Major Choice and Gender Gaps *

Alonso Bucarey Christopher Campos Dante Contreras Pablo Muñoz

October 2024

Abstract

This paper explores how college major returns, preferences, and market design contribute to gender gaps relevant to policy. We leverage data on major preferences to show that majors significantly affect both earnings and fertility, with notable cross-field heterogeneity and modest gender differences in vertical and horizontal returns. Monetary returns largely drive men's major choices, while women prioritize family considerations. Our estimates suggest that 27% of the college gender earnings gap is due to these differing priorities. We also document substantial variation in child penalty gaps across fields, finding that more progressive gender norms within a field are predictive of smaller gender gaps. Finally, we evaluate policy alternatives by simulating counterfactual scenarios within a centralized assignment system. Introducing small gender-based quotas modestly reduces the gender earnings gap without significantly affecting overall efficiency.

Keywords: returns to majors, gender gaps, centralized assignment
JEL Classification: I24, I26, J01, J16

*We thank the Ministerio de Desarrollo Social of Chile for their help in accessing the data. We thank Alessandra Voena and seminar participants at the PUC-Chile, UAI, Universidad de los Andes, Universidad de Chile, and SITE Gender 2024 for their valuable comments and suggestions. The views expressed here are those of the authors and do not reflect the views of the Ministerio de Desarrollo Social. Camila Brown, Connor Fogal, Ryan Lee, Avni Parthan, and Ishira Shrivatsa provided outstanding research assistance. Bucarey: alonsobucareyc@gmail.com. Campos: Chicago Booth School of Business, Christopher.Campos@chicagobooth.edu. Contreras: Universidad de Chile, dcontrer@fen.uchile.cl. Muñoz: Universidad de Chile, pablomh@uchile.cl

1 Introduction

College major choice is a consequential pre-labor market decision that influences numerous life outcomes, including the labor market earnings of men and women (Altonji et al., 2012; Lemieux, 2006; Kirkeboen et al., 2016; Hastings et al., 2013; Oreopoulos and Petronijevic, 2013). While a growing body of evidence documents that gender differences in major choice play an important role in explaining the gender earnings gap among college graduates (e.g., Sloane et al., 2021; Zafar, 2009), still little is known about what factors contribute to these differences in choices (Patnaik et al., 2021). A plausible factor affecting educational choices is the balance of career and family, which has played a central role in women's advancements in the labor market over the last century (Goldin, 2021). Although family considerations are likely an important factor that men and women weigh when choosing potential career paths (Wiswall and Zafar, 2021), limited evidence exists on how fields of study impact family-related outcomes and whether these impacts differ by gender. With a limited understanding of the pecuniary and non-pecuniary returns to majors, many open questions remain about how these factors contribute to gender differences in representation across fields of study and policy-relevant gender gaps.

Previous work has struggled empirically in linking preferences, graduation, earnings, and non-pecuniary factors while methodologically accounting for the self-selection of individuals into majors. The existing literature either documents gender differences in major choice (Zafar, 2009; Bordon et al., 2020; Brenoe and Zolitz, 2020), but does not assess its impact on the gender earnings gap, or studies the role of field of study on earnings (Hastings et al., 2013; Kirkeboen et al., 2016; Dahl et al., 2020) but omits the role of gender differences in preferences. In addition, evidence related to the non-pecuniary returns to the fields of study is scant (Oreopoulos and Petronijevic, 2013; Patnaik et al., 2021) and even more so of gender differences. In this paper, we overcome previous limitations by focusing on the Chilean higher education setting, which allows the linkage of students' preferences for majors to their educational attainment, labor market earnings, and an important non-pecuniary outcome—fertility. The wealth of data allows us to shed light on the various factors that generate rifts in the educational trajectories of men and women at the tertiary level, as well as their consequences.

Descriptive facts motivate our empirical analysis. Similar to findings in Sloane et al. (2021) for the United States, we find that men and women in Chile choose vastly different fields of study and that this representation gap has slowly narrowed over time. A formal decomposition of the earnings gap shows that 82 percent of the raw gender earnings gap among college graduates is due to representation. We simultaneously document substantial heterogeneity in cross-sectional fertility rates by completed major and that these rates negatively correlate with mean earnings ($\rho = -0.57$). These descriptive findings motivate the two focal points of our analysis.

First, the strong negative association between average earnings and fertility points to one explanation for the differences in representation across fields of study that we aim to quantify. To establish such a connection, we must first establish that fields of study causally impact earnings and fertility, with the latter serving as a summary measure for family-related considerations. With estimates of earnings and fertility impacts, we can analyze how each affects major choices, offering clues related to their contribution to the representation and earnings gap. Second, motivated by the fact that representation explains most of the earnings gap, we consider feasible

supply-side policies available in systems with centralized assignment. We focus on policies that can increase enrollment in high-earning majors among students who demand them.

In our conceptual framework, we first distinguish between vertical returns, which represent the general benefits of completing a particular major, and horizontal returns, which capture the idiosyncratic benefits that vary according to individual characteristics and preferences. This distinction helps clarify the assumptions required to causally identify each type of return. Specifically, in settings with centralized assignment systems, certain kinds of selection on gains are plausibly less relevant, while others feature more prominently. In contrast, decentralized systems may require stronger assumptions to identify returns due to greater potential for sorting based on unobserved characteristics. With clarity on the different margins of sorting, we impose standard distributional assumptions on choice behavior along with modest parametric assumptions on potential outcomes that allow us to feasibly estimate vertical and horizontal returns. This framework—borrowing from the school choice literature (Abdulkadiroglu et al., 2020)—leads to an estimation strategy that allows us to decompose the gender earnings gap in new ways. Our approach assumes that the only source of bias in returns to the field of study is due to the correlation between unobserved preference heterogeneity, major choice, and returns, which allows us to construct control functions to flexibly account for self-selection. This framework captures average returns separately for men and women and empirically assesses the importance of preference-specific match effects (Kirkeboen et al., 2016).

Leveraging our framework, we estimate major impacts on earnings and fertility. We document sizable heterogeneity in returns to majors and economically meaningful sorting on gains for both men and women. We find modest gender differences in average returns, with statistically different returns in Science, Engineering, and Teaching, but similar returns in other fields of study. With measures of preference intensity derived from rank-ordered choice data, we further show that match effects are empirically important, indicating that students tend to sort in ways that are consistent with Roy (1951)-like selection. As in Kirkeboen et al. (2016), but through a different empirical lens, we find that individuals with stronger preferences for a given field tend to earn more if they earn a degree in that field of study. However, gender differences in match effects tend to be minimal, suggesting that selection on gains plays a minimal role in explaining within-major earning gaps. Although both vertical and horizontal causal returns contribute to the within-major earnings gap, their contribution is minimal relative to differences in representation, which we return to in our counterfactual analysis.

While a focus on the pecuniary returns to majors follows from the perspective that education is an investment (Becker, 1964), different fields of study expose students to vastly different peer groups and job opportunities (Barrios et al., 2021; Zimmerman, 2019). While those may be perceived as intermediate outcomes affecting the returns to particular fields of study, these factors are also key determinants affecting subsequent fertility decisions. Indeed, we show that the choice of major causally impacts both extensive and intensive margins of fertility, suggesting that this is a key non-pecuniary factor that students might consider while choosing a major. In contrast to earnings results, we do not find evidence of match effects regarding fertility. Moreover, we show that major choice, earnings, and fertility interact in economically meaningful ways by documenting sizable variation in the child penalty gaps across fields of

study. This descriptive exercise reveals a short-run average penalty of 18% with a standard deviation of 13%. These pieces of evidence are, to the best of our knowledge, the first to show that major choice causally impacts fertility and also to document major-specific variation in the child penalty gap.

Our empirical approach assumes that preference heterogeneity governs selection into majors, which we account for using analytical control functions derived from rank-ordered choices. An alternative approach would be to restrict the analysis to the local quasi-random variation embedded in centralized assignment systems (Abdulkadiroglu et al., 2017; Abdulkadiroğlu et al., 2022), limiting the additional flexibility our model gives us to study rich selection patterns. We opt for the former but use the local quasi-random variation to validate our approach. Evidence from a validation exercise (Deming, 2014; Angrist et al., 2017) demonstrates that our preferred model adequately characterizes selection into majors as it produces treatment effects that line up with quasi-experimental approaches leveraging quasi-random assignment to majors (i.e., our estimates are forecast unbiased). The forecast unbiasedness of our estimates holds for both pecuniary and non-pecuniary returns, underscoring the strength of our empirical approach.

Equipped with validated pecuniary and non-pecuniary return estimates, we ask whether gender differences exist in the implicit weights students assign to these outcomes and probe for potential explanations. We provide suggestive evidence that women prioritize family considerations (i.e. fertility impacts) over earning impacts compared to men, as captured by multivariate relationships between estimated major- and cell-specific mean utilities and our estimated fertility and earnings returns. This evidence indicates that these anticipated impacts might generate rifts in the educational and labor market trajectories of men and women before they enter the labor market. To quantify the potential importance, we find that equalizing preferences for family considerations across men and women leads to a 27 percent reduction in the gender earnings gap. This suggests that gender differences in the balance of career and family (Goldin, 2021) influence major choices and have sizable consequences on the broader gender earnings gap.

To shed light on the potential drivers of these gender differences in observed preferences, we consider the role of gender norms as captured by breadwinner norms (Bertrand et al., 2015). We find striking heterogeneity across majors in the degree to which the social norm "a man should earn more than his wife" influences the distribution of relative labor earnings. Consistent with gender norms distorting women's labor supply decisions both before and after birth, we show that gender norm progressivity is highly predictive of improvements in return and child-penalty gaps. Although suggestive, this evidence contributes to the mounting evidence demonstrating how the prescriptive nature of norms potentially plays a role in shaping important labor market outcomes (Akerlof and Kranton, 2000; Bertrand, 2020; Oh, 2023; Bursztyn et al., 2017).

In the final part of our paper, we assess the impact of feasible policies that policymakers could use to address the gender gap and that do not involve changing preferences. Settings with centralized assignment systems can flexibly change assignment rules. With a centralized assignment system in place and equipped with causal estimates of major returns, we can assess the impacts of rule changes on the gender earnings gap of college applicants. Our analysis is

guided by the descriptive evidence underscoring the significant role of differing representation in the high-earning field of Engineering in explaining the gender earnings gap.

We consider two policies that can affect the allocation of Engineering seats. The first policy focuses on a gender-neutral increase in capacity across all Engineering programs, while the second policy holds the number of available Engineering seats constant and reserves a percentage of seats for female applicants. Bursztyn et al. (2023) find that the latter affirmative-action-like policy receives broad support in Chile. We find that any gender-neutral increase in Engineering seats only exacerbates the Engineering representation gap and the overall gender earnings gap.[1] In contrast, policies that reserve even modest shares of Engineering seats to females substantially reduce the representation gap but produce a modest reduction in the gender earnings gap. The reduction in the representation gap is mechanical, while the modest reduction in the gender earnings gap arises from a tradeoff between increased access to higher earnings majors for women and reductions in graduation probabilities (Arcidiacono and Lovenheim, 2016; Arcidiacono et al., 2016). We find surprisingly minimal impacts on efficiency as measured by changes in the aggregate earnings across all students in a given quota counterfactual.

Our work contributes to three distinct but loosely connected literatures. The first relates to an impressive body of work focusing on estimating the returns to college and college selectivity (Card, 2001; Dale and Krueger, 2002; Mountjoy and Hickman, 2021; Mountjoy, 2022; Zimmerman, 2014; Chetty et al., 2023; Mountjoy, 2024), and to a growing body of work that pivots focus to the returns to fields of study (Altonji et al., 2012; Hastings et al., 2013; Kirkeboen et al., 2016; Dahl et al., 2020; Andrews et al., 2022; Lovenheim and Smith, 2023). One reason for the more recent interest in returns to fields of study is the fact that dispersion in field returns is as large as the dispersion in returns to college (Altonji et al., 2012). The large dispersion in returns to majors alludes to the potentially important role of major choice in explaining the gender earnings gap among college graduates (Sloane et al., 2021). While a central component of this paper relates to estimating field-specific returns among college graduates, our analysis makes three additional contributions. First, we make progress in providing a more complete picture of how majors impact outcomes by estimating non-pecuniary returns, something there is scant evidence on (Oreopoulos and Petronijevic, 2013). Second, we allow pecuniary returns to vary by gender to quantify its contribution to the overall gender gap. Third, we leverage rich preference data to understand how preference heterogeneity interacts with match effects and efficiency in the context of reallocative policies of interest.

A related body of work in higher education has documented vast gender differences in major choice (Blakemore and Low, 1984; Turner and Bowen, 1999; Bordon et al., 2020; Neilson et al., 2021; Wiswall and Zafar, 2018; Ahimbisibwe et al., 2024; Wiswall and Zafar, 2021; Zafar, 2009). Bordon et al. (2020) and Ahimbisibwe et al. (2024) find that men apply to selective programs even when they are marginal candidates, while equally qualified women tend to apply less often to these programs. Women's peer groups have also been documented as important determinants of major choice (Brenoe and Zolitz, 2020). Wiswall and Zafar (2021) finds that college students'

[1] This finding results from roughly 30 percent of males listing Engineering programs as their most preferred, while only about 10 percent of women do. Therefore, any gender-neutral increase in capacity results in a disproportionate increase in male representation, worsening the gender earnings gap through both return differentials within Engineering and a disproportionate reallocation of males to high-earning fields.

expectations about how their education will impact their future career and family significantly influence their choice of major and degree completion, with notable differences observed between genders. We contribute to this literature in a variety of ways. First, we show and quantify the importance of representation to the overall gender earnings gap. Second, by estimating both pecuniary and non-pecuniary returns, we can quantify the extent to which educational decisions diverge due to these two important causal factors and quantify the extent to which gender differences in family considerations contribute to the gender earnings gap.

Last, this paper contributes to a growing body of evidence studying the potential effects of affirmative action in systems without centralized assignment systems (Bleemer, 2023; Arcidiacono et al., 2016; Arcidiacono, 2005) and with (Otero et al., 2021; Carlana et al., 2024). Sparked by Abdulkadiroğlu and Sönmez (2003), centralized assignment is an increasingly popular way of allocating seats in education markets, with Neilson (2021) showing that at least 50 countries use centralized assignment. This increase has coincided with an increase in affirmative action policies implemented through adjusting priorities within a given system. Relative to recent work (Ellison and Pathak, 2021; Otero et al., 2021; Carlana et al., 2024), we pivot the focus to gender-based affirmative action and study the potential of such policies in reducing gender earnings gaps among college graduates. In light of the findings in Bursztyn et al. (2023) that there is broad support for gender-based affirmative action, this paper provides evidence of utmost importance to policymakers in Chile and abroad.

The rest of this paper is organized as follows. Section 2 outlines the institutional setting and describes the data we use; Section 3 presents the conceptual framework; Section 4 reports causal estimates of the impacts of major choice on earnings and fertility and gender differences in preferences for these causal impacts; Section 5 argues that variation in gender norms play a role influencing the variation in the various gender gaps we document; in Section 6 we turn to assess various counterfactual assignment policies and their impact on the gender earnings gap; and last, in Section 7 we conclude.

2 Institutional Background, Data, and Descriptive Statistics

This section presents background information on Chile and its centralized college admission system. We discuss the data generated by the college admission system and its linkages with other administrative records that help to facilitate our analysis. We also discuss descriptive evidence regarding gender gaps and college major sorting.

College Admissions in Chile: Higher education admissions are based on high-school GPA and students' test scores at nationwide entrance exams. High school graduates can register for this national admission test (Prueba Selección Universitaria, PSU), an SAT-type exam with four sub-tests: Mathematics, Language, and the choice of taking Science or History. After receiving their scores, students can apply to a major and institution simultaneously (e.g., civil engineering at the University of Chile). Students must submit a rank-ordered list reflecting their preferences (they can rank up to ten college-major combinations). Each major-institution (a program) also ranks applicants by defining weights on the subject test scores and the high school

GPA. Given students' rankings, program-specific weights, and capacities, the system offers students a seat at one program at most so that no student-program pair prefers each other and is not assigned together. This is done using an assignment mechanism built on Gale and Shapley (1962)'s student-proposing deferred acceptance algorithm. This process creates a cutoff for admission at each program, which corresponds to the score of the least qualified admitted student.

Data: Our analysis combines rich administrative data on students' college applications, graduation, fertility, and labor market earnings. Data on college applications and admissions comes from the agency responsible for administering the national college entrance exam (Departamento de Evaluación, Medición y Registro Educacional, DEMRE). For all students who take the college admission exam, we observe their gender, high school identifier and high school location, GPA, and test scores. We also observe the rank-ordered lists of preferences for major institution pairs. Since our focus is on the returns of college graduates, we merge this data with records of students' graduation coming from the Higher Education Information Service (Servicio de Información de la Educación Superior, SIES).

We classify students into fields of study based on the major from which they graduate. Degree programs are classified by field of study mostly based on the OECD Handbook for Internationally Comparative Education Statistics (OECD, 2004). There are eight broad categories: "*Agriculture*", "*Science*", "*Social Sciences, Business and Law*", "*Teaching*", "*Humanities and Arts*", "*Engineering, Manufacturing and Construction*", "*Health and Welfare*", and "*Services*". To ease the exposition, we reclassify "*Social Sciences, Business and Law*" into three separate fields: "*Social Sciences*", "*Business*", and "*Law*". We also separate "*Medicine*" from "*Health and Welfare*" and drop "*Agriculture*" and "*Services*" as they represent small (<5% of graduates) and non-homogeneous college programs in Chile.[2] All in all, we end up with nine fields of study. See Appendix A.1 for details.

Fertility data comes from the Civil Registration System (Registro Civil). As of 2023, we observe the date of birth of all the children of students who applied to college. Labor market earnings come from the Chilean pension system (Superintendencia de Pensiones, SP), which administers the social contributions from all formal public and private labor relationships. This is a universal, high-quality administrative dataset, but it has two limitations. First, it does not include the hours in the workers' contract. However, part-time employment is not that common among college graduate workers in Chile.[3] Second, wages are censored at the social security contribution limit. We address this issue by imputing wages above the limit as in Dustmann et al. (2009) and Card et al. (2013) (see Appendix A.2 for details). We complement the SP data with administrative records from the Unemployment Insurance Administrator Agency (Administradora de Fondos de Cesantía, AFC). This data only covers formal labor relationships in the private sector but we use it for a dual purpose. First, it allows us to reproduce our main results among a subset of individuals in the private sector with a higher wage censoring limit. Second, AFC data contain firm identifiers which allows us to estimate a two-way fixed effects

[2] Appendix Figure A.5 presents the distribution of college admission test scores in each field of study.

[3] According to the household survey CASEN 2017, only 8.6% of prime-age college graduates report to have worked part-time.

model to proxy for firm premia (Abowd et al., 1999; Card et al., 2013) and quantify the extent to which gender differences in firm sorting contribute to the gender earnings gap.

For our analysis, we focus on students who applied to college through the centralized admission system between 2004 and 2007, and who were accepted at least into a program. Table 1 presents descriptive statistics for individuals in our sample for whom we observed labor market earnings in 2019. There is much variation in the number of students, their average admission test scores, and demographics across genders and fields of study. Engineering has the highest number of male graduates, whereas Teaching has the highest number of female graduates. Medicine is the most selective field, while Teaching is the least selective. Compared to Medicine, Teaching graduates come from a more disadvantaged background (e.g., less than 20% of their fathers have a college degree versus 50% of the fathers of graduates from Medicine). Across fields and genders, students from public schools and with public health insurance are significantly underrepresented.

Gender Earnings Gaps in Chile: We now present cross-sectional and time series evidence regarding gender gaps in Chile to provide perspective and benchmark our focal analysis.

We begin with a historical lens documenting trends in gender representation across majors and its potential contribution to the overall gender gap. To do so, we use the largest household survey in Chile (CASEN by its Spanish acronym). Respondents of the 2017 survey wave report their majors and wages, which allows us to decompose changes in the gap in the spirit of Sloane et al. (2021)'s analysis focusing on United States college graduates. Specifically, we construct a major "similarity index" (as a re-normalization of the inverse Duncan-Duncan index) and a "potential wage index" (that assigns to everyone within a major the average hourly wages of prime-age male workers in that major). Appendix Figure A.2 presents the across-cohorts evolution of both indexes. As in the US, majors in Chile have become less segregated over time. As shown by panel (a), the similarity index increased from 0.55 for the 1960 birth cohort to 0.68 for the 1990 birth cohort, indicating that both genders are more evenly spread across majors for more recent cohorts.[4] Coupled with the improvement in gender integration, the earning potential of women has also increased over cohorts. Nonetheless, as shown by panel (b), women still have a lower earnings potential than males due to their sorting into lower-paying majors.

Turning to more recent cross-sectional differences, panels (a) and (b) of Figure 1 show the graduation and log earnings of the graduated individuals in our main estimation sample by gender and field of study. A salient fact from this figure is the "STEM gap" as men graduate disproportionally more from Engineering and Science, while females graduate disproportionally more from Health and Teaching, which are lower-paying fields than Engineering and Science. This suggests that the gender earnings gap might result from differences in preferences for fields of study. One potential culprit partly governing choice differences is a non-pecuniary outcome, fertility, which we show in panel (c). Fields of study with higher earning levels tend to have lower fertility outcomes, with a correlation between earning and fertility returns equal to -0.57.

[4]For the US, using the 2014-2017 American Community Survey, Sloane et al. (2021) reports that the same index increased from 0.55 for the 1950 birth cohort to 0.64 for the 1990 birth cohort.

To quantify the importance of sorting and representation, we decompose the gender earnings gap among college graduates. For this purpose, let μ_j^g stand for the average earnings of individuals of gender g who graduate from major j, and let s_j^g stand for the share of individuals that graduated from field j among those of gender g. Using this notation, we can decompose the overall earnings gender gap among college graduates $\mathbb{E}[\text{Earnings} \mid \text{Female}] - \mathbb{E}[\text{Earnings} \mid \text{Male}]$ into differences in sorting across fields of study and differences in returns within fields of study:

$$\sum_j s_j^F \mu_j^F - \sum_j s_j^M \mu_j^M = \underbrace{\sum_j (s_j^F - s_j^M)\mu_j^F}_{\Delta\text{Representation}} + \underbrace{\sum_j s_j^M (\mu_j^F - \mu_j^M)}_{\Delta\text{Returns}} \tag{1}$$

The first term of Equation (1) captures differences in sorting patterns holding earnings constant; thus, it represents the share of the gap explained by the fact that men and women graduate from different fields of study. The second term holds the sorting patterns constant to capture within-major earnings differences, which may be driven by labor market factors disproportionately affecting women, such as discrimination, child penalties, or differential firm sorting. Figure 2 presents the results from this exercise that decomposes the gender earnings gaps—which corresponds to 19.2% ($4,758 USD per year in favor of men)—into these two components. We see that Engineering and Science are the main contributors to men's higher earnings and that this is mostly due to the low representation of females in these fields.

Three important takeaways from this descriptive evidence inform our main analysis. First, gender differences in earnings and fertility outcomes potentially contribute to the representation gap in economically meaningful ways. Second, the decomposition demonstrates that the representation component is sizable, accounting for 82% of the gender earning gap. Third, the underrepresentation of women in high-earning fields such as Engineering and Science is an important channel explaining the overall earnings gap. The latter takeaway points to policies that can affect the allocation of students to fields of study, which in many settings, including Chile, are much more feasible given centralized assignments (i.e., policymakers have more control over s_j^g in the sense that varying priority structures allows policymakers to address the portion of the gap driven by across majors representation gaps). The scope of such policies in reducing the gender earnings gap hinges on return estimates, rich treatment effect heterogeneity, and the extent to which there is excess demand among women. We formalize these factors next in our conceptual framework.

3 Conceptual Framework

One of the primary goals of our analysis is to study factors contributing to the gender earnings gap among college graduates, which requires us to estimate monetary returns to different fields of study and their impact on fertility. Recent advances in the literature have proposed new ways of estimating treatment effects in systems with centralized assignment (Kirkeboen et al., 2016; Abdulkadiroglu et al., 2017, 2020). Given our focus on major choice and the potential impact of reallocative policies, we adopt a framework that accounts for both vertical and horizontal dimensions of major effects. We start with a largely non-parametric discussion of major

effects, following the approach of Mountjoy and Hickman (2021), and transparently analyze how students sort into majors, considering which types of sorting are most appropriate in our context. We then connect this discussion to our empirical strategy, which draws from the school choice literature (Abdulkadiroglu et al., 2020), integrating a major choice model with potential outcomes for identification.

