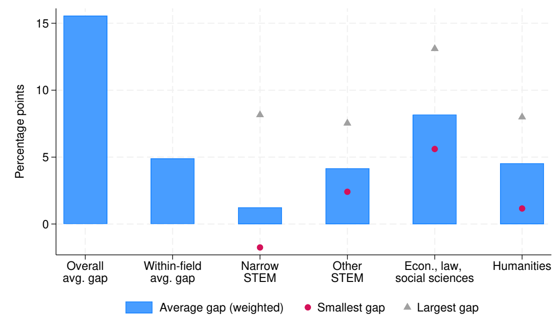
Notes: cohorts graduating between 2011-2014 with a 2nd level or single-cycle degree who are working as private sector employees and not studying 5 years after graduation. The sample includes all individuals for whom we observe all job- and firm- level variables included across *all* Oaxaca-Blinder decomposition regressions (equations 1 and 3). Fields are grouped as in Figure 1. Gaps are expressed in percentage variation relative to boys' values ($100 \times \frac{w_m - w_f}{w_m}$).

Figure 4: Average within-field gender gap in average daily wages, 1 years after graduation, overall and by group of fields

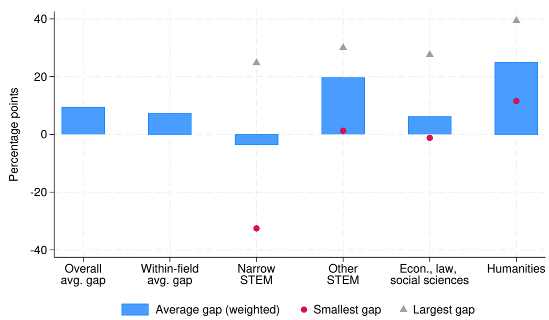
(a) Firm productivity



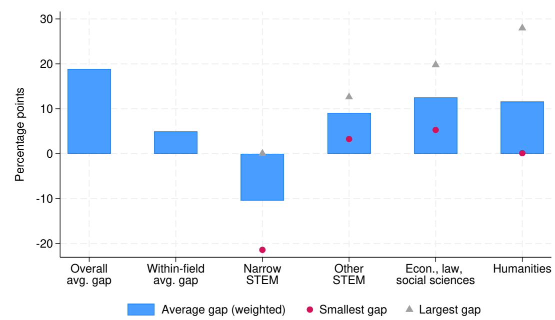
(b) Firm average wage



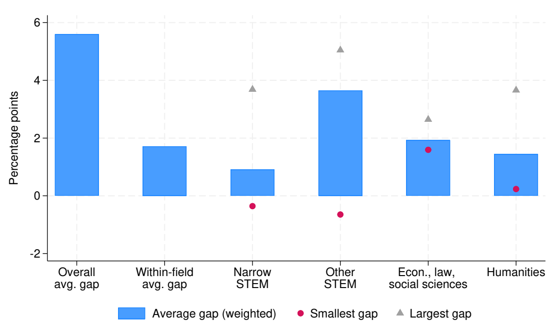
(c) Firm size



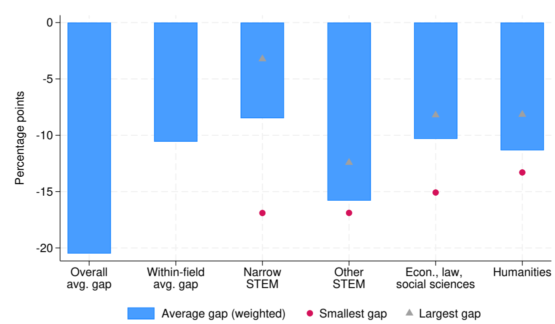
(d) Birthplace-workplace distance



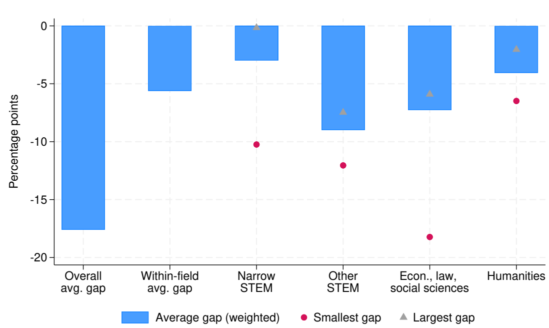
(e) Within firm gender wage gap (90th pct)