Vertical and Horizontal Major Effects: We index the population of accepted college applicants by $i \in \mathcal{I}$ and majors by $j \in \mathcal{J}$. The potential outcome of student i completing major j is:

$$Y_{ij} = Y_{i0} + \mu_{ij} \tag{2}$$

where Y_{i0} is their potential outcome if they do not graduate from college, or a summary measure of their ability, and μ_{ij} is the idiosyncratic return of major j. For expositional purposes, we denote ability as $a_i = Y_{i0}$. Denote $F_i \in \mathcal{J}$ as student i's completed field or major.

Applicants submit a rank-ordered list, $R_i \in \mathcal{R}$, containing a partial ordering of majors, and also report an entrance exam score, $S_i \in \mathcal{S}$. A student type is defined as $\theta_i = (R_i, S_i) \in \Theta$. In previous work, it has been argued that information in college application and admissions portfolios reveals information about student ability (Dale and Krueger, 2002; Mountjoy and Hickman, 2021). In settings with centralized assignments, where access to majors is strongly governed by entrance exam scores, the intuition behind the self-revelation principle also applies. Therefore, we assume that the composition of the rank-ordered list, in addition to entrance exam scores, adequately summarizes the student's ability.

Assumption 1. *Student types proxy for student ability,*

$$a_i = \sum_{\theta \in \Theta} \phi_\theta \mathbb{1}[\theta_i = \theta] + u_i,$$

with $E[u_i|F_i, \theta_i] = E[u_i|\theta_i]$.

Assumption 1 states that any residual variation in ability, not captured by student type θ, is not systematically related to their choice of major. With access to rank-ordered choice data and entrance exams, we can now ask what observational comparisons recover under Assumption 1. With observed outcomes, Y_i, we can further fix $\theta_i = \theta$ and compare the mean outcomes of major j graduates to those of non-graduates ($j = 0$):

$$E[Y_i|F_i = j, \theta] - E[Y_i|F_i = 0, \theta] = \underbrace{E[\mu_{ij}|F_i = j, \theta]}_{ATT_j(\theta)} + \underbrace{E[a_i|F_i = j, \theta] - E[a_i|F_i = 0, \theta]}_{\text{Selection bias} = 0, \text{ by Assumption 1}}. \tag{3}$$

Equation (3) shows that observational comparisons within types θ recover θ-specific average treatment on the treated effects. While empirically interesting and important to some extent, this quantity masks different sources of essential heterogeneity (Heckman et al., 2006) that may be important to policymakers. To see this concretely, we can write the $ATT_j(\theta)$ as:

$$ATT_j(\theta) = \underbrace{E[\mu_{ij} \mid \theta]}_{ATE_j(\theta)} + (E[\mu_{ij} \mid F_i = j, \theta] - E[\mu_{ij} \mid \theta]).$$

The first term is the average treatment effect among students type θ. The latter term captures selection on gains in terms of major choice among those with preferences R and exam score S. For example, selection on gains would be positive among students of type θ if the returns to those who complete major j are larger than those obtained by the average student type θ. In decentralized settings, this kind of sorting is more of a concern as there is a lot of scope for residual sorting even after you reveal your preferences and exam scores. In a setting with centralized assignments and single offers, however, this kind of sorting is less of a concern as a multitude of lotteries govern final access to majors. We summarize this in the following assumption.

Assumption 2. *In settings with single-offer centralized assignment of students to programs and majors, let $Z_i \in \mathcal{J}$ correspond to a student's offered major and let D_i be an indicator for compliance with the assignment. Let F_i correspond to their enrolled or completed major,*

$$F_i = D_i Z_i + (1 - D_i) F_i',$$

where F_i' is a major student i completes in case they do not comply. Compliance, D_i, and residual ability, u_i, are as good as random, conditional on type θ:

$$D_i \perp\!\!\!\perp \mu_{ij} \mid \theta \quad \text{and} \quad D_i \perp\!\!\!\perp u_i \mid \theta.$$

Assumption 2 states that, conditional on type, any non-compliance is unrelated to major returns and student ability. The assumption is analogous to the conditional independence assumption in Angrist et al. (2024), requiring students to comply with their randomized offer for estimation of risk-controlled value-added models. This assumption rules out selection on gains into majors conditional on type, formalized in the following lemma.

Lemma 1. *In settings with a single-offer centralized assignment and compliance that is as good as random conditional on type, selection on gains into majors conditional on type is zero:*

$$E[\mu_{ij} \mid F_i = j, \theta] - E[\mu_{ij} \mid \theta] = 0$$

These results utilize institutional features of centralized assignment systems to rule out the kind of sorting that has raised concerns in estimating the returns to college selectivity (Dale and Krueger, 2002).[5] Nonetheless, the conditional average treatment effect nests additional horizontal effects of interest that cannot be assumed away. Similar to the decomposition above, we can write:

$$ATE_j(\theta) = \underbrace{E[\mu_{ij} \mid \mathcal{I}]}_{ATE_j(\mathcal{I})} + \underbrace{(E[\mu_{ij} \mid \theta] - E[\mu_{ij} \mid \mathcal{I}])}_{M_{ij}}, \tag{4}$$

[5] As Hoxby (2009) points out, "the [Dale and Krueger] strategy generates estimates that rely entirely on the small share of students who make what is a very odd choice. These are students who know that they could choose a much more-selective college and who have already expressed interest in a much more-selective college (they applied), yet, they choose differently than 9 out of 10 students. Almost certainly, these odd students are characterized by omitted variables that affect both their college decision and their later life outcomes." That observation underscores the empirical relevance of this kind of sorting in decentralized markets. In settings with centralized assignment, however, this kind of concern is much attenuated.

where the θ-specific average treatment effect depends on the average treatment effect among the population of college applicants, $\mathcal{I}$, and additional sorting terms. In contrast to the initial sorting terms above, these terms capture how idiosyncratic treatment effects among type θ deviate from the average treatment effect for the typical college applicant. This sorting on gains is analogous to the type considered by Kirkeboen et al. (2016), where preference heterogeneity is associated with potential outcomes directly. This sorting is more closely associated with the formation of the rank-ordered list, an element of θ, governed by inherent preferences, and less so with the actual decisions taken at a later stage.

In the next section, we introduce a choice model that provides a framework to summarize student types, θ, and proxy for student ability in a more tractable, lower-dimensional form. By incorporating elements of this choice model into the potential outcomes, we can effectively capture the sorting on gains, M_{ij}. Through a combination of distributional assumptions on unobserved preference heterogeneity and parametric assumptions on potential outcomes, this approach enables us to recover both the average treatment effect, ATE_j, and the sorting term, M_{ij}. We begin by detailing the structure of the choice model.

Major Choice: We model demand for majors, abstracting away from demand for particular programs or institutions, relative to a left out field of study.[6] The empirical demand model we estimate nests a few assumptions that we now discuss. Student i's indirect utility in outcome j is given by:

$$u_{ij} = \sum_{k \in \mathcal{K}} \rho^k_{c(X_i,G_i)} ATE^k_j(X_i, G_i) + \lambda_{dc(X_i,G_i)} d_{ij} + \lambda_{pc(i)} p_{c(X_i,G_i)j} + \eta_{ij} \tag{5}$$

where d_{ij} is the distance between student i's high school and their most-preferred program corresponding to field j, $p_{c(X_i,G_i)j}$ is the mean net price of field j among students in cell c, and η_{ij} captures any remaining unobserved preference heterogeneity. Cells include two school types, three macro-regions, three financial aid-relevant PSU ranges, and gender.

Equation (5) nests a few empirically convenient assumptions. First, the model assumes students can forecast their mean potential outcomes, and as a consequence, their covariate- and gender-specific average treatment effects denoted by $ATE^k_j(X_i, G_i)$.[7] Second, we parameterize preference heterogeneity to be covariate-specific and aggregated into distinct cells $c(X_i, G_i)$. Third, and related to financial aid and net price, students in Chile face different financial aid schedules governed by rigid discontinuities in the PSU distribution. We account for this in our choice of cell strata, which include different non-overlapping PSU score groups. Therefore, we assume that students' net price is constant within a cell. The three assumptions allow us to summarize the mean utility for major j among those in cell $c(i) = c(X_i, G_i)$ as:

$$\delta_{c(i)j} = \sum_{k \in \mathcal{K}} \rho^k_{c(i)} ATE^k_j(X_i, G_i) + \lambda_{pc(i)} p_{c(i)j}.$$

[6] Collapsing demand estimation is common in the literature. Abdulkadiroglu et al. (2020) estimate demand for schools as opposed to programs, similar to us estimating demand for majors instead of programs. Laverde (2022) uses similar aggregations in the context of pre-K in Boston.

[7] We view rational expectations assumption on mean outcomes as more reasonable than rational expectations on potential outcomes Y_{ij}.

The empirical indirect utility specification is therefore:

$$u_{ij} = \delta_{c(i)j} + \lambda_{dc(i)} d_{ij} + \eta_{ij}.$$

Following standard assumptions in the discrete choice literature, we assume that η_{ij} has an Extreme Value Type 1 distribution conditional on $\delta_{c(i)j}$ and d_{ij}.

Identification of Returns: Our empirical approach links potential outcomes to the choice model by allowing the unobserved preference heterogeneity to affect potential outcomes. For $j = 1, \cdots, J$, the potential outcome of student i is:[8]

$$Y_{ij} = \underbrace{\gamma' X_i + \sum\nolimits_l \psi_l(\eta_{ir} - \bar{\eta})}_{a_i} + \underbrace{\underbrace{\alpha_j + \beta_j G_i}_{ATE_j(G_i)} + \underbrace{\psi_j^*(\eta_{ij} - \bar{\eta})}_{M_{ij}}}_{\mu_{ij}} + e_{ij}\,. \tag{6}$$

Equation (6) imposes a few parametric assumptions relative to Equation (2). First, student ability is fully summarized by elements of X_i, which include entrance exam scores, and unobserved preference heterogeneity. Elements of X_i include student entrance exam scores, and adjusting for an entrance exam score is common in the literature (Altonji et al., 2012; Andrews et al., 2022). In addition to that, we allow the potential outcomes of two individuals with identical exam scores to vary if their underlying preferences are different. The coefficient ψ_l captures this kind of selection on levels, where $\psi_l > 0$ suggests that individuals with stronger unobserved tastes for major l have larger potential outcomes, regardless of the major they complete. Second, the average treatment effect is summarized by α_j and β_j allows for a gender differential. Third, M_{ij} captures the preference-specific sorting on gains discussed above and is linear in the unobserved preference heterogeneity. The ψ_j^* captures the so called match effects, commonly referred to as *selection on gains* in these types of models (Abdulkadiroglu et al., 2020; Einav et al., 2022; Otero et al., 2021; Bruhn et al., 2023). In the case that $\psi_j^* > 0$, students that both have a higher preference for major j and complete major j experience an additional earnings gain, evidence of positive Roy (1951)-like selection.

The conditional expectation of an observed outcome Y_i is given by:

$$E[Y_i|R_i, X_i, G_i, j(i) = j] = \alpha_j + \beta_j G_i + \gamma' X_i + \sum_l \psi_l \lambda_l(\theta(R_i, X_i), G_i, D_i) + \psi_j^* \lambda_j(\theta(R_i, X_i), G_i, D_i), \tag{7}$$

where $\lambda_j(\theta(R_i, X_i), G_i, D_i)$ are control functions derived from rank-ordered choice data contained in R_i and an extreme value type 1 (EVT1) assumption on the unobserved preference heterogeneity. Therefore, we can estimate heterogeneous major returns with estimates of λ_j for each major j.

Define the return to major j relative to the outside option as $\tau_j(\theta(R_i, X_i), G_i, D_i)$. Through the lens of the model above, the conditional average treatment effect of major j is:

$$\tau_j(\theta(R_i, X_i), G_i, D_i) = \alpha_j + \beta_j G_i + \psi_j^* \lambda_j(\theta(R_i, X_i), G_i, D_i). \tag{8}$$

[8] We ignore the index k in the subsequent discussion for expositional ease.

The parameter ψ_j^* governs the degree to which preference heterogeneity, governed by X_i, R_i, G_i, and D_i, maps to variation in treatment effects. The parameter ψ_j^* is identified by comparing students completing the same major but with differing measures of λ_j. The β_j are identified similarly but comparing students completing the same major but of different genders. In contrast, the α_j are identified by comparing students that are identical, including their λ_j, but vary in their completed major. The embedded random assignment of students to majors in centralized assignment systems, along with distance variation, induces the variation in completed majors conditional on their preferences.

What kind of sorting on gains does ψ_j^* identify? To begin, note that $\lambda_j(\theta(R_i, X_i), G_i, D_i)$ does not depend on students' actual completed major. This means that the parametric assumptions imposed above rule out sorting on gains into majors among individuals of the same type, analogous to Assumption 2 discussed above. In contrast, the $\lambda_j(\theta(R_i, X_i), G_i, D_i)$ do capture selection into application portfolios analogous to the sorting reflected in Dale and Krueger (2002) and Kirkeboen et al. (2016).

It is worth mentioning that this approach bears additional resemblance to fallback conditioning presented in Kirkeboen et al. (2016), as treatment effects vary according to the composition of a student's rank-ordered list, R_i. For example, the return to a given field includes a match effect governed by the submitted rank-ordered list, including the fallback option and any subsequent fields a student ranks. This means that treatment effects for major j differ for individuals with different fallbacks, and even among individuals with similar fallbacks, additional differences in the composition of the rank-ordered list produce additional differences in treatment effects. Therefore, our assumptions governing selection into majors play a dual role. They allow us to estimate causal effects by accounting for preference heterogeneity flexibly while simultaneously allowing us to overcome power limitations from other data-demanding approaches emphasizing the importance of fallback conditioning (Kirkeboen et al., 2016; Dahl et al., 2020). Accounting for this heterogeneity in returns is crucial for any reallocative policy.

Estimating the Choice Model: We leverage rich information contained in applicants' rank-ordered lists to estimate the parameters of the choice model that then allow us to construct the control function $\lambda_j(\theta(R_i, X_i), G_i, D_i)$. We rely on the fact that truthful reporting in the Chilean centralized assignment system is a weakly dominant strategy.[9] A student reveals field j as their most preferred if they list it at the top of their list, the next-stated field is their second preferred, and so on. Preferences are summarized by a vector $R_i = (R_{1i}, R_{2i}, \cdots, R_{j(i)i})$ where R_{1i} is the top-listed field and the length of the list $j(i)$ is allowed to vary across students.[10] Thus, under truthful reporting, students' top-ranked field satisfies:

$$R_{1i} = \arg\max_{s \in \mathcal{J}} U_{is}$$

[9] Chile uses a variant of a student-proposing deferred acceptance algorithm for seat allocation in which truthful reporting is a weakly dominant strategy.

[10] For example, some students rank a single field across their list, while others rank multiple fields.

while the remaining options satisfy:

$$R_{ki} = \arg\max_{s \in \mathcal{J} \setminus (R_{1i}, \cdots, R_{k-1i})} U_{is}, \quad k > 1.$$

The EVT1 assumption allows us to express the likelihood of observing the major-specific rank-ordered list as follows:

$$\mathcal{L}(R_i|X_i) = \prod_{k=1}^{j(i)} \frac{\exp(\delta_{kc} + \lambda_{dc} d_{ik})}{\sum_{r \in R_i \setminus R_{1i}, \cdots, R_{k-1}} \exp(\delta_{rc} + \lambda_{dc} d_{ir})}.$$

We estimate preference models separately for the 36 covariate cells defined above (i.e., 2 school types, 3 macro-regions, 3 financial aid-relevant PSU ranges, and gender) via maximum likelihood to obtain a list of major-specific mean utilities.

One potential concern with our estimation approach is potential deviations away from truthful reporting. As Fack et al. (2019) points out, some students may skip the impossible, or constrained list lengths may introduce other empirical challenges (Haeringer and Klijn, 2009). To the latter point, few students rank all fields of study, so that concern does not apply in our setting. However, we may observe some students never ranking Medicine because their PSU score is far from the cutoff, even though they may prefer Medicine programs. Estimating preferences under an *ex-post stability* assumption is an increasingly popular approach that circumvents these issues. We choose not to use this approach in our main analysis for several reasons. To begin, estimating demand is not our end goal, and our focus is to estimate returns to the field of study and field-specific match effects. This requires identifying the λ_j, which is complicated by the implicit extrapolations embedded in the stability assumption.

To avoid these kinds of assumptions, we adopt an approach that utilizes all of the information contained in each individual's rank-ordered list to construct individual-specific control functions, emphasizing that approaches that exclude sizable amounts of data do so at the potential cost of misspecifying control functions, which are essential for identification.[11]

4 Returns to Fields of Study, Preferences, and Gender Gaps

This section reports the pecuniary and non-pecuniary impacts of various fields of study, with estimates of mean returns and match effects that vary by gender. We also show the robustness of our empirical approach by testing for bias in the spirit of Deming (2014) and Angrist et al. (2017). Having documented significant differences in returns across fields of study and gender, we examine whether men and women weigh pecuniary and non-pecuniary outcomes differently and analyze how these differences may contribute to the earnings gap. We conclude this section by providing descriptive evidence of significant heterogeneity in estimated child penalties across fields of study, demonstrating the choice of major, earnings, and fertility interact in the labor market in economically meaningful ways.

[11] As a descriptive assessment of the information embedded in these control functions, Appendix Figure A.3 reports a series of bivariate correlations of our estimated control functions across fields of study. The evidence is consistent with anecdotes and other leading empirical findings (Altonji et al., 2012; Kirkeboen et al., 2016).

4.1 Gender Differences in Returns and Match Effects

Tables 2 and 3 report gender-specific pecuniary and non-pecuniary returns and match effects estimates among the pool of accepted college applicants. Estimates come from the following empirical analog of model (7):

$$\begin{aligned} Y_i = &\sum_j \alpha_j d_{ij} + \sum_j \alpha_j^G d_{ij} G_i + \sum_j \psi_j \hat{\lambda}_{ij} + \sum_j \psi_j^G \hat{\lambda}_{ij} G_i \\ &+ \sum_j \psi_j^* \hat{\lambda}_{ij} d_{ij} + \sum_j \psi_j^{*G} \hat{\lambda}_{ij} d_{ij} G_i + \beta \mathbf{X}_i + \mu_t + \phi_{c(i)} + \varepsilon_{ij}, \end{aligned} \tag{9}$$

where we consider two outcomes: log earnings and an indicator for having at least one child twelve years after applying to college. The variable $d_{ij} \equiv \mathbb{1}(d_i = j)$ is an indicator that equals 1 if individual i completed field j, and 0 otherwise; and G_i is another indicator variable that equals one if an individual i is female. $\hat{\lambda}_{ij}$ is the control function proxying for individual's i preferences for field of study j and $\mathbf{X}_i$ are quadratic polynomials of the college admission PSU score in mathematics and language. We include cohort (i.e., year of application) μ_t and cell $\phi_{c(i)}$ fixed effects, where a cell is defined by two school types (public and private), 3 macro-regions (north, center, and south), and 3 financial aid-relevant PSU ranges ($<$450, $\in$ (450,550), $>$550). To ease interpretation, we standardize $\hat{\lambda}_{ij}$ within a cell.

We use a parametric bootstrapping procedure to account for estimation errors in the estimated control functions $\hat{\lambda}_{ij}$. Concretely, for each cell c, we have an estimate of the asymptotic distribution of $(\delta_{cj}, \lambda_{dc})$, which we can draw from in each bootstrap iteration b to construct different control functions. Thus, we estimate the model (9) using ordinary least squares and the bootstrapped control function $\hat{\lambda}_{ik}^b$. We report the mean estimate across 100 bootstrap iterations and the mean of the standard errors for each of the model's parameters.

Pecuniary impacts: Table 2 reports major effects on log earnings relative to not graduating among the pool of accepted college applicants. Pursuing and completing a college education is a good investment in Chile; the average return (weighted by the number of graduates in each field of study) is 43 log points for women and 51 for men, with sizable differences across fields of study. All fields of study, except for Humanities, exhibit positive returns ranging from 17 to 80 log points. Medicine, Business, and Engineering are the three fields of study with the largest returns, ranging from 55 to 80 log points. The noise-adjusted standard deviation in mean returns for women is 23 log points and 28 for men.[12] This estimated dispersion in returns mirrors similar patterns found in Altonji et al. (2012) in that the heterogeneity in returns rivals the college premium.[13]

[12] We calculate the noise-adjusted variance using the formula: $\frac{1}{J}\sum_{j=1}^{J}\left((\hat{\alpha}_j - \mu_\alpha)^2 + (SE(\hat{\alpha}_j))^2\right)$, where μ_α is the overall mean across field returns $\hat{\alpha}_j$ and $SE(\hat{\alpha}_j)$ is the estimation error.

[13] Appendix Table B.3 reports major effects on total earnings, accounting for potential extensive margin impacts across majors. The results are qualitatively similar. To assess the potential influence of the top-coding on our results, we report estimates using the AFC (private sector) data in Table B.2; this data has a higher censoring limit but it is restricted to only the private sector. The results are also qualitatively similar. Medicine is the most-impacted field of study. One explanation for the differences is the large role the public sector plays as an employer for doctors. Many doctors practice at both public and private institutions, so estimates using AFC data will only capture a portion of their yearly earnings.

We estimate gender differences in returns across all fields of study, which vary but are mostly modest. Among these, men exhibit larger returns in Science and Engineering that are statistically significant. Male returns to Science degrees are roughly 20 log points larger than Female returns to Science, and men have a roughly 7 log points larger return to Engineering. In contrast, estimated returns are nominally higher for women than men in Social Science, Teaching, Humanities, Health, and Business, fields with relatively lower returns. Except for Business, female representation is larger in fields where they have nominally higher estimated returns, and only in Teaching is their larger return statistically different from that of men. Overall, gender differences in mean returns are modest.

Gender differences in average returns inform about one dimension of the gap, but match effects can dilute or amplify the gender gaps. We find strong evidence of selection on gains in that students with larger estimated preferences for a given major—as captured by the control functions—also experience larger returns from that major compared to others with lower estimated preference intensity. We can strongly reject the null hypothesis that men and women do not sort on gains. This evidence is consistent with the growing body of work finding selection on potential earnings gains in higher education settings (Kirkeboen et al., 2016; Otero et al., 2021). We find limited evidence of gender differences in match effects; only one of eight pairs of match effects is statistically different between men and women.

In Appendix B, we further quantify the contribution of differences in mean returns and match effects to the within-major gender gap (Appendix Figure B.1 and assess the importance of gender differences in firm sorting using an AKM decomposition of earnings (Appendix Table B.1. Overall, we find that gender differences in match effects contribute little to within-major gaps and gender differences in firm sorting are also not too economically meaningful.

Non-pecuniary impacts: Table 3 reports the effects of college majors on fertility, measured as the likelihood of having at least one child twelve years after applying to college. The overall mean impacts of majors on fertility differ substantially between men and women. For women, completing college tends to delay fertility, with a mean impact of -0.03, while men experience a larger negative impact on fertility, with a mean effect of -0.11. This indicates that major choice has a more pronounced effect on reducing fertility for men compared to women.[14]

The dispersion in fertility impacts across fields of study is also notable. For women, the noise-adjusted standard deviation of fertility impacts is 0.06, suggesting that the variation across majors is twice as large as the mean impact. This highlights the significant influence that major choice has on fertility decisions later in life. For men, the dispersion is smaller, with a noise-adjusted standard deviation of 0.04, yet it still reflects meaningful variability across different fields of study.

We find no compelling evidence that selection on gains or match effects meaningfully influence fertility outcomes. This suggests that differences in preference intensity for fields of study do not systematically correlate with the fertility treatment effects observed across different majors. This lack of significant match effects is consistent across both genders, indicating that the observed fertility impacts are not driven by selection on gains. It is worth noting, however, that

[14] As shown by Appendix Table B.4, we find similar results if we look at fertility also on the intensive margin, using the number of children as the dependent variable.

we assess match effects based on college application preferences, a stage at which there can be more bias regarding fertility than earning impacts, i.e., motherhood might be an information shock to beliefs Kuziemko et al. (2018).

Finally, the gender differences in mean fertility impacts are large, with men experiencing more substantial negative impacts than women across most fields of study, a common finding insofar as males have children at relatively older ages (Wang et al., 2023). The fact that Teaching and Health are the fields most associated with fertility while Law and Humanities are more associated with childlessness is consistent with the patterns found in Europe and the United States (e.g., Hoem et al., 2006; Michelmore and Musick, 2014). Since a large share of jobs in Teaching and Health are provided by the public sector, this is also consistent with Goldin and Katz (2008) observations that: "women in careers with the greatest predictability and the smallest financial penalty for time out have the most children". These findings underscore the substantial role that major choice plays in shaping fertility outcomes, with clear gender differences and significant variation across fields of study.

4.2 Testing for bias in pecuniary and non-pecuniary returns

The preceding evidence demonstrates striking heterogeneity in pecuniary and non-pecuniary returns to different fields of study. Our empirical approach assumes that accounting for preference heterogeneity captured by partial orderings contained on rank-ordered lists sufficiently accounts for selection bias. Other approaches would further leverage capacity constraints, isolating a multitude of quasi-experiments embedded in centralized assignment (Abdulkadiroglu et al., 2017; Abdulkadiroğlu et al., 2022). Our interest in documenting selection patterns in fields of study precludes us from adopting this approach. Nonetheless, we can leverage our rich data and setting to validate our empirical approach. Intuitively, the validation exercise asks whether the return predictions implied by our empirical approach are commensurate to returns implied by a quasi-experimental approach that uses both preference information and capacity constraints.

The core elements that create these stratified experiments are applicant types, θ_i, defined in Section 3, a vector of degree cutoffs τ, and a bandwidth δ. Abdulkadiroğlu et al. (2022) shows that the multitude of quasi-experiments embedded in centralized assignment produces major-specific propensity scores that account for both preferences and capacity constraints, which we denote by $p_{ij}(\theta_i, \tau, \delta)$. The propensity scores are defined locally in settings such as ours with non-random tiebreakers, so they depend on flexible polynomials of the program-specific running variables, $g(\theta_i, \tau, r_i)$, where r_i is notation for program-specific running variables which are a function of entrance exam scores, R_i, that are an element of θ_i. Last, because not everyone applies to every program, let a_{im} denote the application to a program m indicator. For expositional purposes, let $f_s(p_{ij}, a_{im}, g_r(\theta_i, \tau, r_i, \delta)) = \sum_m \gamma_{mak} a_{im} + \sum_j \gamma_{jpk} p_{ij}(\theta_i, \tau, r_i) + g_r(\theta_i, \tau, r_i, \delta)$ for $s = 1, 2$ summarize the function that nests the local propensity scores, running variables, and application indicators. Conditional on f_1, or equivalently, conditional on a_{im}, p_{ij}, and g_1, offers to different majors—denoted by Z_{ij}—are locally random and this is the benchmark variation we use to validate our empirical approach.

For the validation exercises, let Y_i correspond to an observed outcome of individual i and let $\hat{Y}_i$ be the predicted outcome implied by our preferred empirical approach reported in Table 2 and Table 3. The validation exercise amounts to estimating the following model via two-stage least squares:

$$\begin{aligned}\hat{Y}_i &= \sum_j \pi_j Z_{ij} + f_1(p_{ij}, a_{im}, g_1(R_i, \tau, X_i, r_o, \delta)) + e_{1i} \\ Y_i &= \psi \hat{Y}_i + f_2(p_{ij}, a_{im}, g_2(R_i, \tau, X_i, r_o, \delta)) + e_{2i}. \end{aligned} \tag{10}$$

If the return predictions implied by our preferred empirical approach are commensurate to the returns implied by the quasi-random variation, then $\psi = 1$, and our empirical approach is forecast unbiased. The overidentified nature of the validation exercise further allows for a test of the predictive validity of our approach for each field of study.

We conduct the validation exercise for both fertility and earnings models separately. Table 4 reports test results for three different models of interest. Column 1 reports test results for an uncontrolled model; Column 2 augments the model with quadratic polynomials in PSU scores, cell fixed effects, and cohort fixed effects; Column 3 adds the control functions. Column 1 is a useful starting point to demonstrate the prevalence of bias in models that do not adjust for observable differences that correlate with major choice. The second model contains observables available in a host of papers and serves as a useful data point to gauge the potential bias present in those papers (Altonji et al., 2012, 2016; Andrews et al., 2022). Column 3 is our preferred model corresponding to the specification used to obtain the estimates in Table 2 and Table 3.

Panel (a) shows that uncontrolled earnings models exhibit substantial bias, with a forecast coefficient of 0.72. Models that include cell and cohort fixed effects and polynomials of PSU scores, analogous to SAT scores in other papers, show an improvement with a forecast coefficient of 0.86. The forecast coefficient in Column 2 is marginally significantly different from 1. Our preferred model produces a forecast coefficient of 0.98, underscoring the importance of accounting for preference heterogeneity to improve the average predictive validity of earning impacts. This evidence bodes well for other approaches and settings that use preference data to account for selection bias (Mountjoy and Hickman, 2021; Dale and Krueger, 2002).

Panel (b) reports broadly similar findings for fertility impacts. Uncontrolled models have a substantial forecast bias, and adding rich controls attenuates the forecast bias, but to a lesser extent than in earnings models. Adding control functions that characterize selection into majors produces a forecast coefficient of 0.95 that we cannot reject equals one. The overarching conclusion from the validation exercise is that our preferred model adequately characterizes selection into majors and produces treatment effects that line up with quasi-experimental approaches leveraging locally random assignment to majors.

Finally, we validate a model that focuses solely on the relative returns to different fields of study, excluding the non-graduation margin. These analyses are restricted to the sample of college graduates, similar to Kirkeboen et al. (2016), and assess the sensitivity of our preferred estimates to potential bias from non-graduation, conditional on student type. Appendix Tables B.5 and B.6 present the estimates obtained when the sample is limited to graduates and the omitted field of study is Health. Reassuringly, the relative returns are consistent with those

presented in Table 2 and Table 3. Appendix Table B.7 presents the validation results for these alternative models, showing outcomes that align qualitatively with our preferred models: both pecuniary and non-pecuniary impacts are forecast unbiased.

4.3 Gender Differences in Preferences for Family Considerations and their Contribution to the Earnings Gap

While the connection between majors and earnings follows from canonical human capital models that view education as an investment with pecuniary returns (Becker, 1964), the connection between major choice and fertility is more subtle. Among a sample of "high-ability" undergraduates in the US, Wiswall and Zafar (2021) show that males and females have varying beliefs about future fertility and provide evidence that fertility and earnings are important determinants of choices. We provide an alternative perspective by empirically assessing the relationship between preferences and ex-post earnings and fertility returns across a nationally representative sample in Chile. Through the lens of the choice model, we relate cell-specific mean-utilities with cell-specific returns:

$$\delta_{jc} = \rho^E \bar{Y}^E_{jc} + \rho^F \bar{Y}^F_{jc} + e_{jc}, \tag{11}$$

where the parameters ρ^E and ρ^F yield insights about the importance of earnings and fertility in shaping major choice decisions. To proxy for $\bar{Y}^E_{jc}$ and $\bar{Y}^F_{jc}$ we consider both the mean earnings and fertility of major j graduates in each cell c (residualized with respect to test score polynomials) and the causal estimates of earnings and fertility of major j graduate, from our main model (9), estimated at the cell level.[15]

Table 5 reports estimates of Equation (11). Columns 1-3 report the results obtained when considering the average earnings and fertility by cell and major. Columns 4-6 show the results obtained using the causal impacts on earnings and fertility, estimated by cell. Panel (a) shows that cell-and-major-level mean earnings and fertility independently predict major choice when we pool across genders. Specifically, Column (1) indicates that a one standard deviation increase in earnings corresponds to a 0.28σ rise in mean utility. Likewise, Column (2) reveals that a one standard deviation increase in fertility leads to a 0.21σ increase in mean utility. Column (3) demonstrates that both earnings and fertility remain predictive of choices even when controlling for each other, suggesting that applicants' demand is similarly elastic with respect to both factors. Columns (4) to (6) report quantitatively and qualitatively similar results when considering the cell-and-major-specific causal impacts on earnings and fertility.

Panel (b) explores gender differences in preferences, revealing significant heterogeneity. Demand of both men and women is predicted by earnings and family considerations, but for women, earnings are less predictive, and family considerations are more predictive. This pattern is consistent across both univariate and multivariate preference models. For example, Column (6) indicates that, for women, a one standard deviation increase in causal impacts on fertility, holding earnings constant, is associated with a 0.26σ increase in mean utility. Conversely, when fertility is held constant, a one standard deviation increase in earning impacts are associated with only a 0.16σ increase in mean utility. In contrast, earning impacts (holding

[15] We also control for the average cost of each field of study in this specification.

fertility impacts constant) are associated with a 0.35σ increase in mean utility for men, and their demand is not predicted by family considerations once earnings are held constant. These results suggest that the differing weight placed on family considerations by men and women may significantly contribute to gender disparities.

We perform a simple counterfactual exercise to assess how preferences for family considerations relative to earnings affect the gender earnings gap. For this, we first equalize the importance assigned to family considerations across men and women, i.e., we assign the same importance of fertility impacts $\hat{\rho}^F$ to men and women. Then, we leverage our choice model to predict the shares of individuals of gender g that would apply to each field j (s_j^g) under this counterfactual assignment,[16] and estimate the implied gender gap using Equation (1).

Appendix Figures B.3 and B.4 report the estimates from this exercise. Appendix Figure B.3 shows that equalizing family considerations between genders—while holding other preferences constant—can reduce the earnings gap by 17-24 percent. Notably, assigning male preferences to everyone has a slightly greater effect, largely because it minimizes the influence of family considerations on choice behavior. Appendix Figure B.4 breaks down the changes observed in Appendix Figure B.3. Compared to Figure 2, equalizing preferences for family considerations significantly reduces the contribution of Engineering programs to the overall gender earnings gap while diminishing the offsetting impact of Teaching and Health programs. Nonetheless, differences in representation across fields of study remain the primary driver of the gap in each counterfactual scenario.

Naturally, since fertility impacts might be correlated with other factors, we consider this evidence to be suggestive. Nonetheless, it underscores how a potential gender-conforming channel can contribute to the gender earnings gap by affecting pre-labor market human capital investments.

4.4 The Child Penalty Gap Across Majors

Having established that majors causally impact earnings and fertility, how do they further interact in the labor market? The relationship between these factors is complicated in that changes in earnings have income and substitution effects on fertility. A higher income makes it more feasible to have children but also raises the opportunity cost of having children (Galor and Weil, 1996; Lovenheim and Mumford, 2013). Differing incomes also affect household bargaining decisions (e.g., Chiappori and Mazzocco, 2017). All of this is further compounded by how gender norms affect general labor market behavior and within-household decisions (Bertrand et al., 2015; Akerlof and Kranton, 2000; Jayachandran et al., 2023). Simply put, the relationship is complicated.

We shed light on this issue with descriptive evidence focusing on the child penalty gap (Kleven et al., 2019b,a, 2021, 2023), a summary statistic that encapsulates the aforementioned forces without taking a strong stance on how they interact. We estimate child penalties for college graduates majoring in the nine fields considered throughout the paper. Specifically, we focus on men and women who had their first child after their graduation year, and following

[16] We compute the counterfactual shares s_j^g as $\exp(\hat{\delta}_{jc}) / \sum_c \sum_j \exp(\hat{\delta}_{jc})$, where $\hat{\delta}_{jc}$ is the mean utility predicted by model (11) under the counterfactual assignment of fertility weights $\hat{\rho}^F$.

Kleven et al. (2019b), we estimate the impacts of children on women and men separately to construct the child penalty as the percentage by which women fall behind men due to children at a given point in time. We have a more limited window of coverage compared to existing studies (Kleven et al., 2019a) and can estimate child penalties only up to 4 years after the birth of a child, so our estimates are relatively short-run. See Appendix C for the estimation equation and additional details.

Figure 3 reports child penalty estimates for each field of study. There is a vast amount of heterogeneity, with Science and Social Science exhibiting the largest penalties (36% and 32% percent, respectively) and Business and Medicine exhibiting the lowest child penalties. The average penalty across majors is a 16% reduction in earnings for women relative to men after the birth of their first child, with a noise-adjusted standard deviation of 10%. Although the heterogeneity we document alludes to an empirical interaction between major choice, gender, earnings, and fertility, the evidence reported in Figure 3 is by no means causal.[17]

5 The Role of Gender Norms

What contributes to the various causal and descriptive gender gaps we documented in the previous section? A growing body of work emphasizes the prescriptive role of norms influencing economically meaningful decisions (Oh, 2023; Bursztyn et al., 2017; Bertrand, 2011, 2020; Akerlof and Kranton, 2000). To assess the potential role of gender norms affecting the various gender gaps we document, we first construct measures of major-specific gender norms associated with the notion that "a man should earn more than his wife" (Bertrand et al., 2015). We then test for an association between major-specific gender norms and major-specific return and child penalty gaps, a prediction implied by standard labor supply models with gender differences in norms (Andresen and Nix, 2022).

5.1 Variation in Gender Norms

To study the extent to which gender norms can account for the gaps already documented, we follow the work by Bertrand et al. (2015) and construct a measure of major-specific gender norms associated with the distribution of relative labor earnings within households. We hypothesize that there is variation across fields of study, but we do not take a stance on whether different degree programs causally influence the norms or whether different majors simply attract students with different underlying gender attitudes.

Figure 4 depicts, for each of the nine fields of study, the frequency distribution of relative earnings grouped in 20 bins, along with a lowess (locally weighted scatterplot smoothing) estimate of the distribution on each side of 0.5, the point where the woman starts to earn more than the man.[18] In most fields of study, the distribution of relative earnings exhibits a sharp

[17] An econometric concern is that we observe child penalties for a selected sample of parents who choose to have children earlier. Melentyeva and Riedel (2023) finds that age heterogeneity matters but that child penalties are largest for younger mothers as they are more likely to disconnect from the labor market in years with steep career growth. This evidence suggests that our estimates may be larger than for the typical parents.

[18] This figure considers couple-level observations of female graduates. Appendix Figure C.2 depicts the plots when considering couple-level observations of male graduates. The results are qualitatively similar.

drop at 0.5, with the largest drops in Law, Engineering, Science, and Medicine and the smallest in Teaching, Health, Social Science, and Humanities. Appendix Table C.2 presents the OLS estimates associated with the discontinuity indicators obtained from separate regressions for each of our nine fields of study.

What drives the variation in the estimated cliffs at 0.5? Besides gender norms, a few other features could contribute to the variation in the estimated cliffs. First, majors offering students a menu of jobs with more flexible and less non-linear pay schedules allow for choices that generate a smoother within-household earnings distribution (Goldin, 2024). From that perspective, the estimated cliffs are not a product of norms but a consequence of differences in major-specific job access. To partly gauge the extent of that, we corroborate our norm estimates in publicly available survey data, where we can further probe the sensitivity of the estimates to differences in labor hours. As shown by Appendix Figure C.3, we find broadly similar norm estimates among the subset of couples whose hours worked per week are 40 or more. The patterns in the estimated cliffs using survey data remain broadly the same, with Social Science and Teaching having smaller cliffs than other majors but overall mirroring the estimates in Figure 4.[19] Second, some majors have closer ties to the public sector and variation in public sector access could contribute to variation in the cliffs. We do not find evidence of this as Health, Medicine, and Teaching all provide strong access to the public sector, and we find sizable differences in cliffs across these fields of study. Third, some scholars have noted that cliff estimates may be sensitive to bunching at 0.5, casting doubt on its interpretation as variation in gender norms (Binder and Lam, 2022; Hederos and Stenberg, 2022; Zinovyeva and Tverdostup, 2021). In Appendix C we show that few couples (7%) have equal earnings in our administrative records and that the estimated cliffs remain similar after excluding couples with equal earnings. We, therefore, argue that the variation in the estimated cliffs can be interpreted as variation in breadwinner norms across fields of study.

5.2 Variation in Gender Norms Predicts Gender Gaps

To what extent should variation in gender norms predict variation in the causal and descriptive gender gaps we document? A model where major-specific gender norms induce distortions in women's disutility of work relative to men results in correlations that predict variation in the earnings gap across fields of study. Similarly, a norm-based additional labor supply distortion induced by the birth of a child can lead to relative differences in changes in earnings following the birth of a child (Andresen and Nix, 2022). Therefore, one may expect earnings and child penalty gaps to be smaller in fields of study with more progressive norms, majors where the potential labor supply distortions are the smallest.

In Figure 5, we show that gender gaps in pecuniary and non-pecuniary returns are larger in fields of study that exhibit more support for the norm that "a man should earn more than his wife." A one-standard deviation increase in the progressivity of the gendered breadwinner norm is associated with a 0.87 standard deviation improvement in the earnings gender gap and a 1.09 standard deviation improvement in the fertility gender gap. We also find a correlation

[19]The CASEN survey groups Health and Medicine into a single field of study, so we cannot estimate the cliffs separately for Health and Medicine.

between child penalties and gender norms. A one-standard deviation increase in the gendered breadwinner norm is associated with a 0.54 standard deviation improvement in the child penalty. The latter correlation is weaker than the previous ones and not statistically significant. Still, it is consistent with the work of Kleven (2022) showing that estimated child penalties are smaller in locations with more progressive gender norms, measured from survey responses. Overall, we found suggestive evidence that variation in gender norms and their potentially prescriptive nature plays an important role in explaining variation in economically important gender gaps that vary across majors.

Taking stock of our results: Our descriptive and causal analysis thus far reveals four facts that inform our final counterfactual analysis. First, roughly four-fifths of the overall gender pay gap among college graduates is due to representation. Second, of the remaining fifth, we do not find that gender differences in vertical or horizontal returns contribute significantly to the gender earnings gap. Third, we find that fertility impacts vary across majors and that men and women differentially prioritize these impacts when choosing their fields of study, explaining a sizable share of the gender earnings gap. Fourth, gender norms vary across fields of study and mediate policy-relevant gaps gender gaps.

These findings suggest two avenues to potentially narrow the earnings gap among college graduates. The first, related to the latter two points, relates to the daunting task of changing preferences. While ongoing work explores this possibility indirectly (Del Carpio and Guadalupe, 2022), that is outside the scope of our paper. Instead, motivated by the fact that the representation gap explains 82% of the overall gap, reallocative policies that shuffle students across different fields—while keeping their preferences constant—are an appealing and feasible approach in settings with centralized assignments. This approach allows us to take preferences as given and asks how much progress can be made if we make slight and feasible alterations to the rules of the assignment process.

6 Counterfactual Analysis

In this section, we use the estimated returns to different fields of study by gender and the nature of the admission system to evaluate the impact of different policies on the expected gender earnings gap. In light of our motivating evidence in Figure 2 showing that representation in Engineering is key for the overall gender pay gap, we consider the impact of different capacity expansion and quota policies in Engineering. We consider the gender-neutral policy of expanding all Engineering programs by 10%, 30%, or 50%. Similarly, we consider keeping capacities constant but reserving 10%, 30%, or 50% of seats for female students. Related to gender quotas, Bursztyn et al. (2023) finds that 74 percent of Chilean citizens support affirmative action policies.[20] Capacity expansions are a gender-neutral policy whose potential disparate impacts depend on the nature of preferences among the applicant pool. Quotas, conversely, target certain groups by design and cause displacement of untargeted groups. Both policies

[20]Indeed, in the most recent admission process, exclusive places were given for women in careers related to science and engineering. This measure aims at reducing the gender gap in Chile, a country that ranks as the fourth with the least women graduates in STEM among OECD countries (MINEDUC, 2023, La Tercera, 2023).

have potential efficiency implications but through different channels that we assess. Next, we define the counterfactual objects of interest.

6.1 Conceptual Framework

We begin with the population of college applicants $\mathcal{I}$ and simulate student assignments under alternative capacity and quota policies. The inputs of the simulation analysis are student type $\theta \in \Theta$, college programs' capacities ($\mathbf{q}$), priorities (ω), and a matching algorithm, i.e., a function that takes as inputs the previous elements and outputs a vector of student assignments (ψ). As presented in the conceptual framework, students' type θ are summarized by rank order lists ($\mathbf{R}$) and program-specific scores ($\mathbf{s}$). Changes in the policies we consider correspond to a change in $\mathbf{q}$ for gender-neutral seat expansions or a change in ω for gender quotas, both of which produce a change in the allocation of students ψ.[21]

Students' assigned field is indexed by $j = 0, 1, 2, \cdots, 9$, where $j = 0$ corresponds to no assignment. Potential outcomes defined in Equation (2) are Y_{ij}, where $j = 0$ corresponds to the non-graduation outcome. At the time they are assigned, a student's graduation outcome is not yet observed. We, therefore, define students' expected earnings conditional on assignment j as:

$$\tilde{Y}_{ij} = \begin{cases} p_{ij}Y_{ij} + (1 - p_{ij})Y_{i0} & \text{if } j > 0 \\ Y_{i00} & \text{if } j = 0 \end{cases} \tag{12}$$

where p_{ij} corresponds to the graduation probability in field j if assigned to field j, Y_{ij} is the potential outcome in the state where individual i graduates from j, Y_{i0} is the potential outcome in the state where the assigned individual does not obtain a degree, and Y_{i00} is the potential outcome of the unassigned who also do not graduate from college so that $p_{i0} = 0$.[22]

Policy changes generate changes in assignments to majors that generate counterfactual expected earnings. For a given allocation ψ, student i's assignment is $j(\psi_i)$. The change in expected earnings a student experiences induced by a change from ψ to ψ' is:

$$\Delta_i(\psi, \psi') \equiv \tilde{Y}_{ij(\psi_i')} - \tilde{Y}_{ij(\psi_i)}. \tag{13}$$

We can decompose the change in earnings as:

$$\Delta_i(\psi, \psi') = \underbrace{\left(p_{ij(\psi_i')} - p_{ij(\psi_i)}\right)}_{\text{Change in graduation prob.}} \times \underbrace{\left(Y_{ij(\psi_i')} - Y_{i0}\right)}_{\text{Return to } j} + p_{ij(\psi_i)} \underbrace{\left(Y_{ij(\psi_i')} - Y_{ij(\psi_i)}\right)}_{\text{Change in relative return}}. \tag{14}$$

Equation (14) highlights two key forces that may work in opposite directions. The first term reflects the combination of the change in a student's probability of graduating when transitioning

[21] When introducing quotas, we modify the matching algorithm to consider priority seats, processing open seats first and then reserved seats. See Dur et al. (2018) for the importance of precedence in the processing of reserved seats.

[22] Note that there is a compliance assumption that follows from Assumption 2. In particular, in a setting without centralized assignment, it is likely that a given individual that is accepted to field j can obtain a degree in field k. This would mean that for a given assignment, there may be several non-degenerate p_{ijk} where p_{ijk} corresponds to the probability of graduating from field k if assigned to field j. Assumption 2 rules out this kind of behavior, so $p_{ijk} = 0$ for $j \neq k$.

between assignments and the causal return of their new assignment relative to non-graduation. The second term captures the change in the economic return associated with the new field of study relative to their initial assignment. If a student is reassigned to a field with higher earning potential, this second term will be positive, reflecting an increase in expected earnings. However, if the reassignment results in a lower probability of graduation—due, for example, to mismatch effects (Arcidiacono and Lovenheim, 2016; Arcidiacono et al., 2016)—the first term will be negative, offsetting some or all of the gains from the higher potential earnings.[23] Understanding the empirical magnitude of these two forces is essential for evaluating the impact of reallocative policies on student outcomes.

Let $\tilde{Y}_i(\psi) = \sum_j D_{ij}\tilde{Y}_{ij}$ represent the expected earnings of student i under allocation ψ, where $D_{ij} = \mathbf{1}\{j(\psi_i) = j\}$ indicates whether student i is assigned to major j, and $\sum_j D_{ij} = 1$ due to the single-offer system. The ex-ante gender earnings gap under allocation ψ is defined as:

$$\Delta(\psi) \equiv E[\tilde{Y}_i(\psi) \mid F] - E[\tilde{Y}_i(\psi) \mid M],$$

where the expectation is taken over the female (F) and male (M) populations.

When the allocation changes from ψ to ψ', the corresponding change in the gender earnings gap is given by:

$$\begin{aligned} \Delta\Delta(\psi, \psi') &\equiv \Delta(\psi') - \Delta(\psi) \\ &= E[\Delta_i(\psi, \psi') \mid F] - E[\Delta_i(\psi, \psi') \mid M]. \end{aligned} \tag{15}$$

Equation (15) expresses the change in the earnings gap in terms of the decomposition introduced in Equation (14). This allows us to attribute the overall change in the gender earnings gap to the two components mentioned above, specifically the relative changes in graduation probabilities and shifts in major-specific earnings potential, as the allocation changes from ψ to ψ'. Equation (15) is a key counterfactual object of interest, with elements of Equation (14) further decomposing the sources of changes in the gender earnings gap.

Last, to summarize the efficiency implications of changes in allocations, we consider the aggregate change in expected earnings relative to the aggregate expected earnings in the baseline counterfactual. We define the change in efficiency relative to the baseline aggregate expected earnings as:

$$\mathcal{E}(\psi', \psi) = \frac{\sum_i \Delta_i(\psi, \psi')}{\sum_i \tilde{Y}(\psi)}. \tag{16}$$

Equation (16) summarizes the relative distortion induced by the change in the allocation. A positive value of $\mathcal{E}(\psi', \psi)$ indicates that ψ' leads to higher aggregate expected earnings than ψ, while a negative value implies the opposite. The normalization by baseline aggregate earnings quantifies the changes in terms of the baseline allocation.