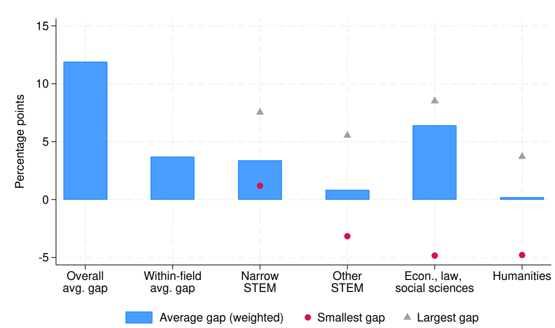
(f) % of women in the firm



(g) % with a part-time contract

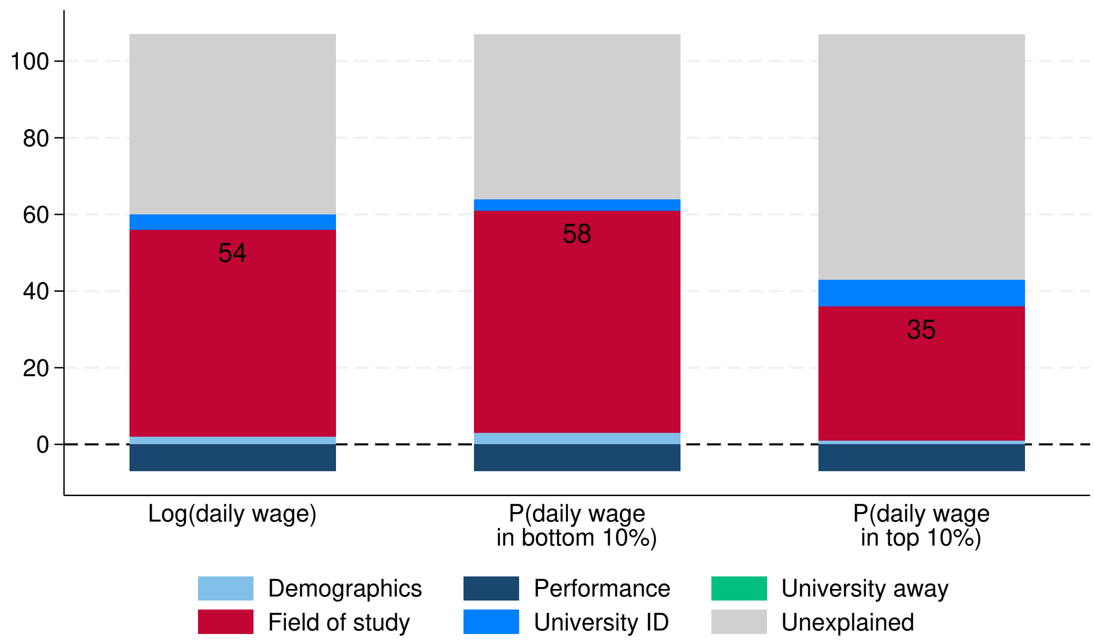


(h) % in high skilled occupations



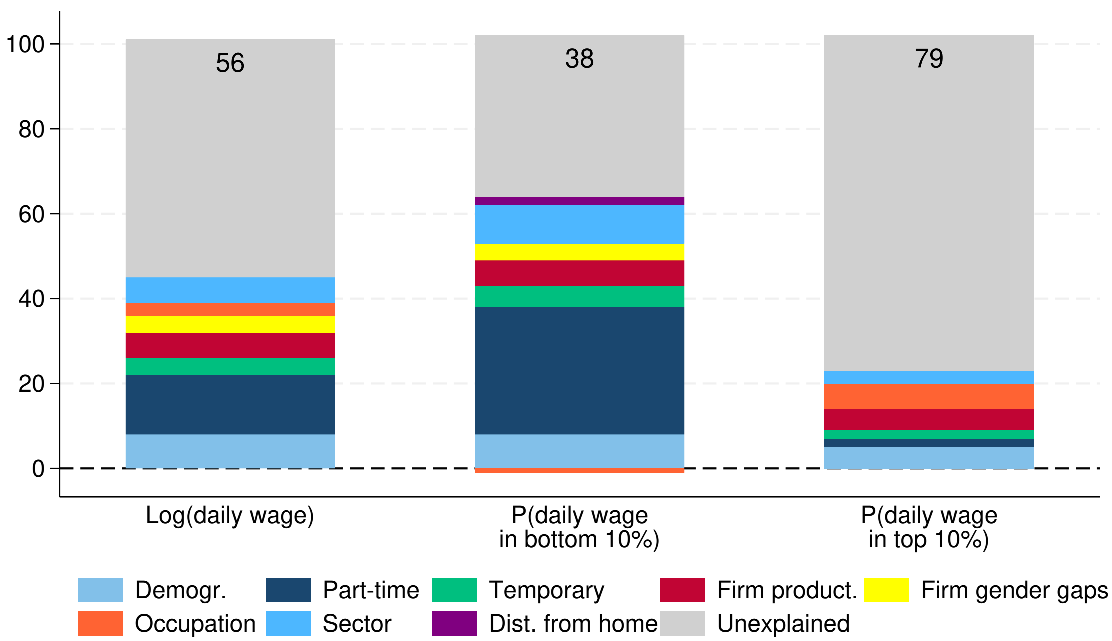
Notes: cohorts graduating between 2011-2018 with a 2nd level or single-cycle degree who are working as private sector employees and not studying 1 year after graduation. Fields are grouped as in Figure 1. Panels a to d: gaps are expressed in percentage variation relative to boys' values ($100 \times \frac{X_m - X_f}{X_m}$); panels e to h: gaps are expressed in absolute difference ($X_m - X_f$).