[23] The empirical evidence on mismatch effects is mixed. In the University of California system, Arcidiacono et al. (2016) find that less prepared minority students at top-ranked campuses would have higher graduation rates had they attended lower-ranked campuses, and Bleemer (2023) finds that minority students experienced worse outcomes as they transitioned into lower tier universities following the ban of California's Proposition 209. In Latin America, Otero et al. (2021) do not find strong evidence of mismatch effects in Brazil, while Carlana et al. (2024) find evidence of mismatch effects for marginal students but not the average student in Chile.

6.2 Empirical Methods

To summarize and decompose changes in the gender earnings gap under different allocations, we need estimates of Y_{ij}, Y_{i00}, and p_{ij} for all i and j. The potential outcome model discussed in Section 3 serves as the starting point to define Y_{ij} among graduates, and we use an auxiliary model to predict Y_{i00}. The p_{ij}, which are instrumental in understanding the potential mismatch effects, are estimated using a random forest model. We discuss each now in turn.

We use estimates of Equation (9) to construct potential earnings $\hat{Y}_{ij}(\psi)$ for each individual that is assigned to a field j in allocation ψ. The causal model is centered on the potential earnings for those accepted (assigned) to a field of study. However, for unassigned students—those who do not receive a college offer—we need to construct their potential earnings, $\hat{Y}_{i00}$. For this purpose, we estimate our causal model using the same rich vector of attributes that are part of our estimating Equation 9 but in a sample of unassigned students. To summarize, the estimates of the causal model provide us with potential earnings for individuals assigned under allocation ψ, and the auxiliary prediction model for Y_{i00} provides us with remaining potential earnings for individuals that are unassigned under allocation ψ.

To estimate the graduation probabilities p_{ij}, we train random forest models tailored to each major. The models incorporate various features, including demographic characteristics, entrance exam scores, preferences revealed through rank-ordered lists and all other relevant attributes that are also part of Equation (9). This approach allows us to capture complex, non-linear relationships between student characteristics and their likelihood of graduating from a given major. To improve model performance, we fine-tune the parameters for each major separately, optimizing them to minimize the out-of-bag (OOB) Brier score. The Brier score, which measures the accuracy of probabilistic predictions, is particularly useful in this context as it reflects both the calibration and sharpness of our predictions. By minimizing the OOB-Brier score, we ensure that our random forest model provides reliable and well-calibrated predictions of graduation probabilities across different majors. Importantly, we find that our estimates $\hat{p}_{ij}$ are forecast unbiased. Appendix D reports additional details and summary statistics for this prediction exercise.

Equipped with estimates of Y_{ij}, Y_{i00}, and p_{ij}, we can construct expected earnings at baseline and in counterfactual scenarios. We account for estimation error and population variation using a bootstrapping procedure. To account for estimation error in the parameters governing $\hat{Y}_{ij}$, we use the parametric bootstrap, sampling from the asymptotic distribution across bootstrap iterations. To account for population uncertainty, we sample applicants with replacement to create different bootstrapped samples. We estimate counterfactual objects of interest across applicant samples and parameter estimate samples and report median estimates along with a 95 percent coverage region.

6.3 Results

Table 6 summarizes our results for various counterfactual policies focused on the field of Engineering. The Table shows the impact of each counterfactual on the representation gap and decomposes the impact of each policy on the factors contributing to the overall gender pay gap, while also assessing the implications for efficiency (as captured by total earnings). The first row

presents the baseline assignment. The overall assignment gap between men and women applicants in Engineering is 14.3 percentage points in column (1). Columns (2) through (4) present the components of the decomposition given by Equation (15), which are relative to the baseline allocation and, therefore, not populated in the first row. Column (2) shows the gender gap in the first term in Equation (14), column (3) presents the gender gap in the second term of the same equation, and column (4) shows the overall gender gap. Column (5) presents the earnings gap between men and women relative to the average earnings of men which at baseline corresponds to 12.5%. Finally, column (6) presents the efficiency implication of the counterfactual as defined in Equation (16). The remaining rows in the table present the different alternative policies. Each number in the table presents the estimated median across bootstrapped samples of returns and students. We present percentiles 2.5 and 97.5 of each estimate across bootstrapped samples in squared brackets. Appendix D.3 presents details on the bootstrap.

The second through fourth rows demonstrate that the representation gap only worsens as we consider gender-neutral capacity expansions of 10%, 30%, and 50%, increasing to 18 percentage points for a policy that increases Engineering seats for each program by fifty percent. This is due to the fact that men exhibit stronger demand for Engineering programs and absorb a majority of the new seats. Appendix Figure A.4 contains additional details. Given the disproportionate excess demand among men, it is unsurprising that the overall gender pay gap weakly increases relative to the baseline assignment, although within the baseline confidence interval. Most of this change is driven by increases in the gender gap in returns, a feature due to men moving up from lower-paying majors to Engineering programs they failed to qualify for in the baseline allocation.

Although an across-the-board increase in Engineering seats does not do much to reduce the gender pay gap, the fifth through seventh rows of Table 6 demonstrate that gender-specific quotas can reduce the gap. We see that quota policies lead to mechanical reductions in the representation gap, reducing it between 17 and 68 percent, depending on the quota size. Despite the sizable gender gap in the causal returns to Engineering, a quota closes the gender earnings gap substantially. A policy that aims to equalize access to Engineering programs with a 50% quota reduces the gender earnings gap by 9 percent.

What are the implications on efficiency? We measure this by aggregating the earnings implied by our counterfactual for all students. Column 6 reports overall small aggregated efficiency implications as measured by Equation (16). We find minimal impacts on efficiency as captured by aggregate earnings. Rows 2 through 4 show that an increase in Engineering seats increase aggregated earnings by up to 0.59%. In turn, quota policies reduce aggregated earnings reflecting the lower return to Engineering for women. A 50% quota in Engineering reduces aggregated earnings by 0.15%.

7 Conclusion

This paper studies the major choice decisions of young adults and the various ways they affect numerous policy-relevant gender gaps. We emphasize the relationship between majors, earnings,

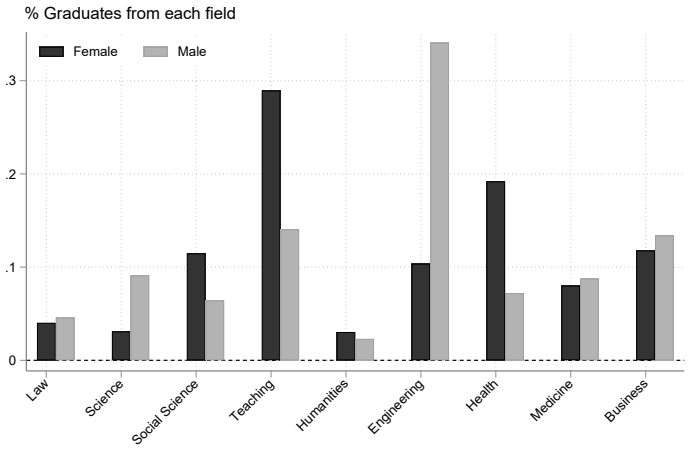
and family considerations. This allows us to elucidate important factors that generate rifts in the educational trajectories of men and women and serve as a precursor for different outcomes.

We document sizable gender differences in how field of study impacts earnings and fertility, with evidence of Roy (1951)-like selection only for monetary returns. Having established that majors affect earnings and fertility outcomes, we then show that men and women differentially prioritize these outcomes, contributing to gender differences in representation across fields of study. Equalizing preferences for family considerations associated with fields of study across men and women reduces the gender earnings gap by 27 percent. To better understand factors contributing to the observed gender differences, we consider the role of gender norms summarized by breadwinner norms. We find that differences in gender norm progressivity across fields of study predict return and child-penalty gaps, suggesting that norms play an important and prescriptive role in governing behavior.

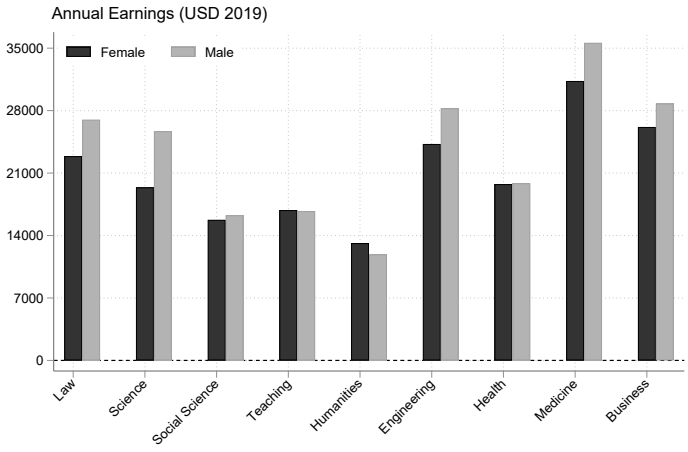
We conclude by assessing whether feasible supply-side policies in systems with centralized assignment can affect the broader gender earnings gap. We find that gender-specific quotas reduce the gender earnings gap among college graduates, while gender-neutral expansions of capacities exacerbate the gap. We also find minimal impacts on the implied efficiency of the centralized match as measured by changes in aggregate earnings.

Although we can quantify the extent to which the balance of career and family affects major choice and subsequently impacts several policy-relevant gaps, we cannot speak to factors that generate these persistent gender differences in preferences. On the policy side, our work does not consider behavioral responses to changes in assignment policies and is partial equilibrium in nature. Understanding the factors that generate gender differences in preferences and potential behavioral responses to affirmative action policies continues to be an important avenue for future research.

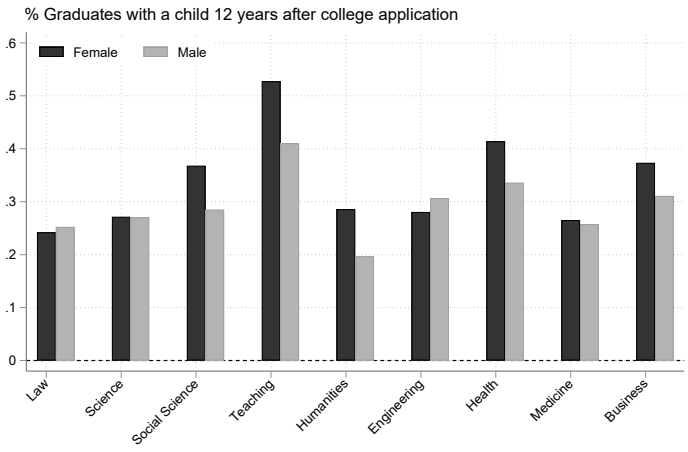
Figure 1: Graduation, Earnings,and Fertility



(a) Graduation



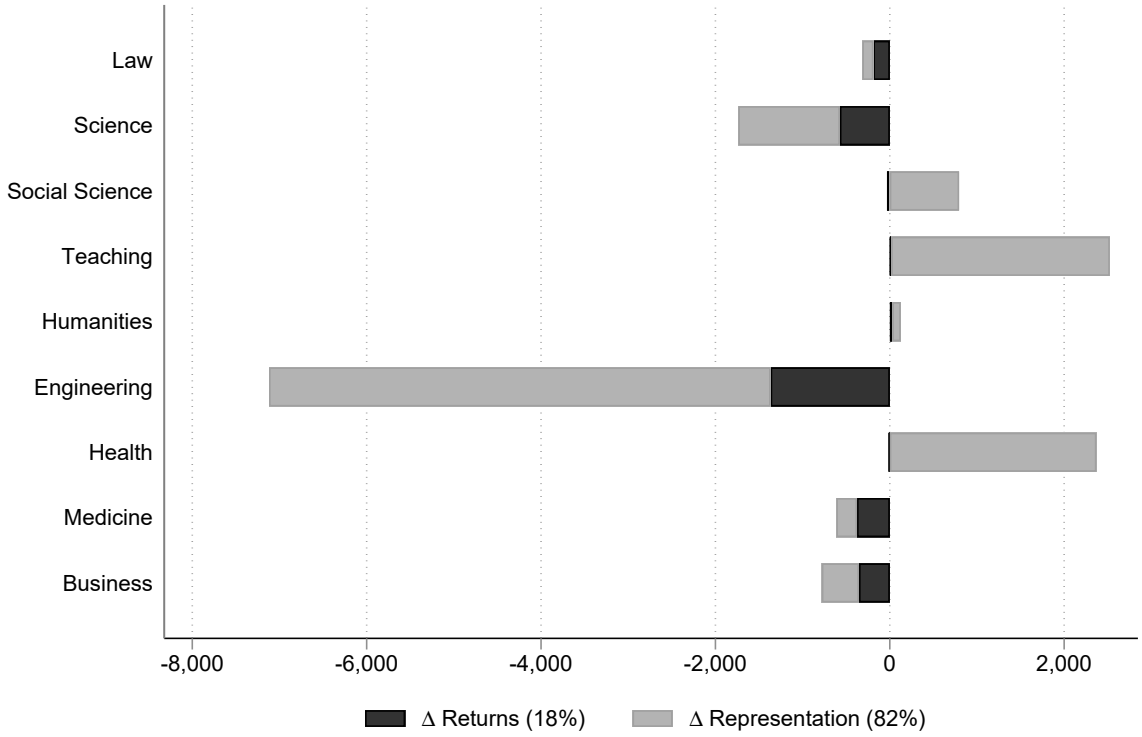
(b) Earnings



(c) Fertility

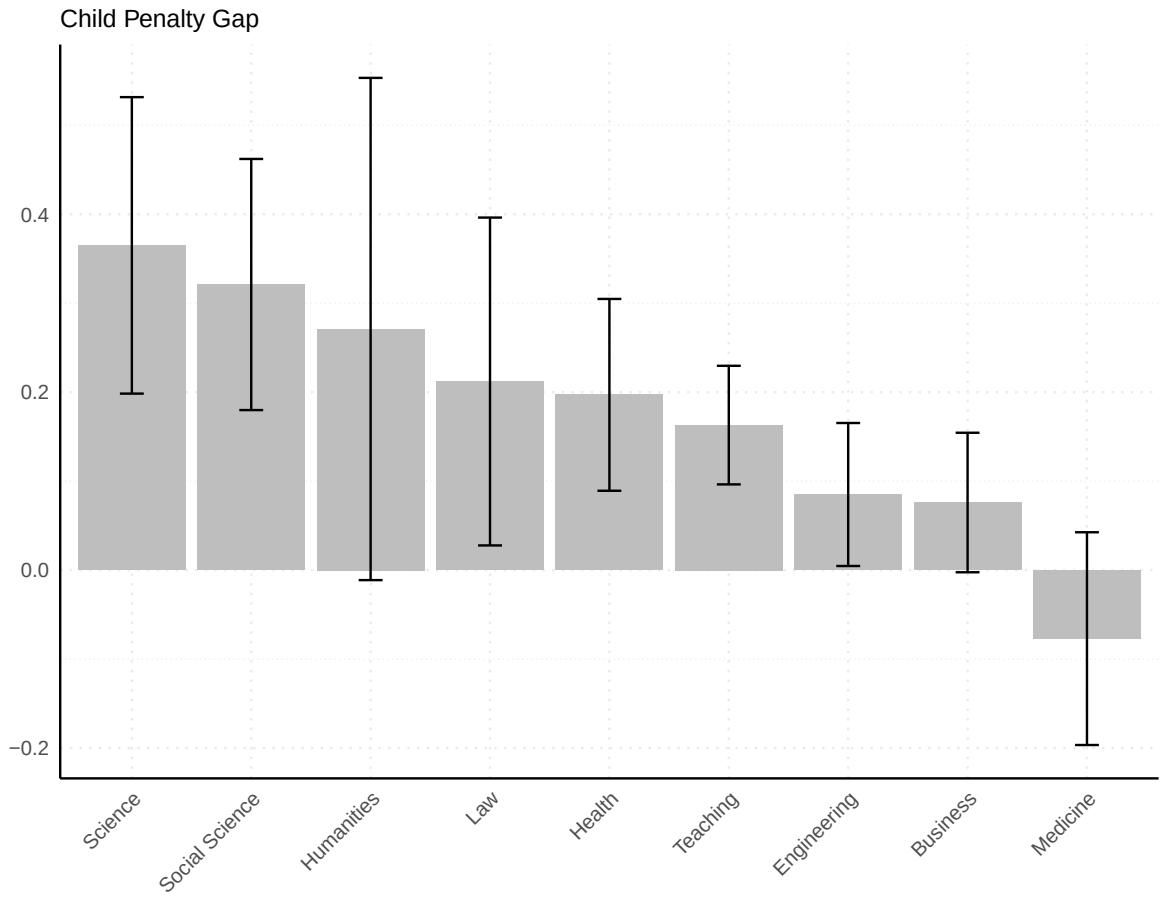
Notes: Panel (a) shows the share of male and female graduates in each field of study. Panel (b) shows the annual earnings (as of 2019) of male and female graduates in each field of study. Both panels consider students who were employed in 2019. Panel (c) shows the share of students with at least one child twelve years after their college application. All panels consider students who applied through the centralized admission system between 2004 and 2007 and graduated between 2007 and 2019.

Figure 2: Earnings Gap: Sorting versus Returns



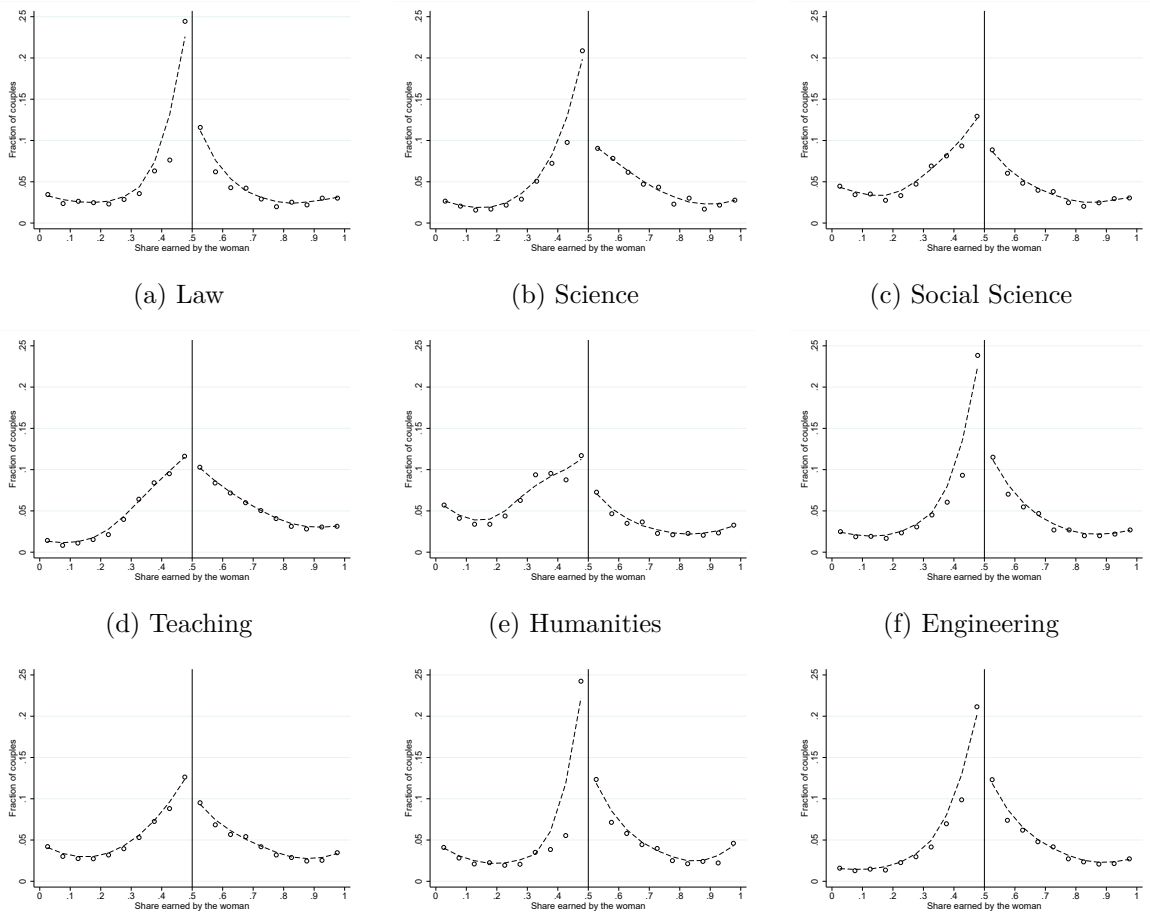
Notes: This figure shows the result from the exercise described by Equation (1) that decomposes the gender earnings gap into differences in sorting between fields of study and differences in returns within fields of study. We consider the annual earnings in 2019 of students who applied through the centralized admission system between 2004 and 2007, graduated between 2007 and 2019, and were employed in 2019.

Figure 3: Child Penalty Gaps Across Fields of Study



Notes: This figure shows the estimated child penalty gap four years after birth (and the corresponding confidence interval). We consider college graduates who had their first child between 2016 and 2017 and after graduation. Earnings are not conditional on employment status, and the effects include extensive and intensive margins. Estimation follows Kleven et al. (2019b) and for additional details, see Appendix C.

Figure 4: Gender Norms by Field of Study



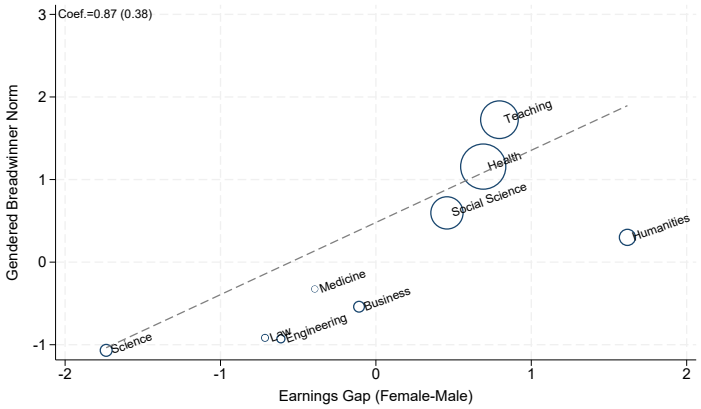
(a) Law (b) Science (c) Social Science

(d) Teaching (e) Humanities (f) Engineering

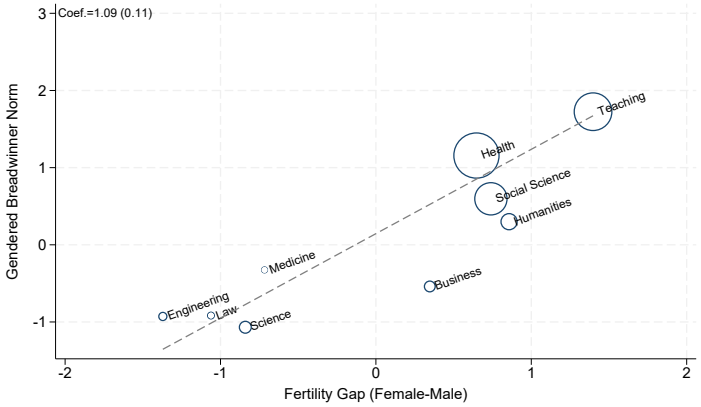
(g) Health (h) Medicine (i) Business

Notes: These figures report the relative income distribution for females who graduated from different fields of study. The sample includes couples of parents who both earn positive income. We use the observation from the third and fourth quarters before childbirth. Each dot is the fraction of couples in a 0.05 relative income bin. The vertical line indicates the relative income share=0.5. The dashed line is the lowess smoother applied to the distribution, allowing for a break at 0.5.

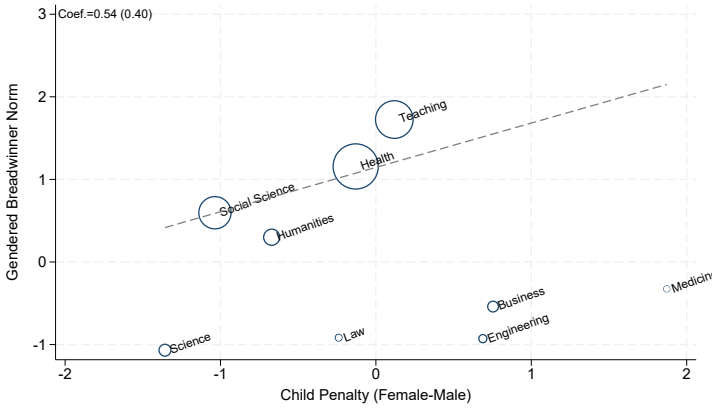
Figure 5: Gender Norms Correlations



(a) Gender Earnings Gap



(b) Gender Fertility Gap



(c) Child Penalties

Notes: These figures plot empirical Bayes-adjusted breadwinner norm estimates (i.e., "a man should earn more than his wife") by field of study against average earnings gaps in Panel (a), average fertility gaps in Panel (b), and child penalties in Panel (c). All variables are standardized to have a mean of zero and a standard deviation of one. Observations are weighted by the inverse of the variance of the breadwinner norm estimates. Estimates of a slope coefficient from a linear regression of estimated norms on each outcome are displayed in the top left of each panel, with robust standard errors in parenthesis.

Table 1: Descriptive Statistics

	Non-Grad (1)	Law (2)	Science (3)	Social Science (4)	Teaching (5)	Humanities (6)	Engineering (7)	Health (8)	Medicine (9)	Business (10)
					Panel A: Female					
Average Test Score	0.03	1.00	0.57	0.41	-0.02	0.69	0.76	0.57	1.53	0.40
Father is College Grad	0.20	0.46	0.29	0.34	0.17	0.47	0.38	0.26	0.52	0.28
Mother is College Grad	0.15	0.41	0.24	0.29	0.13	0.39	0.30	0.22	0.46	0.23
Public Health Insurance	0.59	0.37	0.52	0.47	0.62	0.36	0.45	0.53	0.32	0.50
Public School	0.29	0.14	0.24	0.21	0.33	0.12	0.15	0.25	0.10	0.20
Observations	10,477	1,697	1,307	4,833	12,184	1,280	4,365	8,085	3,374	4,961
					Panel B: Male					
Average Test Score	0.24	1.18	0.63	0.59	0.18	0.64	0.84	0.65	1.72	0.65
Father is College Grad	0.22	0.50	0.29	0.34	0.17	0.35	0.37	0.23	0.50	0.36
Mother is College Grad	0.17	0.44	0.23	0.30	0.14	0.29	0.28	0.19	0.44	0.27
Public Health Insurance	0.59	0.32	0.51	0.47	0.65	0.46	0.45	0.58	0.33	0.45
Public School	0.29	0.16	0.26	0.25	0.35	0.22	0.20	0.33	0.18	0.20
Observations	15,494	1,584	3,146	2,223	4,861	797	11,772	2,493	3,030	4,634

Notes: This table presents the average test scores and demographic characteristics of all students in our main analysis sample, by gender and field of study. Our main analysis sample includes all students who applied and were accepted into a program through the centralized admission system between 2004 and 2007 and who were employed in 2019. Student characteristics come from the forms that students must fill in before taking the college admission exam.

Table 2: Gender Differences in Pecuniary Returns: Log Earnings in 2019

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Female				
Returns	0.364	0.333	0.214	0.415	-0.056	0.554	0.347	0.779	0.606
	(0.069)	(0.053)	(0.026)	(0.010)	(0.054)	(0.027)	(0.022)	(0.036)	(0.018)
Selection on Gains	0.062	0.029	0.038	-0.004	0.081	0.011	0.059	-0.035	0.017
	(0.023)	(0.026)	(0.014)	(0.008)	(0.022)	(0.015)	(0.011)	(0.014)	(0.012)
					Panel B: Male				
Returns	0.437	0.536	0.171	0.321	-0.184	0.623	0.338	0.801	0.598
	(0.081)	(0.028)	(0.043)	(0.018)	(0.070)	(0.015)	(0.036)	(0.042)	(0.022)
Selection on Gains	0.088	0.039	0.068	0.058	0.095	0.013	0.076	0.003	0.050
	(0.027)	(0.016)	(0.019)	(0.010)	(0.025)	(0.012)	(0.017)	(0.019)	(0.010)
Est. Females = Est. Males									
Δ Returns (p-val)	0.491	0.001	0.376	0.000	0.145	0.020	0.818	0.689	0.751
Δ Match Effects (p-val)	0.458	0.730	0.193	0.000	0.674	0.926	0.401	0.100	0.028

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9). We display the averages—of coefficients, robust standard errors, and p-values—across 100 regressions, each of which uses a different set of control functions obtained after parametric bootstrap over the ranked ordered logit estimation. The sample includes 102,597 observations of students who applied and were accepted into a program through the centralized admission system between 2004 and 2007 and were employed in 2019. Panel (a) focuses on parameter estimates for Female students, and Panel (b) focuses on Male students. In each panel, we report return estimates and selection on gains parameters, also referred to as match effects in the text. Underneath each estimate in parentheses is the bootstrapped standard error. The final two rows of the table report p-values from two additional statistical tests. The Δ Returns p-values correspond to the null hypothesis that the Female and Male returns are equivalent for a given column. The Δ Match Effects p-values correspond to a similar null hypothesis that the Female and Male match effects are equivalent for a given column.

Table 3: Gender Differences in Non-Pecuniary Returns: Any child 12 years after college application

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Female				
Returns	-0.156	-0.116	-0.066	0.045	-0.094	-0.071	-0.010	-0.073	-0.054
	(0.025)	(0.023)	(0.012)	(0.008)	(0.020)	(0.015)	(0.011)	(0.018)	(0.010)
Selection on Gains	0.019	-0.006	0.006	-0.011	0.008	-0.013	-0.001	-0.002	0.022
	(0.009)	(0.011)	(0.007)	(0.005)	(0.009)	(0.008)	(0.006)	(0.007)	(0.007)
					Panel B: Male				
Returns	-0.142	-0.166	-0.148	-0.070	-0.210	-0.109	-0.107	-0.113	-0.126
	(0.026)	(0.012)	(0.017)	(0.012)	(0.021)	(0.008)	(0.016)	(0.017)	(0.011)
Selection on Gains	0.009	0.013	0.011	-0.006	0.016	0.001	0.005	0.015	0.003
	(0.009)	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)	(0.008)	(0.005)
Est. Females = Est. Males									
Δ Returns (p-val)	0.686	0.052	0.000	0.000	0.000	0.018	0.000	0.102	0.000
Δ Match Effects (p-val)	0.410	0.151	0.629	0.576	0.528	0.176	0.520	0.109	0.021

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9), with health as the omitted field of study. We display the averages—of coefficients, robust standard errors, and p-values—across 100 regressions, each of which uses a different set of control functions obtained after parametric bootstrap over the ranked ordered logit estimation. The sample includes 137,339 observations of students who applied and were accepted into a program through the centralized admission system between 2004 and 2007. Panel (a) focuses on parameter estimates for Female students and Panel (b) focuses on Male students. In each panel, we report return estimates and selection on gains parameters, also referred to as match effects in the text. Underneath each estimate in parentheses is the bootstrapped standard error. The final two rows of the table report p-values from two additional statistical tests. The Δ Returns p-values correspond to the null hypothesis that the Female and Male returns are equivalent for a given column. The Δ Match Effects p-values correspond to a similar null hypothesis that the Female and Male match effects are equivalent for a given column.

Table 4: Testing for Bias

	Panel (a): Earnings		
	Model 1	Model 2	Model 3
	(1)	(2)	(3)
Forecast Coefficient ϕ	0.717	0.856	0.984
	(0.066)	(0.078)	(0.091)
Covariates in the model:			
Controls	No	Yes	Yes
Control function	No	No	Yes
Tests:			
First-stage F- Statistic:	251	213	171
p-values			
Forecast bias ($\phi = 1$)	0.000	0.065	0.861
Overidentification	0.068	0.212	0.087
Observations	44,684	44,684	44,684
	Panel (b): Fertility		
	Model 1	Model 2	Model 3
	(1)	(2)	(3)
Forecast Coefficient ϕ	0.555	0.768	0.946
	(0.109)	(0.144)	(0.176)
Covariates in the model:			
Controls	No	Yes	Yes
Control function	No	No	Yes
Tests:			
First-stage F-statistic	471	321	239
p-values			
Forecast Bias ($\phi = 1$)	0.000	0.107	0.758
Overidentification	0.047	0.110	0.116
Observations	59,859	59,859	59,859

Notes: This table reports the results of tests for bias following Equation (10). Column 1 reports test results for an uncontrolled model; Column 2 augments the model with quadratic polynomials in PSU scores, cell fixed effects, and cohort fixed effects; Column 3 adds the control functions. The forecast bias test checks whether the forecast coefficient equals 1, and the overidentification test checks the IV model's overidentifying restrictions. Panel (a) considers logged earnings in 2019 as the outcome, and Panel (b) considers fertility (any child 12 years after college application) as the outcome.

Table 5: Preference Estimates

	Mean Utility					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel (a): Pooled Estimates					
Earnings	0.282		0.278	0.252		0.254
	(0.052)		(0.049)	(0.051)		(0.051)
Fertility		0.207	0.202		0.142	0.146
		(0.053)	(0.051)		(0.055)	(0.054)
	Panel (b): Estimates by Gender					
Earnings	0.437		0.432	0.354		0.353
	(0.062)		(0.062)	(0.073)		(0.073)
× Female	-0.336		-0.326	-0.213		-0.190
	(0.111)		(0.101)	(0.108)		(0.104)
Fertility		0.091	0.042		0.038	0.023
		(0.075)	(0.064)		(0.075)	(0.074)
× Female		0.265	0.315		0.206	0.234
		(0.107)	(0.099)		(0.108)	(0.106)
Observations	324	324	324	324	324	324

Notes: This table reports estimates from bivariate and multivariate regressions of mean utilities on pecuniary and non-pecuniary returns among graduates at the cell-major level, following Equation (11). In columns (1)-(3), we proxy mean returns with mean earnings and fertility of major j graduates for each cell c. In columns (4)-(6), we proxy mean returns with the causal estimates of earnings and fertility of major j graduates for each cell c. Cells are defined by macro-region, test score ranges, high school type, and gender. All specifications include cell fixed effects and control for the average cost of each field of study, which is standardized to be mean zero with a standard deviation equal to one. Standard errors are robust and reported in parentheses.

Table 6: Counterfactuals Admissions to Engineering

	Representation Gap	Gender gap in $(p_{ij'} - p_{ij})(Y_{ij'} - Y_{i0})$	Gender gap in $p_{ij}(Y_{ij'} - Y_{ij})$	Gender gap in change of returns	Gender Earnings Gap	Efficiency
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline	0.1428	-	-	-	0.1254	-
		-	-	-	[0.1223 , 0.1280]	-
Engineering expand capacity 10%	0.1544	3.4506	-2.4603	1.0219	0.1259	0.0016
		[2.8186 , 4.0187]	[-2.8858 , -2.0899]	[0.4466 , 1.4986]	[0.1229 , 0.1282]	[0.0015 , 0.0018]
Engineering expand capacity 30%	0.1733	8.7872	-6.4658	2.2405	0.1264	0.0041
		[7.7103 , 9.7727]	[-7.1903 , -5.8276]	[1.4648 , 3.2558]	[0.1233 , 0.1289]	[0.0038 , 0.0043]
Engineering expand capacity 50%	0.1838	11.8291	-8.7651	3.0569	0.1267	0.0059
		[10.7881 , 12.8786]	[-9.5343 , -8.0653]	[1.9445 , 4.1409]	[0.1235 , 0.1292]	[0.0056 , 0.0062]
Engineering quota 10%	0.1186	-5.1966	1.3860	-3.7908	0.1230	0.0000
		[-5.6132 , -4.6156]	[1.1243 , 1.7272]	[-4.1695 , -3.3354]	[0.1200 , 0.1254]	[-0.0001 , 0.0001]
Engineering quota 30%	0.0771	-14.3804	3.3777	-10.9573	0.1183	-0.0004
		[-15.1470 , -13.6291]	[2.8650 , 3.8744]	[-11.5181 , -10.2658]	[0.1155 , 0.1207]	[-0.0006 , -0.0002]
Engineering quota 50%	0.0464	-21.4086	4.5523	-16.7852	0.1147	-0.0015
		[-22.6974 , -20.0632]	[4.0026 , 5.4636]	[-17.7921 , -15.7491]	[0.1118 , 0.1169]	[-0.0017 , -0.0013]

Notes: This table presents the results from different counterfactual exercises that modify the admission process in Engineering either by increasing capacity or by prioritizing the admission of women using quotas (see Section 6 for details). Column (1) reports the representation gap in Engineering under different counterfactual assignments. Columns (2) and (3) report changes in the elements of $\Delta_i(\psi, \psi')$ defined in Equation (14) that contribute to the overall change, $\Delta\Delta(\psi, \psi)$, defined in Equation (15) and reported in Column (4). Column (5) reports the gender earnings gap as a share of men's average earnings: $\frac{\Delta(\psi)}{E[Y_i|M]}$. Column (6) reports the aggregate change in expected earnings across all students, our measure of efficiency. The table presents median estimates across 100 bootstrapped samples of returns and students. Brackets present the 2.5 and 97.5 percentiles across bootstrapped samples. See Appendix D.3 for details.

References

Abdulkadiroğlu, Atila and Tayfun Sönmez, "School choice: A mechanism design approach," *American economic review*, 2003, *93* (3), 729–747.

Abdulkadiroglu, Atila, Joshua D Angrist, Yusuke Narita, and Parag A Pathak, "Research design meets market design: Using centralized assignment for impact evaluation," *Econometrica*, 2017, *85* (5), 1373–1432.

Abdulkadiroğlu, Atila, Joshua D Angrist, Yusuke Narita, and Parag Pathak, "Breaking ties: Regression discontinuity design meets market design," *Econometrica*, 2022, *90* (1), 117–151.

Abdulkadiroglu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters, "Do parents value school effectiveness?," *American Economic Review*, 2020, *110* (5), 1502–39.

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Ahimbisibwe, Isaac, Adam Altjmed, Georgy Artemov, Andrés Barrios-Fernández, Aspasia Bizopoulou, Martti Kaila, Jin-Tan Liu, Rigissa Megalokonomou, José Montalbán, Christopher Neilson, Jintao Sun, Sebastián Otero, and Xiaoyang Ye, "The STEM Major Gender Gap: Evidence from Coordinated College Application Platforms Across Five Continents," June 2024. Latest Version.

Akerlof, George A and Rachel E Kranton, "Economics and identity," *The quarterly journal of economics*, 2000, *115* (3), 715–753.

Altonji, Joseph G, Erica Blom, and Costas Meghir, "Heterogeneity in human capital investments: High school curriculum, college major, and careers," *Annu. Rev. Econ.*, 2012, *4* (1), 185–223.

_ , Peter Arcidiacono, and Arnaud Maurel, "The analysis of field choice in college and graduate school: Determinants and wage effects," in "Handbook of the Economics of Education," Vol. 5, Elsevier, 2016, pp. 305–396.

Andresen, Martin Eckhoff and Emily Nix, "What causes the child penalty? Evidence from adopting and same-sex couples," *Journal of labor economics*, 2022, *40* (4), 971–1004.

Andrews, Rodney J, Scott A Imberman, Michael F Lovenheim, and Kevin M Stange, "The returns to college major choice: Average and distributional effects, career trajectories, and earnings variability," Technical Report, National Bureau of Economic Research 2022.

Angrist, Joshua D, Peter D Hull, Parag A Pathak, and Christopher R Walters, "Leveraging lotteries for school value-added: Testing and estimation," *The Quarterly Journal of Economics*, 2017, *132* (2), 871–919.

Angrist, Joshua, Peter Hull, Parag A Pathak, and Christopher Walters, "Credible school value-added with undersubscribed school lotteries," *Review of Economics and Statistics*, 2024, *106* (1), 1–19.

Arcidiacono, Peter, "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?," *Econometrica*, 2005, *73* (5), 1477–1524.

_ and Michael Lovenheim, "Affirmative action and the quality–fit trade-off," *Journal of Economic Literature*, 2016, *54* (1), 3–51.

_ , Esteban M Aucejo, and V Joseph Hotz, "University differences in the graduation of minorities in STEM fields: Evidence from California," *American Economic Review*, 2016, *106* (3), 525–562.

Barrios, Andrés, Christopher Neilson, and Seth D Zimmerman, "Elite universities and the intergenerational transmission of human and social capital," *Available at SSRN 4071712*, 2021.

Becker, Gary, "Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education," Technical Report, National Bureau of Economic Research, Inc 1964.

Bertrand, Marianne, "New perspectives on gender," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1543–1590.

_ , "Gender in the twenty-first century," in "AEA Papers and proceedings," Vol. 110 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2020, pp. 1–24.

_ **, Emir Kamenica, and Jessica Pan**, "Gender identity and relative income within households," *The Quarterly Journal of Economics*, 2015, *130* (2), 571–614.

Binder, Ariel J and David Lam, "Is there a male-breadwinner norm? the hazards of inferring preferences from marriage market outcomes," *Journal of Human Resources*, 2022, *57* (6), 1885–1914.

Blakemore, Arthur E and Stuart A Low, "Sex differences in occupational selection: The case of college majors," *The Review of Economics and Statistics*, 1984, pp. 157–163.

Bleemer, Zachary, "Affirmative action and its race-neutral alternatives," *Journal of Public Economics*, 2023, *220*, 104839.

Bordon, Paola, Catalina Canals, and Alejandra Mizala, "The gender gap in college major choice in Chile," *Economics of Education Review*, 2020, *77*, 102011.

Breiman, Leo, "Random forests," *Machine learning*, 2001, *45*, 5–32.

Brenoe, Anne Ardila and Ulf Zolitz, "Exposure to more female peers widens the gender gap in STEM participation," *Journal of Labor Economics*, 2020, *38* (4), 1009–1054.

Bruhn, Jesse M, Christopher Campos, and Eric Chyn, "Who Benefits from Remote Schooling? Self-Selection and Match Effects," Technical Report, National Bureau of Economic Research 2023.

Bursztyn, Leonardo, Alexander W Cappelen, Bertil Tungodden, Alessandra Voena, and David H Yanagizawa-Drott, "How are gender norms perceived?," Technical Report, National Bureau of Economic Research 2023.

_ **, Thomas Fujiwara, and Amanda Pallais**, "âActing wifeâ: Marriage market incentives and labor market investments," *American Economic Review*, 2017, *107* (11), 3288–3319.

Card, David, "Estimating the return to schooling: Progress on some persistent econometric problems," *Econometrica*, 2001, *69* (5), 1127–1160.

_ **, Ana Rute Cardoso, and Patrick Kline**, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *The Quarterly journal of economics*, 2016, *131* (2), 633–686.

_ **,** _ **, Joerg Heining, and Patrick Kline**, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

_ **, Jörg Heining, and Patrick Kline**, "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly journal of economics*, 2013, *128* (3), 967–1015.

Carlana, Michela, Enrico Miglino, and Michela M Tincani, "How Far Can Inclusion Go? The Long-term Impacts of Preferential College Admissions," Technical Report, National Bureau of Economic Research 2024.

Carpio, Lucia Del and Maria Guadalupe, "More women in tech? Evidence from a field experiment addressing social identity," *Management Science*, 2022, *68* (5), 3196–3218.

Chetty, Raj, David J Deming, and John N Friedman, "Diversifying societyâs leaders? The causal effects of admission to highly selective private colleges," Technical Report, National Bureau of Economic Research 2023.

Chiappori, Pierre-Andre and Maurizio Mazzocco, "Static and intertemporal household decisions," *Journal of Economic Literature*, 2017, *55* (3), 985–1045.

Dahl, Gordon, Dan-Olof Rooth, and Anders Stenberg, "Long-run returns to field of study in secondary school," Technical Report, National Bureau of Economic Research 2020.

Dale, Stacy Berg and Alan B Krueger, "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables," *The Quarterly Journal of Economics*, 2002, *117* (4), 1491–1527.

Deming, David J, "Using school choice lotteries to test measures of school effectiveness," *American Economic Review*, 2014, *104* (5), 406–411.

Dur, Umut, Scott Duke Kominers, Parag A. Pathak, and Tayfun Sonmez, "Reserve Design: Unintended Consequences and the Demise of Boston's Walk Zones," *Journal of Political Economy*, 2018, *126* (6), 2457–2479.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg, "Revisiting the German wage structure," *The Quarterly journal of economics*, 2009, *124* (2), 843–881.

Einav, Liran, Amy Finkelstein, and Neale Mahoney, "Producing Health: Measuring Value Added of Nursing Homes," Technical Report, National Bureau of Economic Research 2022.

Ellison, Glenn and Parag A Pathak, "The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools," *American Economic Review*, 2021, *111* (3), 943–75.

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

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

Galor, Oded and David N Weil, "The Gender Gap, Fertility, and Growth," *The American Economic Review*, 1996, pp. 374–387.

Gartner, Hermann et al., "The imputation of wages above the contribution limit with the German IAB employment sample," *FDZ Methodenreport*, 2005, *2* (2005), 2005.

Goldin, Claudia, *Career and Family: Womenâs Century-Long Journey toward Equity*, Princeton University Press, 2021.

_ , "Nobel Lecture: An Evolving Economic Force," *American Economic Review*, 2024, *114* (6), 1515–1539.

_ **and Lawrence F Katz**, "Transitions: Career and family life cycles of the educational elite," *American Economic Review*, 2008, *98* (2), 363–369.

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

Hastie, Trevor, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman, *The elements of statistical learning: data mining, inference, and prediction*, Vol. 2, Springer, 2009.

Hastings, Justine S, Christopher A Neilson, and Seth D Zimmerman, "Are some degrees worth more than others? Evidence from college admission cutoffs in Chile," Technical Report, National Bureau of Economic Research 2013.

Heckman, James J, Sergio Urzua, and Edward Vytlacil, "Understanding instrumental variables in models with essential heterogeneity," *The review of economics and statistics*, 2006, *88* (3), 389–432.

Hederos, Karin and Anders Stenberg, "Gender identity and relative income within households: Evidence from Sweden," *The Scandinavian Journal of Economics*, 2022, *124* (3), 744–772.

Hoem, Jan M, Gerda Neyer, and Gunnar Andersson, "Education and childlessness The relationship between educational field, educational level, and childlessness among Swedish women born in 1955-59," *Demographic research*, 2006, *14*, 331–380.

Hoxby, Caroline M, "The changing selectivity of American colleges," *Journal of Economic perspectives*, 2009, *23* (4), 95–118.

Jayachandran, Seema, Lea Nassal, Matthew Notowidigdo, Marie Paul, Heather Sarsons, and Elin Sundberg, "Moving to opportunity, together," 2023.

Kirkeboen, Lars J, Edwin Leuven, and Magne Mogstad, "Field of study, earnings, and self-selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1111.

Kleven, Henrik, "The geography of child penalties and gender norms: Evidence from the United States," Technical Report, National Bureau of Economic Research 2022.

_ , Camille Landais, and Gabriel Leite-Mariante, "The child penalty atlas," Technical Report, National Bureau of Economic Research 2023.

_ , _ , and Jakob Egholt Sogaard, "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 2019, *11* (4), 181–209.

_ , _ , and Jakob Egholt Søgaard, "Does biology drive child penalties? evidence from biological and adoptive families," *American Economic Review: Insights*, 2021, *3* (2), 183–198.

_ , _ , Johanna Posch, Andreas Steinhauer, and Josef ZweimuÌller, "Child penalties across countries: Evidence and explanations," in "AEA Papers and Proceedings," Vol. 109 2019, pp. 122–26.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington, "The mommy effect: Do women anticipate the employment effects of motherhood?," Technical Report, National Bureau of Economic Research 2018.

Laverde, Mariana, "Distance to schools and equal access in school choice systems," *Unpublished Manuscript*, 2022.

Lemieux, Thomas, "Postsecondary education and increasing wage inequality," *American Economic Review*, 2006, *96* (2), 195–199.

Lovenheim, Michael and Jonathan Smith, "Returns to different postsecondary investments: Institution type, academic programs, and credentials," in "Handbook of the Economics of Education," Vol. 6, Elsevier, 2023, pp. 187–318.

Lovenheim, Michael F and Kevin J Mumford, "Do family wealth shocks affect fertility choices? Evidence from the housing market," *Review of Economics and Statistics*, 2013, *95* (2), 464–475.

Melentyeva, Valentina and Lukas Riedel, "Child penalty estimation and mothers' age at first birth," Technical Report, ECONtribute Discussion Paper 2023.

Michelmore, Katherine and Kelly Musick, "Fertility patterns of college graduates by field of study, US women born 1960–79," *Population studies*, 2014, *68* (3), 359–374.

Mountjoy, Jack, "Community colleges and upward mobility," *American Economic Review*, 2022, *112* (8), 2580–2630.

_ , "Marginal Returns to Public Universities," Technical Report, National Bureau of Economic Research 2024.

_ and Brent R Hickman, "The Returns to College (s): Relative Value-Added and Match Effects in Higher Education," Technical Report, National Bureau of Economic Research 2021.

Neilson, Christopher, "The Rise of Centralized Choice and Assignment Mechanisms in Education Markets Around the World," 2021.