Figure 5: The role of education choices for the unconditional gender gap - Oaxaca decomposition



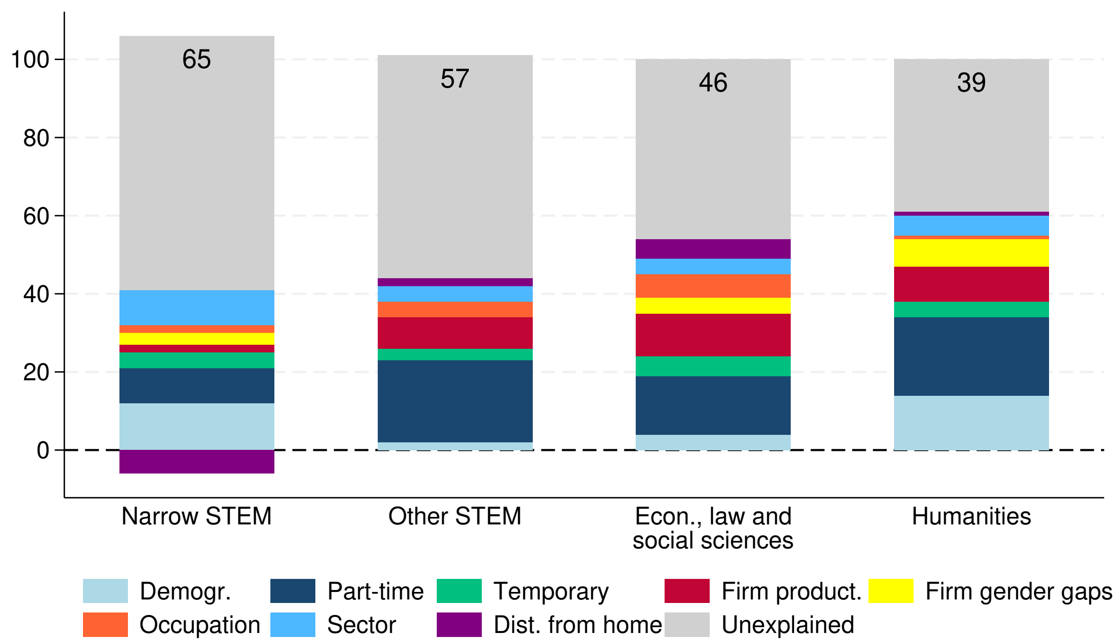
Notes: cohorts graduating between 2011-14 in a 2nd level and single-cycle degree who are working as private sector employees and not studying 5 years after graduation. The sample is further restricted to individuals with non-missing values for all the controls contained in the vector X in equations 1 and 3. Overall differences normalized to 100. "Demographics" includes: marital status, a set of fixed effects for ventiles for household income, region of birth, and cohort fixed effects. "Performance" includes the final grade at university and the age at graduation. "University away" is a dummy that takes the value 1 if the university is located in a region different from that of birth. "Field of study" includes a set of dummies for 27 majors. "University ID" includes a set of dummies for all the various universities. "Unexplained" represents the part of the gap not explained by gender disparities in the characteristics and choices listed above. Number of observations = 129810.

Figure 6: Average contribution of jobs' and firms' characteristics for the within-field gender gap 5 years after graduation - Oaxaca decomposition



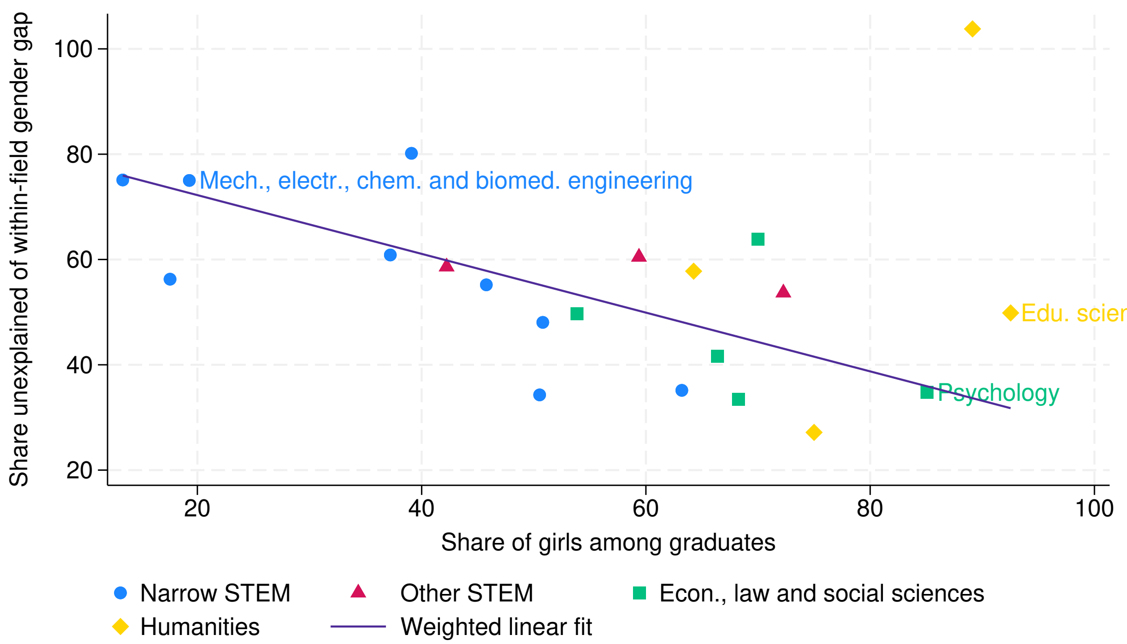
Notes: cohorts graduating between 2011-14 in a 2nd level and single-cycle degree who are working as private sector employees and not studying 5 years after graduation. The sample is further restricted to individuals with non-missing values for all the controls contained in the vector X in equations 1 and 3. Overall differences normalized to 100. "Demogr." includes: marital status, dummies for ventiles for household income, region of birth, and cohort dummies. "Part-time" ("temporary") indicates if the worker was employed part-time (with a fixed-term contract) in the main job in the year. "Firm product." includes: firm size, age, value added per worker, and average education level of the workforce. "Firm gender gaps" includes: the share of women in the firm, and the firm-level wage gap at the average and the 90th percentile. "Occupation" consists of 2-digit occupation fixed effects. "Sector" indicates 2-digit sector fixed effects. Distance from home is the distance between the birth municipality and the work municipality, aggregated in 10 bins. Number of observations = 127270.