_ , **Federico Huneeus, Conrad Miller, Seth Zimmerman et al.**, "Firm Sorting, College Major, and the Gender Earnings Gap," Technical Report 2021.

OECD, *OECD handbook for internationally comparative education statistics: concepts, standards, definitions and classifications*, OECD Paris, 2004.

Oh, Suanna, "Does identity affect labor supply?," *American Economic Review*, 2023, *113* (8), 2055–2083.

Oreopoulos, Philip and Uros Petronijevic, "Making college worth it: A review of research on the returns to higher education," 2013.

Otero, Sebastián, Nano Barahona, and Cauê Dobbin, "Affirmative action in centralized college admission systems: Evidence from Brazil," *Unpublished manuscript*, 2021.

Patnaik, Arpita, Matthew Wiswall, and Basit Zafar, "College majors 1," *The Routledge handbook of the economics of education*, 2021, pp. 415–457.

Roy, Andrew Donald, "Some thoughts on the distribution of earnings," *Oxford economic papers*, 1951, *3* (2), 135–146.

Sloane, Carolyn M, Erik G Hurst, and Dan A Black, "College Majors, Occupations, and the Gender Wage Gap," *Journal of Economic Perspectives*, 2021, *35* (4), 223–48.

Turner, Sarah E and William G Bowen, "Choice of major: The changing (unchanging) gender gap," *ILR Review*, 1999, *52* (2), 289–313.

Wang, Richard J, Samer I Al-Saffar, Jeffrey Rogers, and Matthew W Hahn, "Human generation times across the past 250,000 years," *Science Advances*, 2023, *9* (1).

Wiswall, Matthew and Basit Zafar, "Preference for the workplace, investment in human capital, and gender," *The Quarterly Journal of Economics*, 2018, *133* (1), 457–507.

_ **and** _ , "Human capital investments and expectations about career and family," *Journal of Political Economy*, 2021, *129* (5), 1361–1424.

Zafar, Basit, "College Major Choice and the Gender Gap. Staff Report No. 364.," *Federal Reserve Bank of New York*, 2009.

Zimmerman, Seth D, "The returns to college admission for academically marginal students," *Journal of Labor Economics*, 2014, *32* (4), 711–754.

_ , "Elite colleges and upward mobility to top jobs and top incomes," *American Economic Review*, 2019, *109* (1), 1–47.

Zinovyeva, Natalia and Maryna Tverdostup, "Gender identity, coworking spouses, and relative income within households," *American Economic Journal: Applied Economics*, 2021, *13* (4), 258–284.

Online Appendix for:

Major Choice and Gender Gaps

Alonso Bucarey Christopher Campos Dante Contreras Pablo Muñoz

October 2024

Table of Contents

A Data Appendix **2**

A.1 Major Classification . . . 2

A.2 Data Imputations . . . 2

A.3 Descriptive Evidence . . . 3

A.4 AKM Estimation . . . 3

B Additional Results **11**

B.1 Returns' Dynamics . . . 11

B.2 Role of Firm Premia . . . 11

B.3 Additional Results . . . 11

B.4 Counterfactual Preferences . . . 12

C Child Penalties and Gender Norms **25**

C.1 Child Penalty Gap . . . 25

C.2 Gender Norms . . . 25

D Counterfactual Details **34**

D.1 Deferred Acceptance Algorithm . . . 34

D.2 Prediction Accuracy and Summary Statistics . . . 34

D.3 Counterfactual Bootstrap Details . . . 35

A Data Appendix

This section discusses additional details regarding the data used in our analysis. We begin by providing details about our categorization of majors that are crucial in our analysis. Then, since earnings records are top-coded, we discuss our imputation procedure. We also report additional descriptive evidence related to trends in gender representation, preferences, and standardized test scores mentioned in the paper. Last, we report details of our AKM (Abowd et al., 1999) estimation procedure used to study the role of firm sorting.

A.1 Major Classification

We classify students into fields of study based on the major from which they graduate. Degree programs are classified by field of study mostly based on the OECD Handbook for Internationally Comparative Education Statistics (OECD, 2004). There are eight broad categories: "*Agriculture*", "*Science*", "*Social Sciences, Business and Law*", "*Teaching*", "*Humanities and Arts*", "*Engineering, Manufacturing and Construction*", "*Health and Welfare*", and "*Services*".

We reclassify the category "*Social Sciences, Business and Law*" into three separate fields of study "*Social Sciences*", "*Business*", and "*Law*". The field of "*Social Sciences*" includes the degrees of anthropology, library science, political sciences, social communication, geography, journalism, psychology, sociology, and social work. The field of "*Business*" includes the degrees: commercial engineering, accounting, commerce engineering, business administration, marketing engineering, logistics engineering, foreign trade engineering, management control engineering, human resources engineering, finance engineering, public management, advertising, and public relations; while the field of "*Law*" includes law degrees.

We also separate "*Medicine*" from "*Health and Welfare*". In particular, we include all degrees covered by the "Medical Law" in Chile (Ley N 19,664) into "*Medicine*". These degrees are in medicine, dentistry, pharmaceutical chemistry, and biochemistry. Finally, we drop "*Agriculture*" and "*Services*" as they represent very few graduates (less than 5%) from non-homogeneous college programs in Chile.

A.2 Data Imputations

To impute wages above the social security contribution limit, we proceed as in Dustmann et al. (2009) and Card et al. (2013). First, we fit a series of Tobit models to log wages separately by gender, including the type of high school, test-score range, and region controls.

Then, we impute an uncensored value for each censored observation using the estimated parameters of these models and random drawings ϵ from a truncated distribution. Following Gartner et al. (2005):

$$\epsilon_i = \Phi^{-1}\left(u \times \left[1 - \Phi\left(\frac{c - X_i'\hat{\beta}}{\hat{\sigma}}\right)\right] + \Phi\left(\frac{c - X_i'\hat{\beta}}{\hat{\sigma}}\right)\right),$$

where $u \sim$ U[0,1], c is the social security contribution limit), $X'\hat{\beta}$ is the tobit prediction, and $\hat{\sigma}$ is the standard deviation of the tobit error. Figure A.1 below presents both the distribution of the original log earnings and the imputation-adjusted log earnings used in our analysis.

A.3 Descriptive Evidence

We present descriptive evidence that motivates and complements our work.

Historical Trends: Figure A.2 documents trends in gender representation across majors and its potential contribution to the overall gender gap. In the spirit of Sloane et al. (2021), and leveraging the largest household survey in Chile, the figure shows the major "similarity index" (as a re- normalization of the inverse Duncan-Duncan index) and a "potential wage index" (that assigns to everyone within a major the average hourly wages of prime-age male workers in that major). This figure shows that majors in Chile have become less segregated over time. See section 2 for details.

Fields of Study: To characterize preferences for fields of study, Figure A.4 reports fallback major prevalence and Figure A.3 reports a series of bivariate correlations of our estimated control functions. Overall we find sensible patterns, i.e., that applicants with larger estimated tastes for Medicine also have larger tastes for Health programs, but those preferences are negatively correlated with those for Social Science or Humanities. Similarly, males and females with larger tastes for Social Science have larger tastes for Humanities and Law but have less intense preferences for Science and Engineering. Finally, Figure A.5 presents the distribution of college admission test scores in each of the nine field of study we consider.

A.4 AKM Estimation

Following the seminal study of the French labor market by Abowd et al. (1999), and more recent studies (Card et al., 2013, 2016, 2018), we estimate a so-called AKM model for log wages that includes additive effects for workers and firms. Specifically, our model for the log wage of person i in year t takes the form:

$$\ln w_{it} = \alpha_i + \psi_{J(i,t)} + \mathbf{X'}_{it}\beta + \varepsilon_{it}, \tag{17}$$

where α_i is a "person effect" capturing the (time-invariant) portable component of earnings ability, $\psi_{J(i,t)}$ is a function indicating the employer of worker i in year t, and ε_{it} is an unobserved time-varying error capturing shocks to human capital, person-specific job match effects, and other factors. The vector $\mathbf{X}_{it}$ includes year fixed effects and a third-order polynomial of age interacted with five education categories (some primary education, some secondary education, complete secondary or incomplete higher education, complete higher education, and a "no education data" category). Following Card et al. (2013), we recenter age around 40.

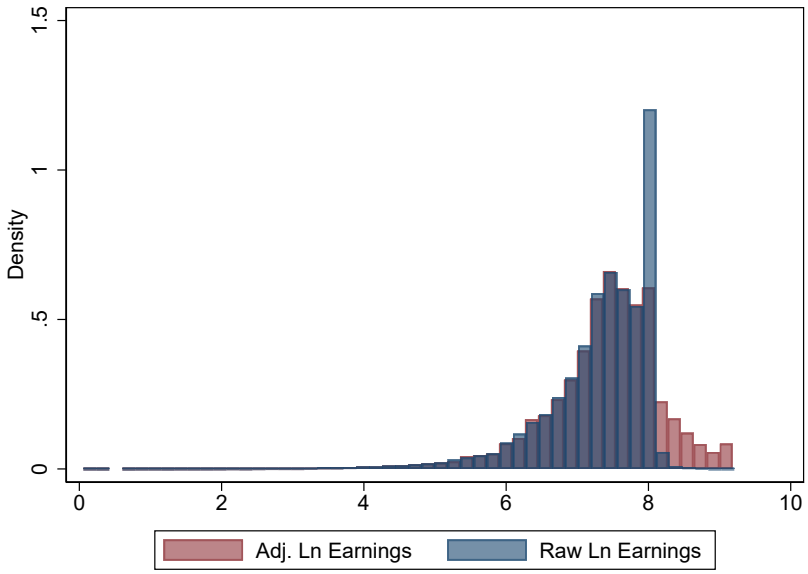
Importantly, the fixed effects in Equation (17) are identified by movers and can only be compared within connected sets. Thus, we estimate our model within the largest connected set of workers and firms, including 59,969,972 observations of 8,588,072 workers and 757,698 firms over an eleven-year period (2009-2019). Our estimates show that 18% of the variance of wages is explained by firms. This is consistent with previous studies of wage determination with worker and firm fixed effects, which typically find that firm-specific premiums explain 20% of overall wage variation (Card et al., 2018).

For our analysis, we want to estimate firm-specific premiums $\psi_{J(i,t)}$ and a key assumption for identification of these parameters is "exogenous mobility", i.e., worker mobility must be uncorrelated with the time-varying residual components of wages. Following Card et al. (2013), we present a simple event-study analysis of the wage changes experienced by workers moving between different groups of firms. For this, we define firm quartiles based on the average pay of each firm. As explained in Card et al. (2018), if the AKM model is correct and firms offer proportional wage premiums for all their employees, workers who move to higher-paying firms will, on average, experience pay raises. At the same time, those who move in the opposite direction will experience pay cuts. Moreover, the gains and losses for movers in opposite directions between any two groups of firms will be symmetric.[24]

Figure A.6 present our event study. Reassuringly, we find that the wage profiles exhibit steplike patterns. When workers move to higher-paying establishments, their wages rise; when they move to lower-paying establishments, their wages fall. The wage changes for people who stay in the same firm type are close to 0.

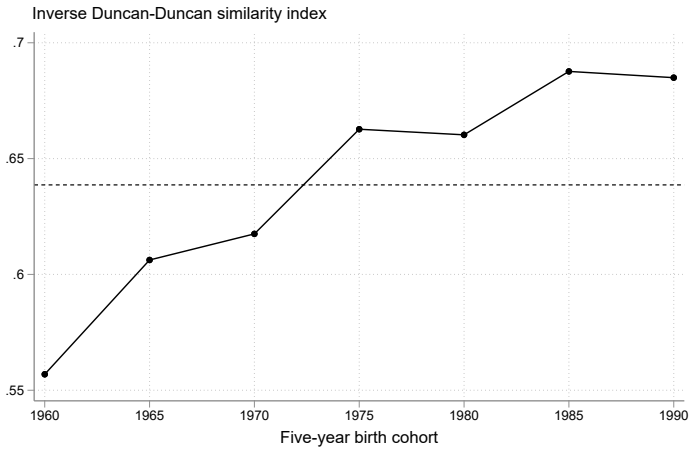
[24]This is in contrast to models of mobility linked to the worker-and-firm-specific match component of wages that imply that movers will tend to experience positive wage gains regardless of the direction of their move, violating the symmetry prediction.

Figure A.1: Log Earnings Distribution

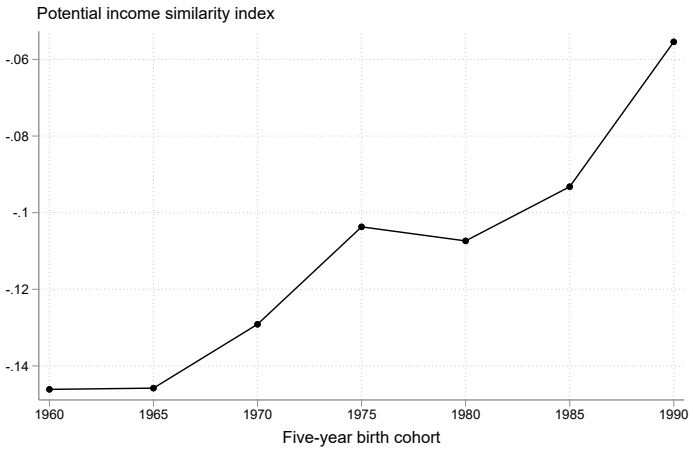


Notes: This figure presents the histogram of the original log earnings and the imputation-adjusted log earnings. We adjust earnings at the contribution limit following Dustmann et al. (2009) and Card et al. (2013). For details, see section A.2.

Figure A.2: Similarity and Potential Wage Indexes by Field of Study across Cohorts cohorts



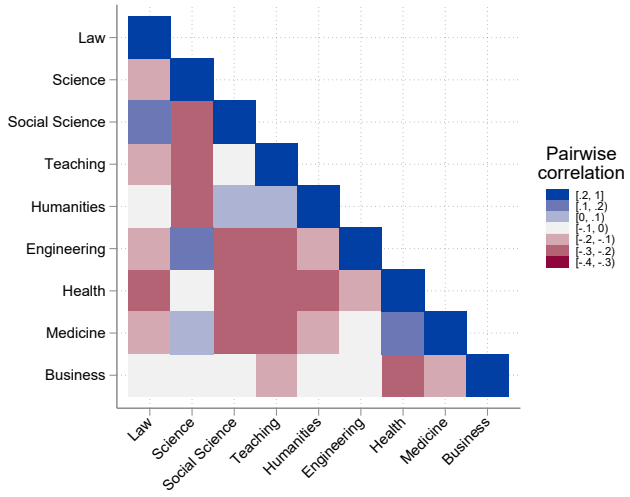
(a) Gender Similarity Index



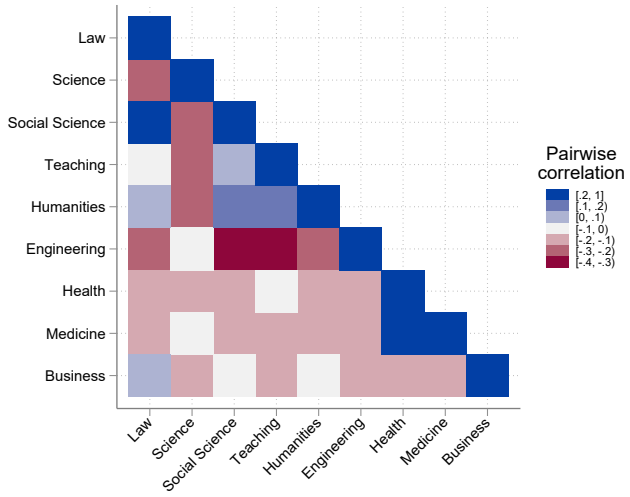
(b) Potential Wage Index

Notes: This figure presents trends in the similarity index and wage potential by field of study across cohorts, as in Sloane et al. (2021). Data comes from the largest household survey in Chile and is restricted to those with at least a bachelor's degree. Panel (a) plots the renormalized, inverse Duncan-Duncan index for different cohorts of Chilean college graduates. Panel (b) plots the potential wage index for different cohorts of Chilean college graduates.

Figure A.3: Pairwise Correlations of "Unobserved" Preferences



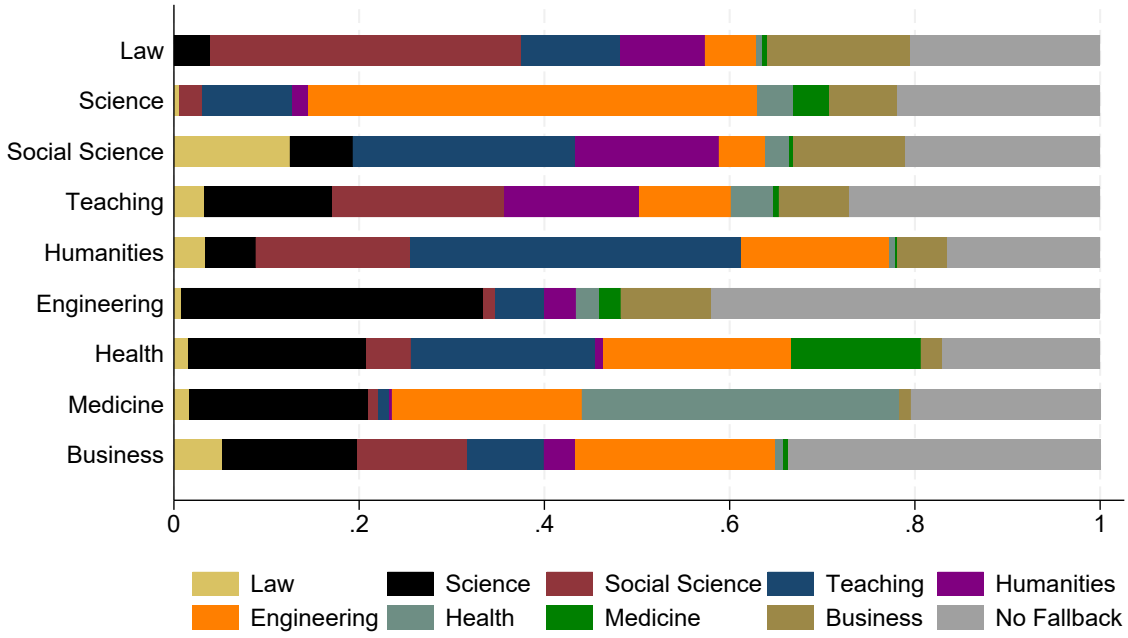
(a) Female



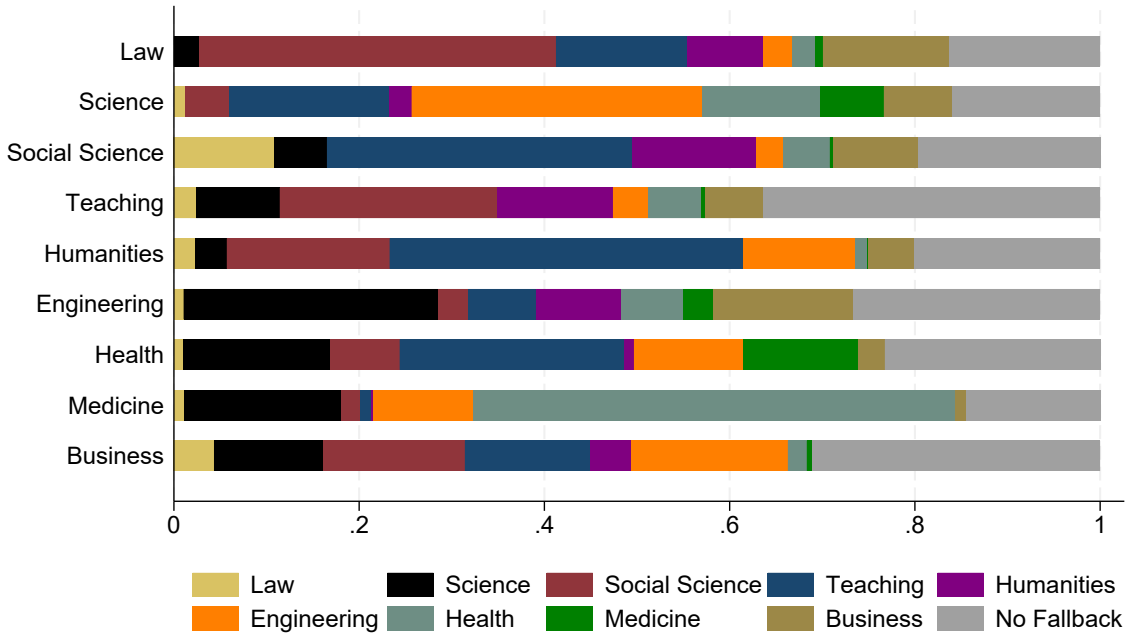
(b) Male

Notes: This figure shows the pairwise correlations between students' estimated field-specific control functions, proxies for preference intensity. Panel (a) shows the correlation matrix for unobserved female preferences. Panel (b) shows the correlation matrix for unobserved male preferences. We consider all students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed in 2019.

Figure A.4: Applicants' Fallbacks



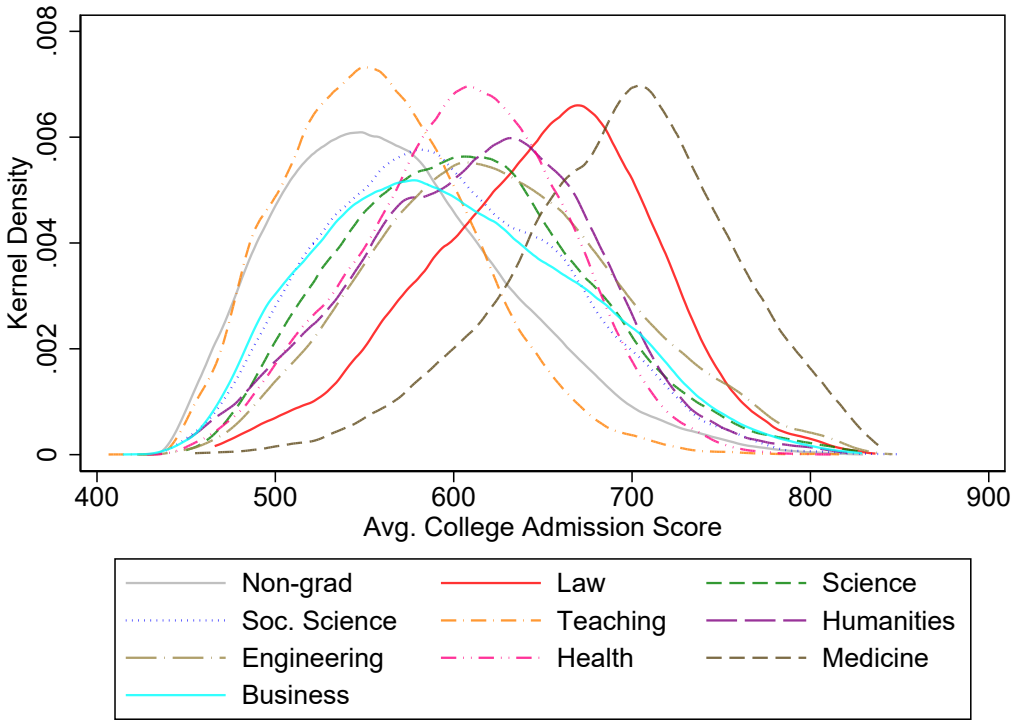
(a) Male



(b) Female

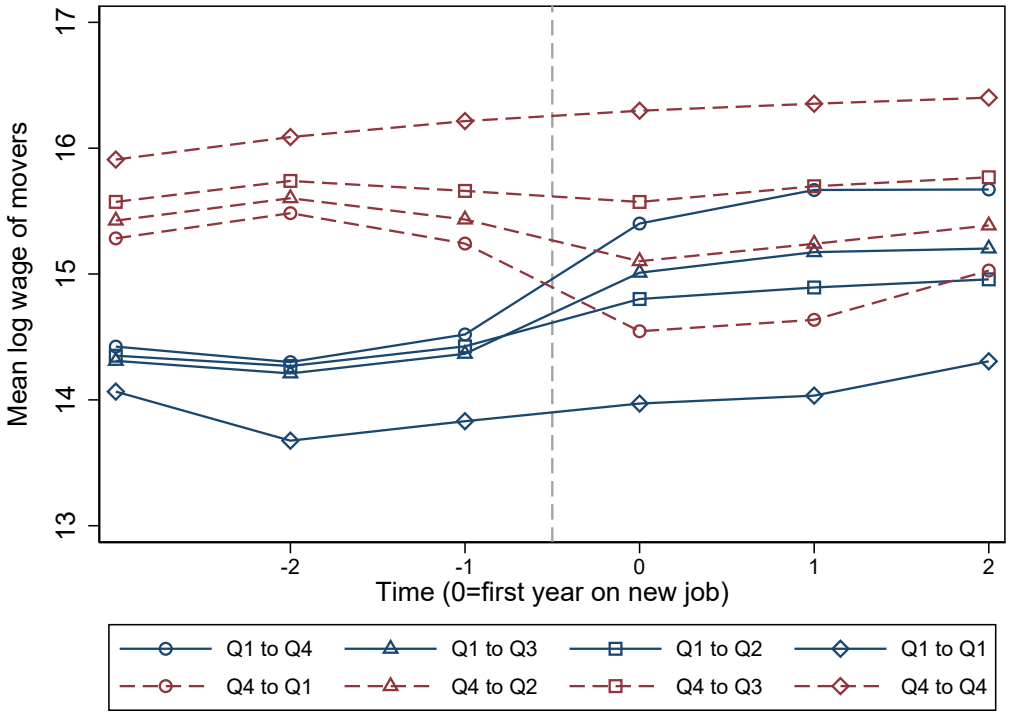
Notes: This figure shows students' fallback fields of study based on their ranked order lists. Among students ranking a given field as their first choice (y-axis), we compute the share of students ranking each of the remaining fields of study as a second option (i.e., fallback). Panels (a) and (b) show the relevance of each fallback for males and females, respectively. We consider all students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed in 2019.

Figure A.5: Kernel Density of College Admission Scores



Notes: This figure shows the kernel densities of the college admission score (average score in the two mandatory exams of mathematics and language) of students graduating from different fields of study. We consider all students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed in 2019.

Figure A.6: Mean Wages of Job Changers



Notes: This figure shows the mean log wages of job changers classified by the quartile of the wages paid at origin and destination firms, in the spirit of Card et al. (2018). We consider workers and firms in the Unemployment Insurance dataset between 2009 and 2019 who are within the largest connected set.

B Additional Results

This section provides additional results mentioned in the paper. We provide evidence on the dynamics of the gender earnings gap through the lens of a decomposition between differences in returns and differences in match effects. We also focus on the role of gender differences in sorting to firms with different pay premiums, viewed through the lens of the canonical AKM model Abowd et al. (1999). We then discuss a set of additional results that all point to qualitatively similar conclusions as the results in the main paper.

B.1 Returns' Dynamics

Appendix Figure B.1 reports dynamics in the gender earning gap over time for separate fields of study. Each subfigure reports the within-major gender earnings gap in black, the portion of that gap due to differences in average returns (α_j), in gray, and the portion due to differences in realized match effects ($\psi_j^* \times \bar{\lambda}_j$), is reported in maroon. We report these contributions for a total of eight years, starting with seven years after applying to college and ending with 14 years post-application. Besides a few cases, match effects play a minimal role in explaining the within-major gap.

B.2 Role of Firm Premia

Following Abowd et al. (1999), and leveraging records from the Unemployment Insurance Administrator Agency (which only include private sector workers but has firm identifiers), we estimate firm premia from a two-way fixed-effects model.[25] Our estimates of gender-specific returns in terms of firm premiums are presented in Appendix Table B.1. We find heterogeneous impacts of fields of study on firm fixed-effects, with the fields of Science, Engineering, and Business being associated with the largest premia. However, we fail to find evidence of significant gender gaps in firm premia across fields of study, except for Teaching, where women sort into firms paying larger premia than men. Mostly, we find null evidence of heterogeneity in match effects.[26] In terms of dynamics, Appendix Figure B.2 shows the relative importance of the return and firm premia channels to account for the gender earnings gap over time. Firm premia seem to play a modest role in explaining the earnings gap dynamic in most fields, except for the cases of Teaching and Engineering. In both cases, the differences in firm premia favor women initially but diminish over time.

B.3 Additional Results

Private Sector Earnings Data: Appendix Table B.2 reports estimates using the private sector data discussed in Section 2. The overarching results are qualitatively similar to our main estimates. In this case, however, the return to Medicine is relatively smaller. This stems from the fact that many medical doctors are employed in the public sector in Chile, so estimates

[25] Details on the estimation of firm fixed-effects are provided in Appendix A.4.

[26] For reference, we also estimate our main model using log earnings with data from the Unemployment Insurance. These results are presented in Appendix Table B.2.

using private sector data will invariably miss a significant portion of their labor market. For this reason, we use the pension system data as our primary evidence.

Total Earnings: Appendix Table B.3 reports estimates using the pension data but considering total yearly earnings instead of log earnings as the dependent variable. This allows us to include individuals not employed in the Chilean formal labor market in 2019. We find similar results when considering this outcome variable encompassing the intensive and extensive margins of the labor market. We find differences in returns across fields of study and gender. In line with our main results, men exhibit larger earnings in Science and Engineering, and women exhibit higher earnings in Teaching and Humanities, fields with relatively lower returns.

Alternative Non-Pecuniary Returns: The fertility results in the paper are on the extensive margin of having a child by a certain time after graduation. Appendix Table B.4 reports the estimates obtained when the outcome is the total number of children. This captures both extensive and intensive margin effects. Reassuringly, we find similar results when accounting for the intensive margin of fertility.

Focus on College Graduates: The results presented in the main body of the paper consider applicants who were accepted into a field of study, independently of whether they graduated or not. As a robustness check, in this appendix, we exclude the non-graduation margin and restrict our analysis to a sample of college graduates, using Health as the omitted field of study. Appendix Tables B.5 and B.6 present our main results when considering this alternative sample. We find that the relative returns are strikingly consistent with those presented in Tables 2 and 3 of the paper. Table B.7 further presents the validation results for these alternative models. We see that the outcomes align qualitatively with our preferred model in the sense that both pecuniary and non-pecuniary impacts are forecast unbiased.

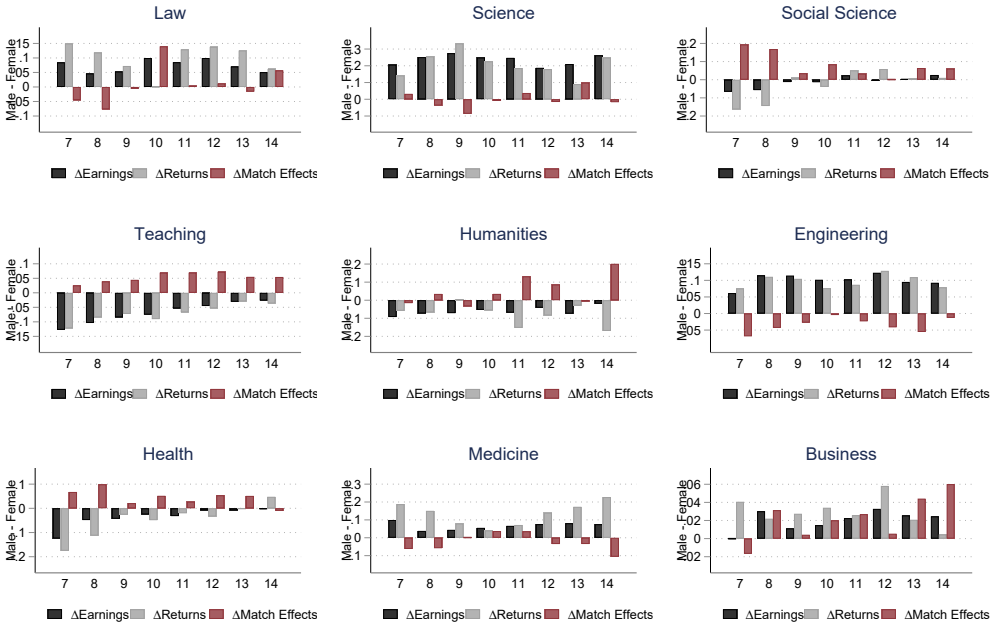
B.4 Counterfactual Preferences

To assess the extent to which family considerations relative to earnings lead to gender gaps, we perform a simple counterfactual exercise in which we equalize the importance assigned to family considerations across men and women. For details, see section 4.

The first bar of Figure B.3 reports the baseline and observed gender earnings gap among our sample of students. The second bar, referred to as "Male Preferences", assigns all students the preferences for family considerations of men. Similarly, the third bar, referred to as "Female Preferences", assigns all students the preferences for family considerations of women. For each counterfactual bar, we use estimates reported in Table 5 in the main body of the paper to construct counterfactual enrollment shares. Our exercise shows that equalizing family considerations between genders—while holding other preferences constant—can reduce the earnings gap by 17-24 percent. Figure B.4 offers additional evidence, based upon the decomposition presented in Equation (1), and shows that equalizing preferences for family considerations significantly reduces the contribution of Engineering programs to the overall gender earnings gap while also diminishing the offsetting impact of Teaching and Health programs.

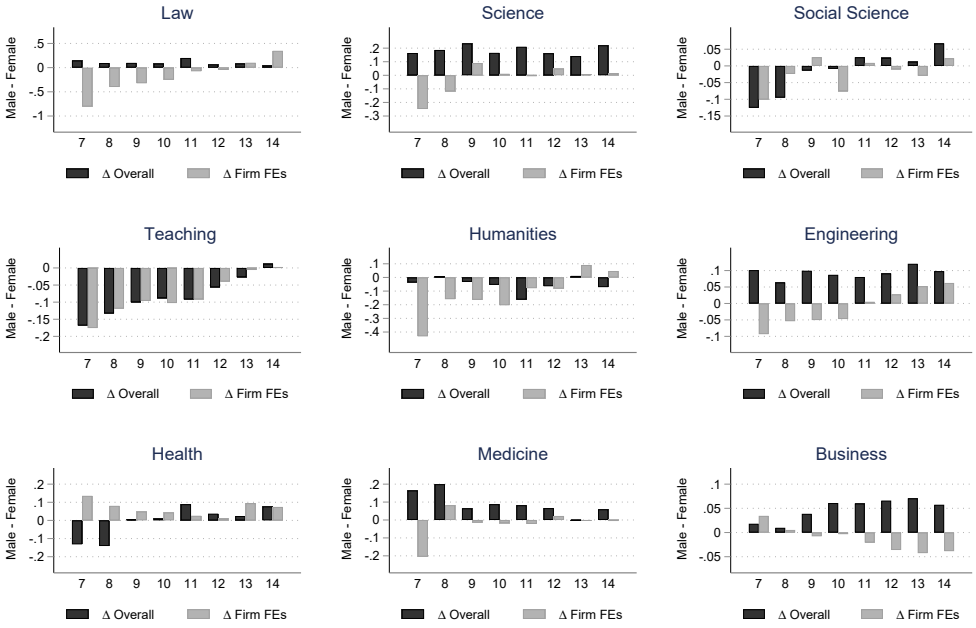
As discussed in section 4.3, differences in representation across fields of study seem to be the primary driver of the gap in each counterfactual scenario, reflecting the relatively modest gender differences in the causal impacts of majors.

Figure B.1: Gender Earnings Gap Over Time



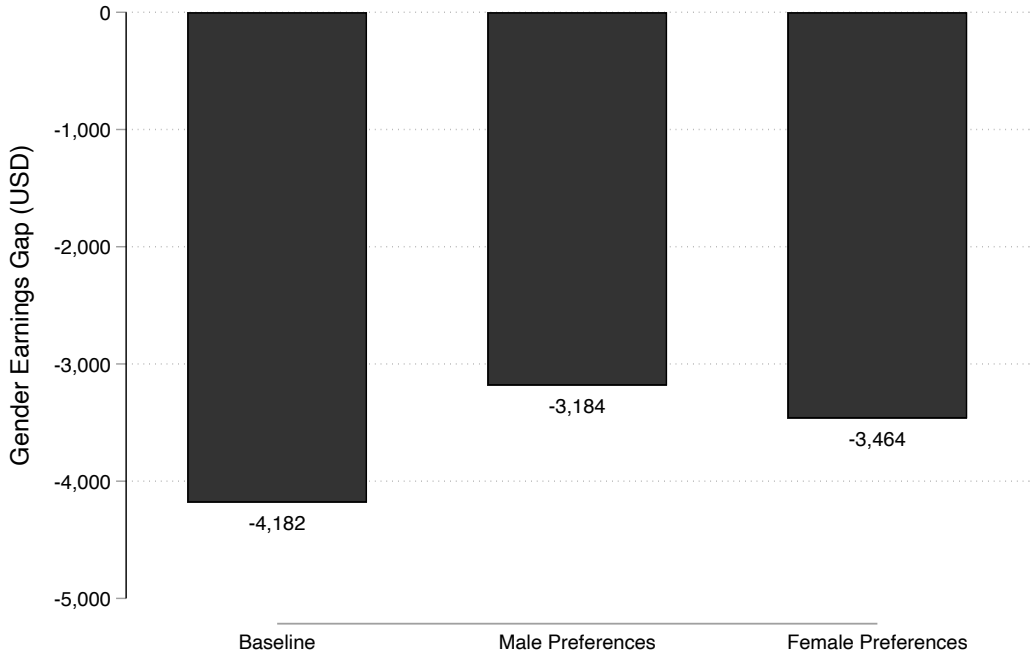
Notes: This figure shows the results obtained from estimating Equation (9) using labor market earnings at different time horizons (after high school graduation) as the dependent variable. For each field of study, the figure presents the overall gender earnings gap and its components: the gender gap in returns and the gender gap in match effects. The gender gap in level effects is omitted. We consider all students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed at a given time horizon after they graduated from high school.

Figure B.2: Gender Earnings Gap Over Time: The Role of Firms



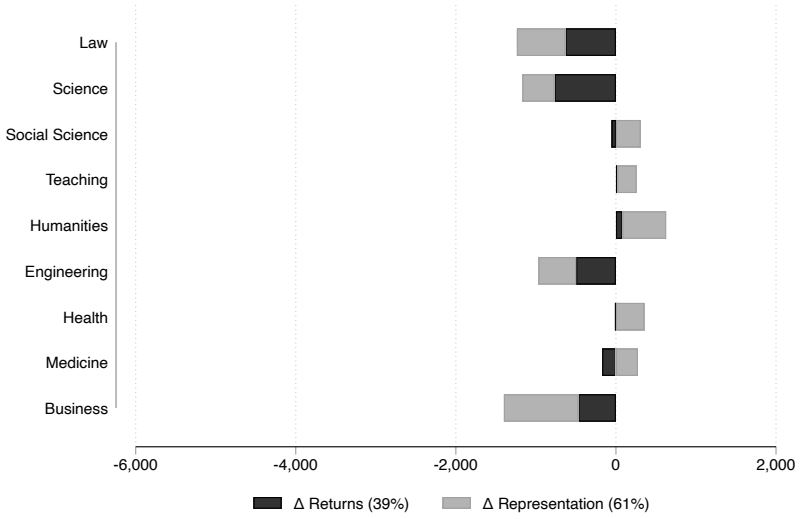
Notes: This figure shows the results obtained from estimating Equation (9) using labor market earnings and firm premia at different time horizons (after high school graduation) as the dependent variable. We only consider workers in the private sector since firm identifiers are only available for them. For each field of study, the figure presents the overall gender earnings gap alongside the gender gap in firm premia (i.e., firm fixed effects estimated following Card et al. (2013); see Section A.4 for details). We consider all students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed at a given time horizon after they graduated from high school.

Figure B.3: Baseline and Counterfactual Earnings Gap

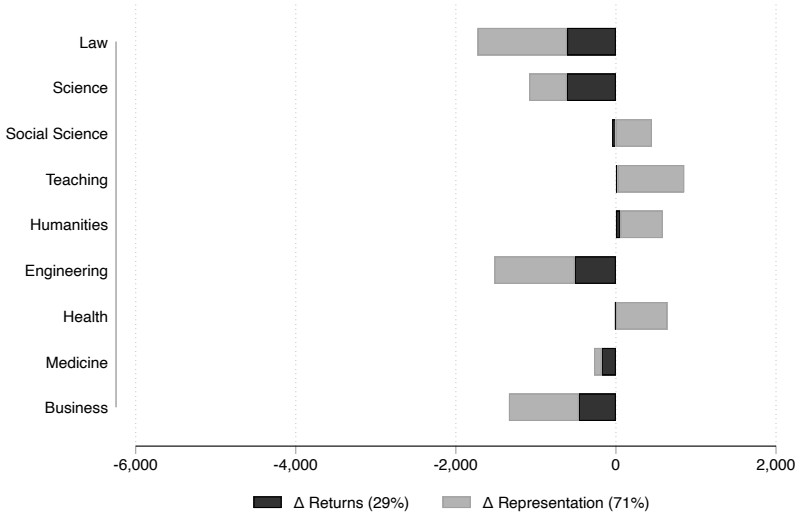


Notes: This figure reports the gender earnings gap in US dollars in three scenarios. We use the annual earnings 2019 of students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed in 2019. The first bar reports the baseline and observed gender earnings gap among this sample of students. The second bar, referred to as Male Preferences, assigns all students the preferences for family considerations of men. Similarly, the third bar, referred to as Female Preferences, assigns all students the preferences for family considerations of women. For each counterfactual bar, we use estimates reported in Table 5, in the main body of the paper, to construct counterfactual enrollment shares.

Figure B.4: Baseline and Counterfactual Earnings Gap: Sorting versus Returns



(a) Baseline



(b) Counterfactual

Notes: These figures show the result from the exercise described by Equation (1) that decomposes the gender earnings gap into differences in sorting between fields of study and differences in returns within fields of study. We consider the annual earnings 2019 of students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed in 2019. Panel (a) plots the decomposition when we assign male preferences for family considerations to every observation. Panel (b) plots the decomposition when we assign everyone female preferences for family considerations. The predictions use estimates reported in Table 5 in the main text.

Table B.1: Gender Differences in Returns and Match Effects: Firm Premia

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Female				
Returns	-0.084	0.223	0.080	0.107	-0.008	0.218	0.070	0.102	0.218
	(0.139)	(0.059)	(0.023)	(0.015)	(0.051)	(0.040)	(0.023)	(0.025)	(0.023)
Selection on Gains	0.081	0.015	0.008	0.009	0.050	-0.014	0.041	-0.005	0.016
	(0.046)	(0.028)	(0.013)	(0.011)	(0.021)	(0.021)	(0.013)	(0.010)	(0.012)
					Panel B: Male				
Returns	-0.010	0.155	0.074	0.026	-0.091	0.226	0.122	0.098	0.234
	(0.155)	(0.035)	(0.035)	(0.032)	(0.082)	(0.021)	(0.043)	(0.030)	(0.026)
Selection on Gains	0.066	0.050	0.010	0.038	0.066	0.005	-0.006	-0.007	0.005
	(0.044)	(0.021)	(0.015)	(0.018)	(0.029)	(0.018)	(0.018)	(0.012)	(0.014)
Est. Females = Est. Males									
Δ Returns (p-val)	0.724	0.321	0.880	0.019	0.384	0.865	0.269	0.929	0.625
Δ Match Effects (p-val)	0.817	0.317	0.913	0.161	0.662	0.479	0.036	0.876	0.548

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9). We consider the firm premia estimated from an AKM model (Abowd et al., 1999; Card et al., 2013) as the dependent variable. The sample includes 77,414 observations of students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed in the private sector in 2019.

Table B.2: Gender Differences in Returns and Match Effects: Private Sector Data

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Female				
Returns	0.391	0.376	0.222	0.284	-0.023	0.517	0.246	0.429	0.525
	(0.071)	(0.045)	(0.020)	(0.011)	(0.047)	(0.022)	(0.023)	(0.054)	(0.018)
Selection on Gains	0.058	0.009	0.039	0.021	0.069	0.016	0.065	-0.013	0.019
	(0.023)	(0.022)	(0.011)	(0.008)	(0.019)	(0.013)	(0.013)	(0.023)	(0.009)
					Panel B: Male				
Returns	0.434	0.531	0.192	0.193	-0.124	0.614	0.278	0.428	0.580
	(0.092)	(0.023)	(0.036)	(0.021)	(0.065)	(0.013)	(0.049)	(0.069)	(0.019)
Selection on Gains	0.105	0.036	0.060	0.066	0.087	0.004	0.066	0.032	0.052
	(0.030)	(0.014)	(0.016)	(0.011)	(0.023)	(0.011)	(0.019)	(0.027)	(0.010)
Est. Females = Est. Males									
Δ Returns (p-val)	0.710	0.002	0.459	0.000	0.203	0.000	0.551	0.983	0.025
Δ Match Effects (p-val)	0.224	0.307	0.289	0.001	0.532	0.489	0.942	0.203	0.014

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9). We consider the earnings of workers in the private sector. The sample includes 78,336 observations of students who applied and were accepted through the centralized admission system between 2004 and 2007 and were employed in the private sector in 2019.

Table B.3: Gender Differences in Returns and Match Effects: Total Earnings

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Female				
Returns	7,079	4,261	3,212	7,245	-2,402	10,040	7,442	17,444	12,056
	(1,185)	(961)	(336)	(148)	(558)	(660)	(286)	(956)	(483)
Selection on Gains	1,418	197	539	-196	1,245	486	1,008	40	1,002
	(417)	(480)	(188)	(115)	(235)	(362)	(172)	(415)	(234)
					Panel B: Male				
Returns	6,042	9,170	2,507	6,451	-2,739	14,226	7,013	21,592	14,093
	(1,368)	(698)	(538)	(258)	(671)	(374)	(604)	(1,200)	(542)
Selection on Gains	2,450	1,128	1,289	625	1,548	744	1,221	38	1,396
	(482)	(413)	(253)	(155)	(253)	(325)	(264)	(478)	(284)
Est. Females = Est. Males									
Δ Returns (p-val)	0.566	0.000	0.262	0.005	0.696	0.000	0.511	0.007	0.005
Δ Match Effects (p-val)	0.105	0.141	0.017	0.000	0.379	0.596	0.499	0.997	0.285

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9). We consider the sum of earnings in 2019 (in USD) instead of the log of earnings to account for the extensive margin. We restrict the sample to students who applied and were accepted through the centralized admission system between 2004 and 2005 and were employed in 2019. The sample includes 137,339 observations.

Table B.4: Fertility: Number of children 12 years after college application

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Panel A: Female				
Returns	-0.184	-0.239	-0.072	0.119	-0.151	-0.063	0.087	-0.006	-0.005
	(0.050)	(0.040)	(0.022)	(0.015)	(0.040)	(0.028)	(0.019)	(0.033)	(0.023)
Selection on Gains	0.060	0.001	-0.006	-0.020	0.022	-0.016	-0.010	0.018	0.012
	(0.018)	(0.020)	(0.012)	(0.011)	(0.016)	(0.016)	(0.011)	(0.015)	(0.011)
					Panel B: Male				
Returns	-0.241	-0.345	-0.277	-0.136	-0.436	-0.126	-0.142	-0.144	-0.122
	(0.049)	(0.023)	(0.030)	(0.022)	(0.036)	(0.015)	(0.031)	(0.037)	(0.021)
Selection on Gains	0.058	0.043	0.012	-0.003	0.047	-0.008	0.020	0.046	-0.018
	(0.018)	(0.014)	(0.014)	(0.012)	(0.014)	(0.013)	(0.013)	(0.015)	(0.011)
Est. Females = Est. Males									
Δ Returns (p-val)	0.415	0.018	0.000	0.000	0.000	0.040	0.000	0.004	0.000
Δ Match Effects (p-val)	0.929	0.075	0.338	0.294	0.239	0.687	0.091	0.177	0.059

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9). The sample includes 137,339 observations of students who applied and were accepted through the centralized admission system between 2004 and 2007.

Table B.5: Sample of College Graduates: Log Earnings in 2019

	Law	Science	Social Science	Teaching	Humanities	Engineering	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Returns	0.035	-0.015	-0.162	-0.008	-0.401	0.132	0.277	0.199
	(0.046)	(0.039)	(0.020)	(0.011)	(0.043)	(0.018)	(0.027)	(0.016)
Selection on Gains	0.023	0.003	0.018	-0.005	0.045	0.036	-0.008	0.037
	(0.016)	(0.021)	(0.010)	(0.007)	(0.016)	(0.012)	(0.013)	(0.008)
				Panel B: Male				
Returns	0.057	0.179	-0.233	-0.070	-0.519	0.235	0.397	0.263
	(0.054)	(0.026)	(0.029)	(0.014)	(0.048)	(0.015)	(0.031)	(0.018)
Selection on Gains	0.050	0.038	0.044	0.024	0.064	0.043	-0.013	0.036
	(0.018)	(0.012)	(0.015)	(0.010)	(0.018)	(0.010)	(0.015)	(0.009)
Est. Females = Est. Males								
Δ Returns (p-val)	0.756	0.000	0.035	0.000	0.063	0.000	0.003	0.002
Δ Match Effects (p-val)	0.265	0.155	0.165	0.014	0.453	0.646	0.804	0.943

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9), when estimated in a sample of college graduates, with health as the omitted field of study. We display the averages—of coefficients, robust standard errors, and p-values—across 100 regressions, each of which uses a different set of control functions obtained after parametric bootstrap over the ranked ordered logit estimation. The sample includes 111,831 observations of students who applied through the centralized admission system between 2004 and 2007, graduated between 2007 and 2019, and were employed in 2019. Panel (a) focuses on parameter estimates for Female students and Panel (b) focuses on Male students. In each panel, we report return estimates and selection on gains parameters, also referred to as match effects in the text. Underneath each estimate in parentheses is the bootstrapped standard error. The final two rows of the table report p-values from two additional statistical tests. The Δ Returns p-values correspond to the null hypothesis that the Female and Male returns are equivalent for a given column. The Δ Match Effects p-values correspond to a similar null hypothesis that the Female and Male match effects are equivalent for a given column.