Figure 7: Average contribution to the within-field average gender gap in daily wage 5 years after graduation - Oaxaca decomposition



Notes: cohorts graduating between 2011-2014 in a 2nd level and single cycle degree who are working as private sector employees and not studying 5 years after graduation. The sample is further restricted to individuals with non-missing values for all the controls contained in the vector X in equations 1 and 3. The overall difference is normalized to 100. Fields are grouped as in Figure 1. Among humanities, we exclude literature because it is an outlier. Number of observations = 121043.

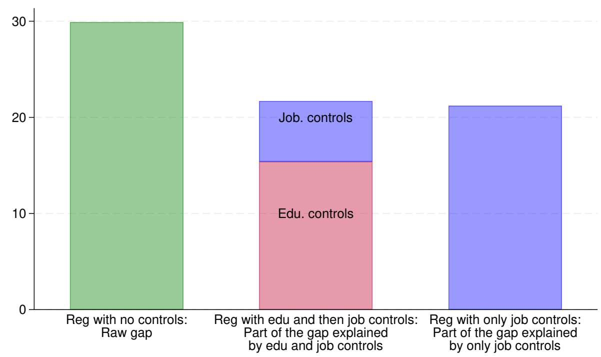
Figure 8: The unexplained component by field of study



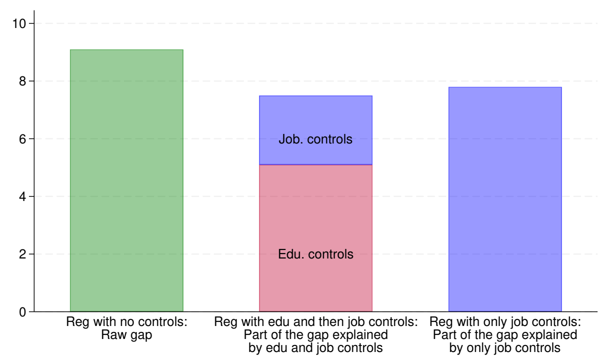
Notes: cohorts graduating between 2011-2014 in a 2nd level and single cycle degree who are working as private sector employees and not studying 5 years after graduation. The sample is further restricted to individuals with non-missing values for all the controls contained in the vector X in equations 1 and 3. Fields are grouped as in Figure 1. The fitted line excludes literature, as it is an outlier.

Figure 9: Gender pay gap 5 years after graduation, contribution of jobs conditional on education controls and of jobs alone

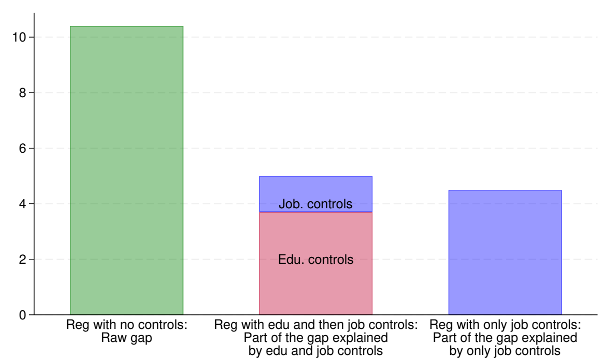
(a) Log(daily wages)



(b) Pr(daily wage in the bottom 10%)



(c) Pr(daily wage in the top 10%)



Notes: cohorts graduating between 2011-2014 in a 2nd level and single cycle degree who are working as private sector employees and not studying 5 years after graduation. The bars are based on the coefficients estimated as in table 2. The gap in daily wages is computed as (M-F)/M, the gap in the probability of lying at the bottom of the distribution as (F-M), that in the probability of lying at the top as (M-F).

Tables

Table 1: Job characteristics 5 years after graduation at university

	Boys		Girls			
	mean	s.d.	mean	s.d.	Difference	t-stat
	(1)	(2)	(3)	(4)	(5)	(6)
Daily wage	79.053	79.654	58.917	31.03	20.136***	(66.207)
Type of contract						
Temporary	0.188	0.391	0.288	0.453	-0.100***	(-50.605)
Part-time	0.128	0.334	0.314	0.464	-0.186***	(-100.125)
Occupation						
Managers	0.005	0.073	0.003	0.054	0.002***	(7.876)
Professionals	0.34	0.474	0.294	0.456	0.046***	(20.768)
Technicians	0.307	0.461	0.238	0.426	0.070***	(32.897)
Clerical support workers	0.239	0.426	0.327	0.469	-0.089***	(-42.267)
Services and sales workers	0.068	0.251	0.114	0.318	-0.046***	(-34.938)
Skilled agric. and craft workers	0.012	0.108	0.004	0.063	0.008***	(17.817)
Plant and machine operators	0.01	0.101	0.004	0.064	0.006***	(14.958)
Elementary occupations	0.019	0.137	0.016	0.125	0.003***	(5.354)
Sector						
Agriculture and industry	0.328	0.47	0.174	0.379	0.154***	(74.773)
Private market services	0.601	0.49	0.623	0.485	-0.022***	(-9.469)
Private non-market services	0.071	0.257	0.203	0.402	-0.132***	(-85.950)
Firm-level characteristics						
Distance birth-workplace (km)	247.915	380.582	209.858	359.387	38.057***	(21.287)
% of women	0.356	0.212	0.568	0.257	-0.212***	(-183.744)
Avg years of education	14.453	2.066	14.493	1.993	-0.040***	(-3.961)
Avg monthly wage	2922.959	1029.88	2495.642	992.37	427.317***	(84.825)
Age	22.82	18.611	20.039	17.15	2.780***	(31.267)
Value added	70.123	38.043	56.441	37.787	13.682***	(70.703)
Size	4647.893	18580.85	4915.123	19551.33	-267.230**	(-2.831)
Gender wage gap (mean)	0.116	0.176	0.1	0.223	0.016***	(16.033)
Gender wage gap (90th pct)	0.141	0.268	0.091	0.356	0.049***	(31.170)
% white collars	0.574	0.281	0.6	0.311	-0.026***	(-17.673)
Observations	75917		109231			185148

Notes: cohorts graduating between 2011-2014 in a 2nd level and single cycle degree who are working as private sector employees and not studying 5 years after graduation.

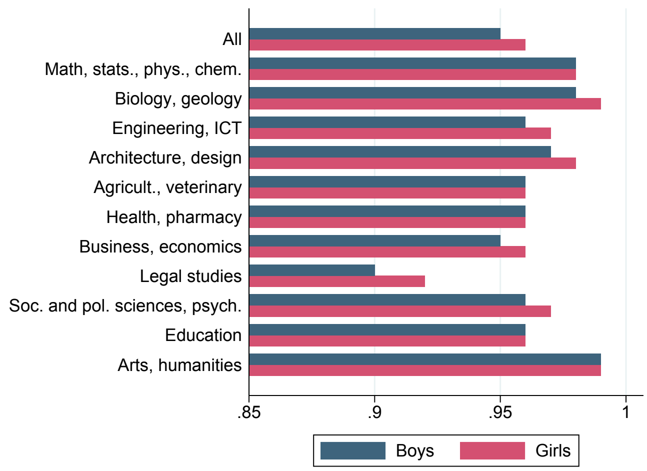
Table 2: Gender gap (girls-boys) 5 years after graduation

	Raw (1)	Only edu contr. (2)	All contr. (3)	Only job contr. (4)
		Average (log) daily wage		
Female	-0.299*** (0.003)	-0.145*** (0.003)	-0.082*** (0.002)	-0.087*** (0.002)
% of column 1 gap explained:		51.5	72.6	70.9
N	138444	138440	138376	138380

Notes: cohorts graduating between 2011-2014 in a 2nd level and single cycle degree who are working as private sector employees and not studying 5 years after graduation. Columns (1) displays the raw gap; column (2) controls for demographics, background, as well as education choices, and academic achievement. Controls consist of: region of birth fixed effects, marital status, and socio-economic background as captured by ventiles of parents' total income; age and final grade at graduation, as well as their squares; university major fixed effects; university id fixed effects. Column (3) controls also for job and employer attributes, which consist of: dummies for whether the contract is part-time or temporary, respectively; age, size, and productivity (i.e. value added per worker) of the employer, as well as their squares; share of workforce in the firm who is female and average workforce education, and their squares; distance between the birth municipality and the work municipality, aggregated in 10 bins; 5-digit occupation fixed effects and 6-digit sector fixed effects.

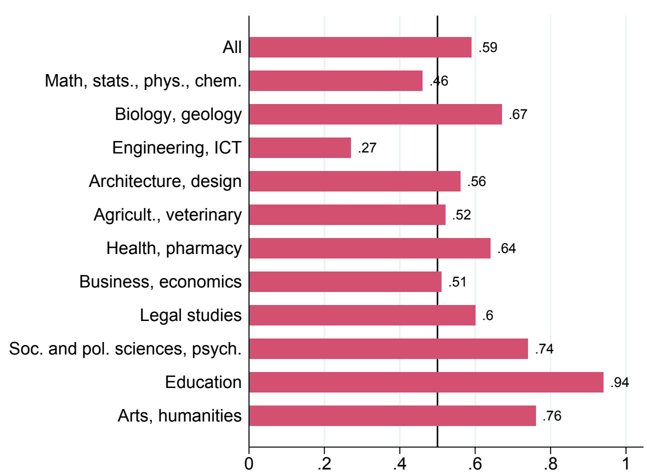
A Appendix: additional figures and tables

Figure A.1: Average final grade of 2nd cycle and single-cycle university graduates, by gender and group of majors



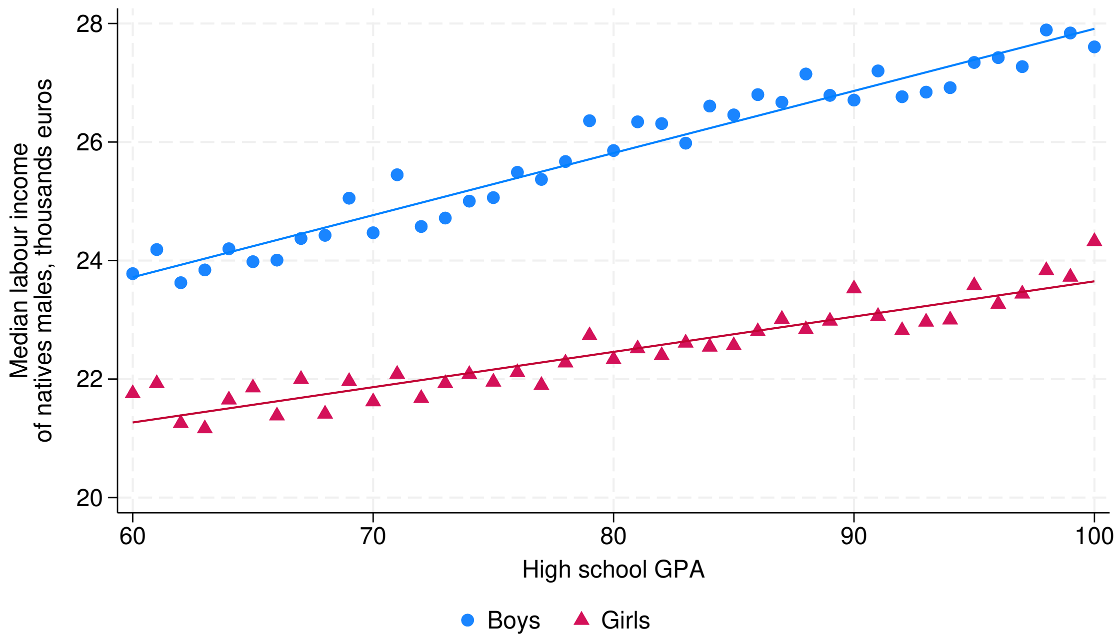
Notes: 2011-18 cohorts of graduates from 2nd cycle and single-cycle degrees. Final grades range from 60 to 110 *cum Laude*.

Figure A.2: Share of females among graduates from 2nd cycle and single-cycle degrees, by group of majors



Notes: 2011-18 cohorts of graduates from 2nd cycle and single-cycle degrees.

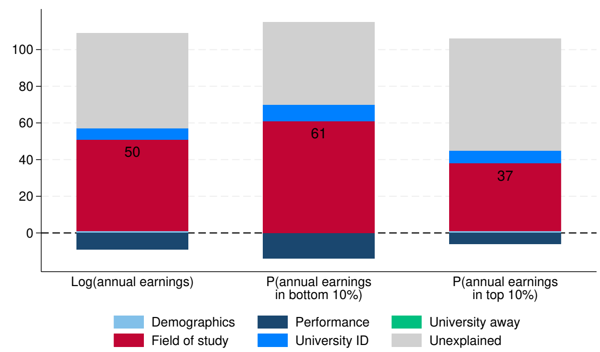
Figure A.3: Financial returns of fields of study 5 years after graduation along the ability distribution, by gender



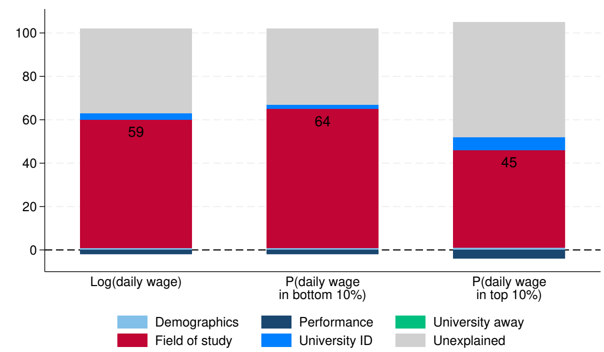
Notes: Separately by gender, this figure plots against ability (measured by the final grade in high school) the financial returns of fields of study choices. The financial return of a major is measured by the median labor income of its male native graduates 5 years after graduation. The measure of ability is available for students who both graduate from high school and obtain a 2nd cycle or single-cycle degree over 2011-18. Financial returns are computed on male native students who graduated from university in 2011-14 and 5 years after graduation work and no longer study.

Figure A.4: The role of education choices for the unconditional gender gap - Oaxaca decomposition, robustness with different samples and outcomes

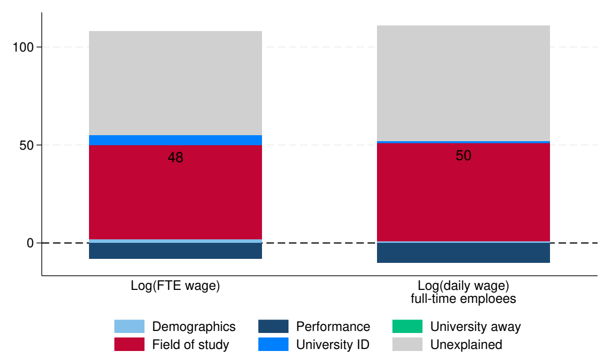
(a) All workers, 5 years after graduation



(b) Private sector employees, 1 year after graduation



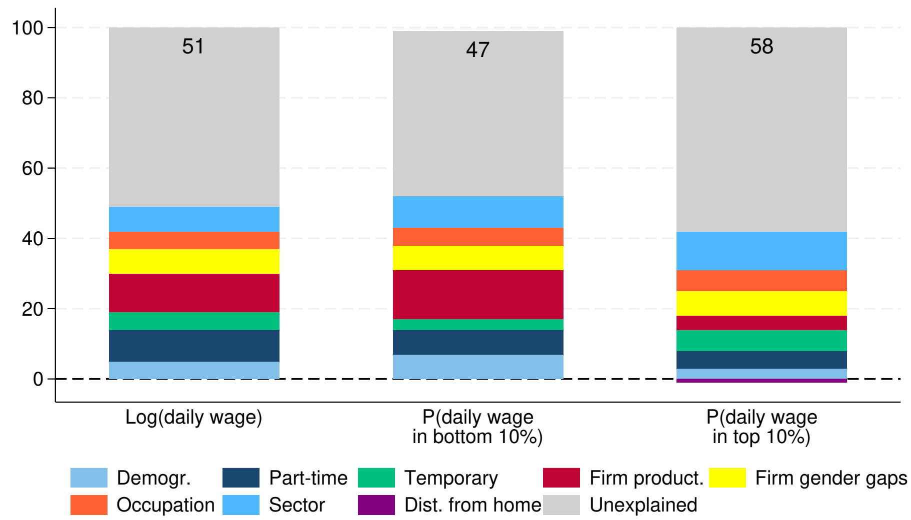
(c) Private sector employees, 5 years after graduation, different outcomes



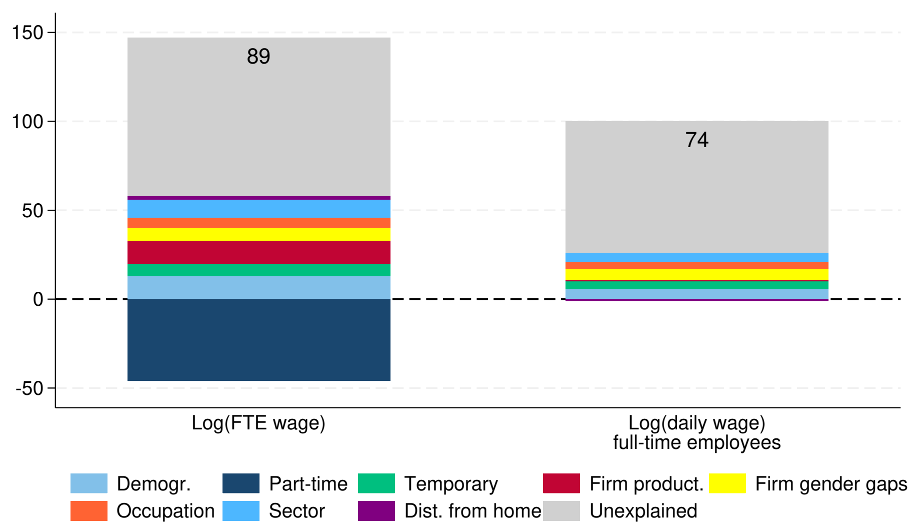
Notes: panel a) cohorts graduating between 2011-14 in a 2nd level and single-cycle degree who are working and not studying 5 years after graduation and with non-missing values for all the controls contained in the vector X in equations 1. Panel b) cohorts graduating between 2011-18 in a 2nd level who are working as private sector employees and not studying 1 year after graduation; the sample is further restricted to individuals with non-missing values for all the controls contained in the vector X in equations 1 and 3. Panel c) the sample is the same as in Figure 5. Overall differences normalized to 100. Variables grouped as in Figure the same as in Figure 5.

Figure A.5: Average contribution of job and firm characteristics for the within-field gender gap - Oaxaca decomposition, robustness with different samples and outcomes

(a) Private-sector employees, 1 year after graduation



(b) Private sector employees, 5 years after graduation, different outcomes



Notes: Panel a) cohorts graduating between 2011-14 in a 2nd level and single-cycle degree who are working as private sector employees and not studying 5 years after graduation. The sample is further restricted to individuals with non-missing values for all the controls contained in the vector X in equations 1 and 3. Panel b) the sample is the same as in Figure 6. Overall differences normalized to 100. Variables grouped as in Figure 6

B Appendix: sample selection

Starting from the full population of graduates, the sample for the main analysis is obtained according to the following sequential steps. *(i)* We select the graduates who 1 (5) year(s) after graduation work in Italy (and do not keep studying).[27] (*ii*) Among those selected in step (*i*), we restrict the attention to those who - according to tax records - draw the main source of labor income from dependent employment. (*iii*) Among those selected in step (*ii*), we retain only the employees who we can successfully and unambiguously match with their employer - based on mandatory reporting forms that firms fill every time they start or end a new contract with a worker. (*iv*) Finally, among those selected in step (*iii*), we focus on private-sector employees. For each step s, Table B.1 reports the share of individuals, in total and separately by gender, that we retain from step $s-1$ and the final sample size as a percentage of the full population.

Table B.1: Sample selection - sequential steps

	All	Men	Women
	A. 1 year after graduation		
Step (*i*)	63%	64%	62%
Step (*ii*)	88%	85%	91%
Step (*iii*)	75%	75%	76%
Step (*iv*)	87%	92%	84%
Final sample (% of population)	36%	38%	36%
	B. 5 years after graduation		
Step (*i*)	74%	74%	75%
Step (*ii*)	80%	81%	77%
Step (*iii*)	85%	83%	86%
Step (*iv*)	81%	89%	77%
Final sample (% of population)	41%	44%	38%

Notes: cohorts graduating in 2011-18 from a 2nd cycle or a single-cycle degree in Panel A; cohorts graduating between 2011-2014 in Panel B. Each row reports the percentage of individuals who is selected in step s, out of the pool selected in step $s-1$.

[27] As explained in footnote 20, we cannot distinguish those who have moved abroad from those who have remained in Italy but neither kept studying nor started working. Hence, the share of individuals who work should not be interpreted in this context as the employment rate of graduates.

C Appendix: sample comparison

Appendix Table C.1 compares 5 samples: all employed graduates (column 1); private-sector employees (column 2); public-sector employees (column 3); self-employed (column 4); and private-sector employees on whom we perform the Oaxaca decomposition (column 5). The samples in columns (2) and (5) are those used for the analysis in Sections 4 and 5.2, respectively. The share of female graduates among private-sector employees is lower than among public-sector employees but higher than among the self-employed. Proxying the socio-economic background with parental income at the time of graduation, private-sector employees are more advantaged than public-sector ones but less so than the self-employed. Comparing the full sample of private-sector employees with the narrower sample on whom the Oaxaca decomposition is performed, it emerges that in the Oaxaca sub-sample the share of females is slightly lower (54% resp. 58%), indicating that attrition out of this sample due to one or more control variables missing is higher for women than for men. Moreover, in the sub-sample on whom we perform the Oaxaca decomposition the early career gender gap is slightly larger (30% resp. 25%). This suggests that women for whom all controls used in the Oaxaca decomposition are available might be negatively selected compared to the average. Nonetheless, the magnitudes of gaps are similar, indicating that findings on the role of fields of study choices and firms' characteristics in shaping the gender wage gap could be generalized to the full population of private sector employees.

Table C.1: Main samples and other samples - basic socio-demographic characteristics

	All employed	Private employees	Public employees	Self-employed	Private employees Oaxaca sample
	(1)	(2)	(3)	(4)	(5)
% female (1 yr)	58.3	57.6	75.0	50.8	53.7
% female (5 yr)	60.2	59.0	77.3	54.7	55.1
% annual earnings gap (1 yr)				21.3	
% annual earnings gap (5yr)				28.3	
% daily wage gap (1 yr)		20.8	17.5		28.0
% daily wage gap (5 yr)		25.5	14.0		30.1
Parental income at grad.	51613	50044	43389	63050	50044

Notes: cohorts graduating in 2011-18 from 2nd cycle or single-cycle degrees for outcomes 1 year after graduation ("1 yr") and cohorts graduating in 2011-14 for outcomes 5 years after graduation ("5 yr"). The statistics on parental income at the time of graduation are computed on the 2011-18 cohorts. For self-employed days worked are not available. The percentage gender gap in daily wages or annual earnings is computed as $100 \times \frac{w_m - w_f}{w_m}$, where f indexes the women and m indexes the men. Across all samples, the focus is on graduates who 1 (5) year(s) after graduation work and no longer study.