Table B.6: Sample of College Graduates: Any child 12 years after college application

	Law	Science	Social Science	Teaching	Humanities	Engineering	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Female				
Returns	-0.108	-0.079	-0.017	0.099	-0.094	-0.044	-0.077	-0.000
	(0.017)	(0.018)	(0.010)	(0.007)	(0.015)	(0.010)	(0.013)	(0.009)
Selection on Gains	0.009	-0.002	-0.001	-0.004	0.024	-0.005	-0.001	0.008
	(0.006)	(0.009)	(0.005)	(0.004)	(0.006)	(0.007)	(0.006)	(0.005)
				Panel B: Male				
Returns	-0.114	-0.110	-0.111	-0.025	-0.194	-0.073	-0.072	-0.076
	(0.018)	(0.012)	(0.012)	(0.009)	(0.015)	(0.008)	(0.013)	(0.009)
Selection on Gains	-0.002	-0.001	0.003	-0.005	0.018	0.006	0.013	-0.005
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)
Est. Females = Est. Males								
Δ Returns (p-val)	0.788	0.137	0.000	0.000	0.000	0.009	0.779	0.000
Δ Match Effects (p-val)	0.223	0.891	0.557	0.830	0.443	0.206	0.112	0.054

Notes: This table presents the estimates obtained from our main regression model, presented in Equation (9), when estimated in a sample of college graduates, with health as the omitted field of study. We display the averages—of coefficients, robust standard errors, and p-values—across 100 regressions, each of which uses a different set of control functions obtained after parametric bootstrap over the ranked ordered logit estimation. The sample includes 140,690 observations of students who applied through the centralized admission system between 2004 and 2007 and graduated between 2007 and 2019. Panel (a) focuses on parameter estimates for Female students and Panel (b) focuses on Male students. In each panel, we report return estimates and selection on gains parameters, also referred to as match effects in the text. Underneath each estimate in parentheses is the bootstrapped standard error. The final two rows of the table report p-values from two additional statistical tests. The Δ Returns p-values correspond to the null hypothesis that the Female and Male returns are equivalent for a given column. The Δ Match Effects p-values correspond to a similar null hypothesis that the Female and Male match effects are equivalent for a given column.

Table B.7: Sample of College Graduates: Testing for Bias

	Panel (a): Earnings		
	Model 1	Model 2	Model 3
	(1)	(2)	(3)
Forecast Coefficient ϕ	0.693	0.891	0.991
	(0.069)	(0.087)	(0.097)
Covariates in the model:			
Controls	No	Yes	Yes
Control function	No	No	Yes
Tests:			
First-stage F- Statistic:	565	366	339
p-values			
Forecast bias ($\phi = 1$)	0.000	0.210	0.923
Overidentification	0.051	0.178	0.157
	Panel (b): Fertility		
	Model 1	Model 2	Model 3
	(1)	(2)	(3)
Forecast Coefficient ϕ	0.531	0.701	0.838
	(0.097)	(0.125)	(0.156)
Covariates in the model:			
Controls	No	Yes	Yes
Control function	No	No	Yes
Tests:			
First-stage F-statistic	823	508	383
p-values			
Forecast Bias ($\phi = 1$)	0.000	0.018	0.298
Overidentification	0.101	0.154	0.069

Notes: This table reports the results of tests for bias following Equation (4). Column 1 reports test results for an uncontrolled model; Column 2 augments the model with quadratic polynomials in PSU scores, cell fixed effects, and cohort fixed effects; Column 3 adds the control functions. The forecast bias test checks whether the forecast coefficient equals 1, and the overidentification test checks the IV modelâs overidentifying restrictions. Panel (a) considers logged earnings in 2019 as the outcome, and Panel (b) considers fertility (any child 12 years after college application) as the outcome.

C Child Penalties and Gender Norms

This section provides additional details on how we estimate the child penalty gaps and gender norms by field of study presented in the main body of the paper.

C.1 Child Penalty Gap

Following Kleven et al. (2019b), we estimate child penalties with the model:

$$\tilde{Y}_{it} = \sum_{k \neq -2} \beta_j D_{it}^k + \mu_a + \gamma_t + \nu_{it}, \tag{18}$$

where $D_{it}^k \equiv \mathbb{1}[t = c + k]$ is a binary variable that indicates the time relative to the period in which the first child was born c. μ_a are age-fixed effects that account for the life cycle effects on the outcome variable, and γ_t are time-fixed effects that control for temporal trends at the 6-month frequency. We omit the time binary variable corresponding to $t = -2$ since, in our case, it represents the period of child conception. Our primary outcome variable is labor market earnings in period t relative to the period of child conception t^*, i.e., $\tilde{Y}_{it} = Y_{it}/Y_{it^*}$.

We estimate the impacts of children on women and men separately and define the child penalty at event time t as $\hat{\beta}_t^m - \hat{\beta}_t^w$, which measures the percentage by which women fall behind men due to children. Figure C.1 shows the effects of parenthood on earnings across fields of study. To use a symmetric window, and since we have earnings starting only in 2015, we focus on college graduates whose first child was born in 2016 or 2017. This figure also reports the estimated child penalty at time $t = 4$. We have a more limited window of coverage compared to existing studies (Kleven et al., 2019a) and can estimate child penalties only up to 4 years after the birth of a child, so our estimates are relatively short-run. Nonetheless, most evidence shows a somewhat immediate drop in mothers' earnings, which changes minimally over time.

Appendix table C.1 complements previous evidence and reports the estimated child penalty before child conception, at time $t = -4$, and after it, at time $t = 4$.

C.2 Gender Norms

In this subsection, we discuss how we estimate our proxies for major-specific gender norms and report additional robustness checks that aid in interpreting the variation as due to norms instead of other labor market factors.

Estimation Details: In our analysis, we focus on graduates who had their first child after they obtained their college degree and on their partners (i.e., a couple). For each couple of parents, we use their labor market earnings from the third and fourth quarter before childbirth and focus on couples where both members earn positive income and are between 20 and 40 years old. We define $\textit{Relative Earnings}_i$ as $\frac{\textit{Woman Earnings}_i}{\textit{Woman Earnings}_i + \textit{Man Earnings}_i}$, where i indexes the couple, and $\textit{Woman Earnings}_i$ and $\textit{Male Earnings}_i$ are the total labor income of the will-be mothers and fathers, respectively.

Our proxy for the progressivity of gender norms is measured by the size of the cliff of $\textit{Relative Earnings}_i$ at 0.5. To gauge the magnitude of the drop at the point where the female starts to earn more than the male, we calculate the fraction of couples in each of twenty 0.05 relative income bins and project it on a discontinuity indicator while controlling for quadratic polynomials at each side of the 0.5 cutoff. Appendix Table C.2 presents the OLS estimates associated with the discontinuity indicators obtained from separate regressions for each of our nine fields of study. Panel (a) presents the results when we focus on female graduates (and their partners), and Panel (b) repeats the exercise but focuses on male graduates (and their partners). The estimates in Panel (a) are analogous to the estimates in Figure 4 in the main body of the paper, and the estimates in Panel (b) are analogous to the estimates in Figure C.2 which

reports the distribution of relative income when we consider male graduates (who eventually become fathers) instead of female graduates. Overall, we find similar variation in gender norms across fields of study, independently of whether we focus on male or female graduates (and their partners).[27]

Addressing Differences in Labor Market Attachment: Appendix Figure C.3 shows complementary evidence on the distribution of relative income using publicly available survey data from Chile. We observe hours worked in the survey data so we can restrict the sample to couples where both partners work at least 40 hours per week. This allows us to remove potential differences in labor market attachment associated with differences in the field of study. Reassuringly, the results reported in this figure are qualitatively similar to our preferred results.

Addressing Bunching Concerns: Since the seminal work by Bertrand et al. (2015), other scholars have reexamined this phenomenon in various contexts (Zinovyeva and Tverdostup, 2021; Binder and Lam, 2022; Hederos and Stenberg, 2022). Some have noted that cliff estimates may be sensitive to bunching at 0.5, potentially eliminating their presence in certain cases. Based on this evidence, they argue that cliff estimates are not reliable proxies for social or gender norms.

It is important to emphasize that, unlike some previous studies, we use high-quality administrative data that avoids common issues associated with survey data, such as i) self-reporting bias, ii) imputation for non-responses, and iii) rounding to reduce disclosure risk.

Nonetheless, we examine the prevalence of bunching at 0.5 to assess whether this effect drives our cliff estimates. Focusing on couples where the female's relative earnings fall between 0.45 and 0.55, as in Binder and Lam (2022), we find that around 7% of couples have equal earnings. However, as shown by Table C.3, the estimated cliffs remain similar to those reported in Table C.2 after excluding couples of parents with equal earnings.

Appendix Figure C.4 further investigates the potential impact of bunching on interpreting differences in cliffs as variations in gender norms. Panel (a) presents the cliffs when pooling data across fields of study, while Panel (b) shows the cliffs after removing cases where parents have equal earnings. Reassuringly, the cliffs are similar across panels. We further support our findings with Figure C.5 which replicates Figure 4 after excluding parents with equal earnings. Overall, we observe consistent variation in gender norms across fields of study. The robustness of our results strengthens our interpretation of the variation in cliffs as indicative of differences in gender norms.

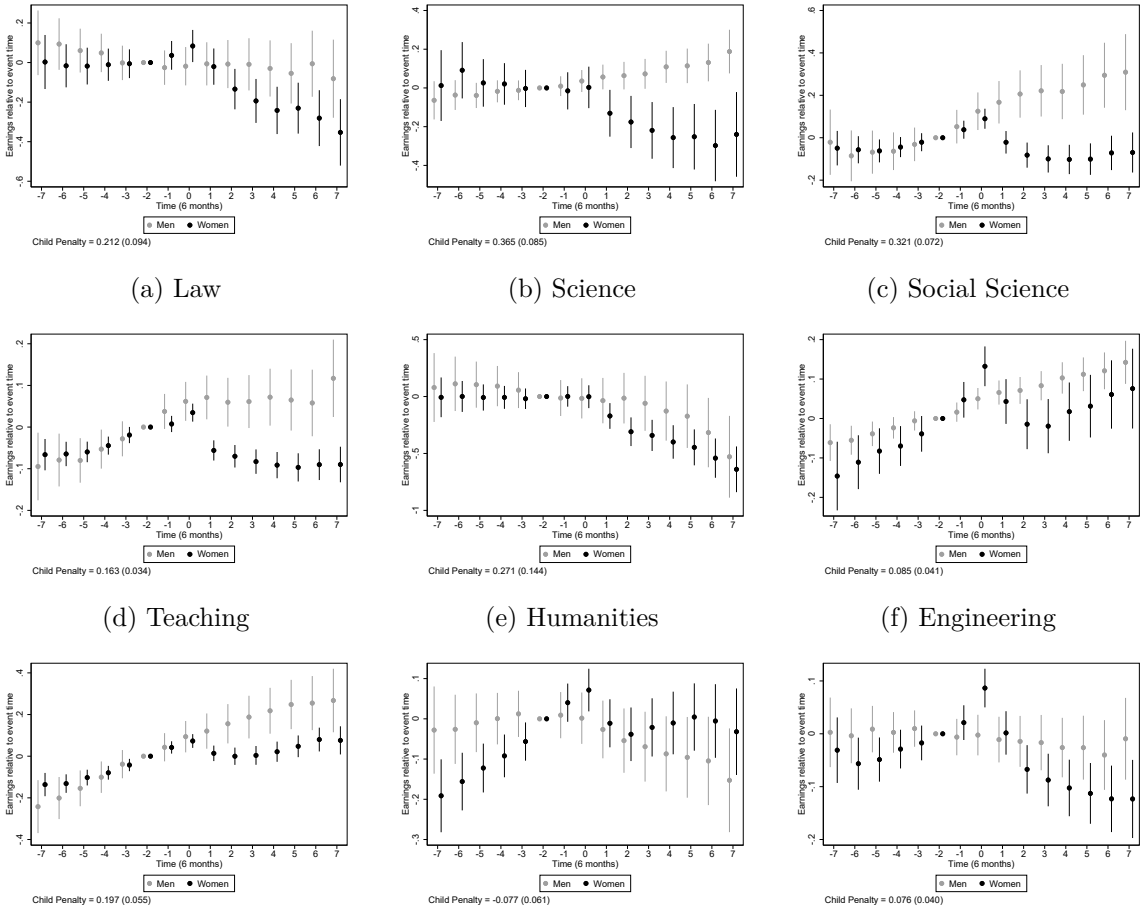
[27] Two exceptions are Humanities and Medicine, which show smaller and larger drops (at the point where the female starts to earn more than the male), respectively, when the graduate is male.

Table C.1: Child Penalties

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Male-Female Gap Before ($t = -4$)	0.060	-0.038	-0.020	-0.008	0.099	0.046	-0.021	0.093	0.032
	(0.064)	(0.058)	(0.049)	(0.023)	(0.098)	(0.028)	(0.038)	(0.042)	(0.027)
Male-Female Gap After ($t = 4$)	0.212	0.365	0.321	0.163	0.271	0.085	0.197	-0.077	0.076
	(0.094)	(0.085)	(0.072)	(0.034)	(0.144)	(0.041)	(0.055)	(0.061)	(0.040)
Number of observations	24,155	24,436	60,385	131,242	19,859	95,012	71,754	27,935	80,003

Notes: This table presents estimates obtained from estimating equation 18. Specifically, it shows the estimated child penalty (and its standard error) before child conception, at time $t = -4$, and after it, at time $t = 4$.

Figure C.1: Child Penalty by Field of Study



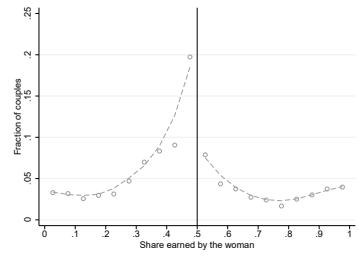
(a) Law (b) Science (c) Social Science

(d) Teaching (e) Humanities (f) Engineering

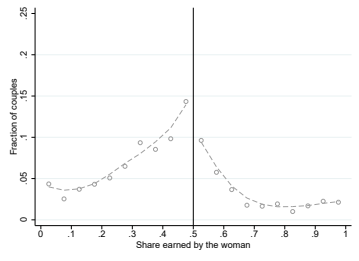
(g) Health (h) Medicine (i) Business

Notes: These figures show the percentage effects of parenthood on earnings across event time for men and women. Each panel corresponds to one of the nine fields of study considered in our analysis. We consider college graduates who had their first child between 2016 and 2017 and after their graduation. The figure also displays the average child penalty at the end of our time window (4 years after childbirth). Earnings are not conditional on employment status, and the effects include extensive and intensive margins.

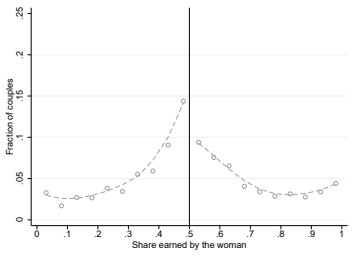
Figure C.2: Gender Norms by Field of Study



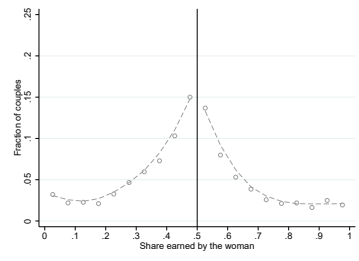
(a) Law



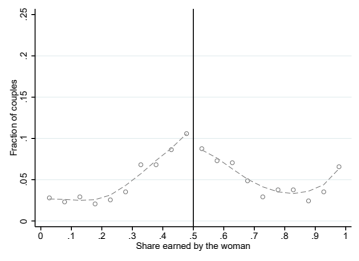
(b) Science



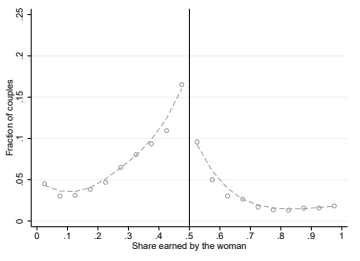
(c) Social Science



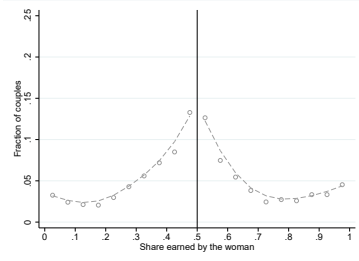
(d) Teaching



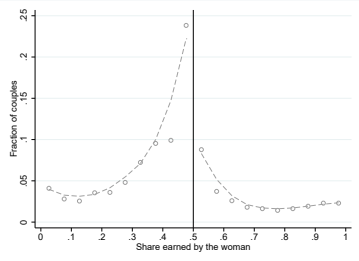
(e) Humanities



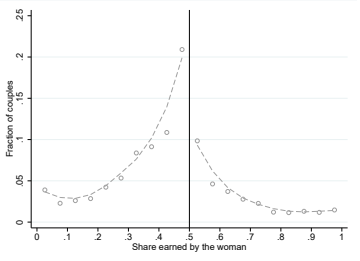
(f) Engineering



(g) Health



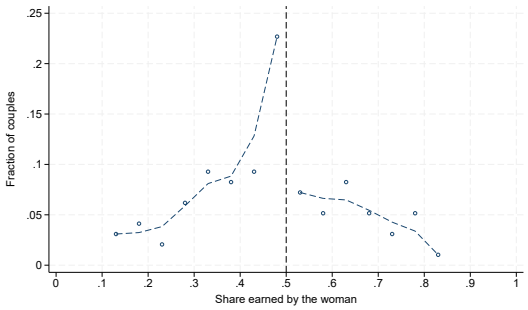
(h) Medicine



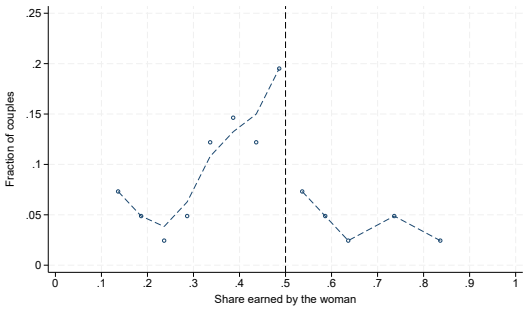
(i) Business

Notes: These figures report the distribution of relative income for males who graduated from different fields of study. The sample includes couples of parents who both earn positive income. We use the observation from the third and fourth quarters before childbirth. Each dot is the fraction of couples in a 0.05 relative income bin. The vertical line indicates the relative income share=0.5. The dashed line is the lowess smoother applied to the distribution, allowing for a break at 0.5.

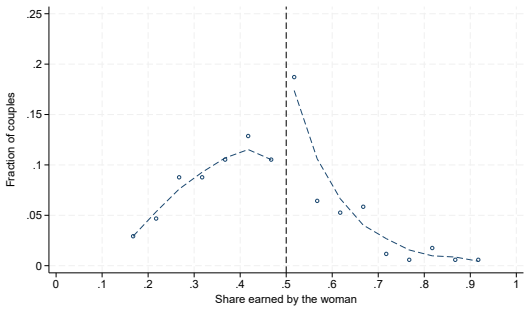
Figure C.3: Gender Norms by Field of Study: Full-time workers (CASEN survey)



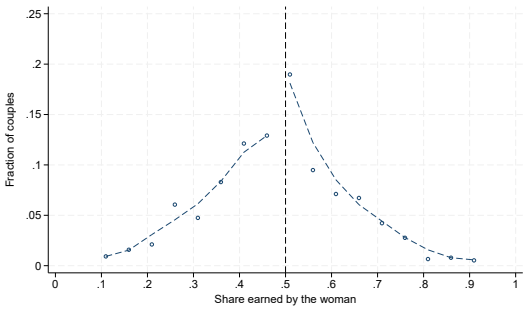
(a) Law



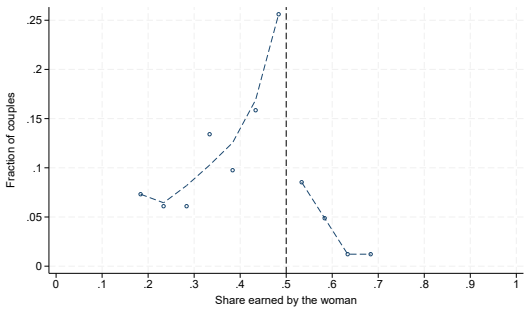
(b) Science



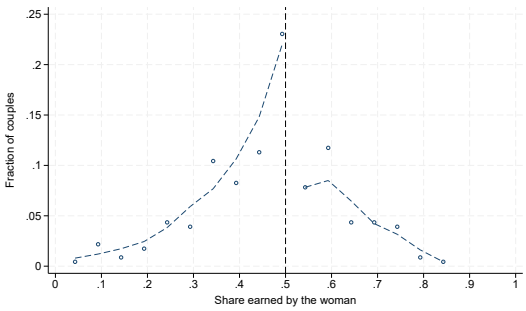
(c) Social Science



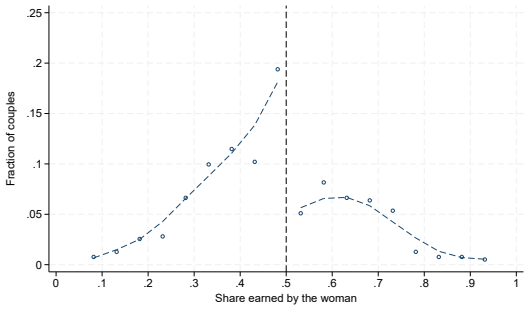
(d) Teaching



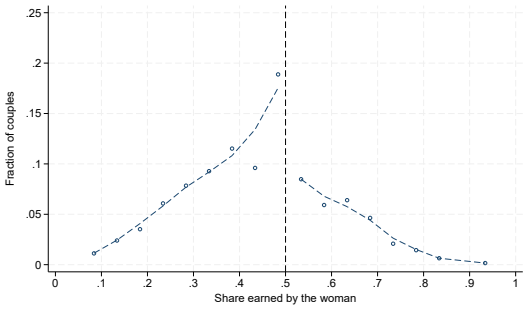
(e) Humanities



(f) Engineering



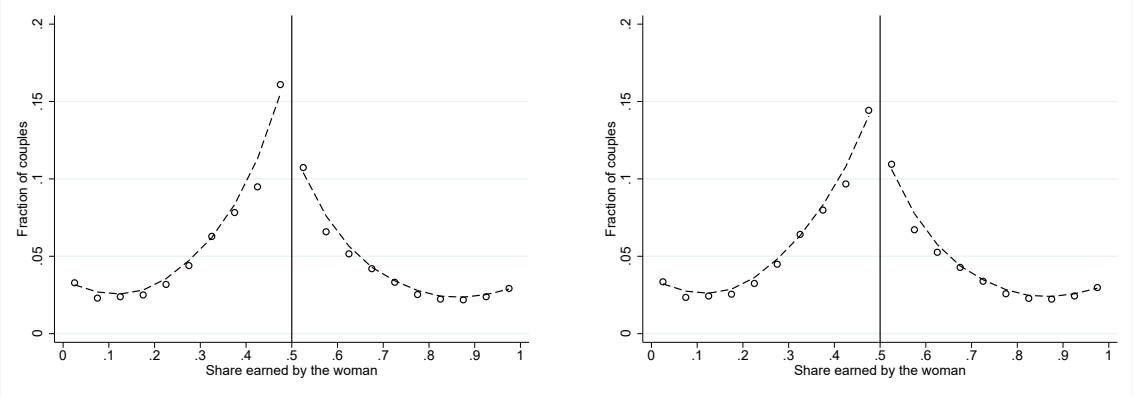
(g) Health and Medicine



(h) Business

Notes: These figures report the distribution of relative income separately for females who graduated from different fields of study. The sample includes couples in the household survey CASEN, where both earn positive income and have full-time jobs (work more than 40 hours a week). Each dot is the fraction of couples in a 0.05 relative income bin. The vertical line indicates the relative income share=0.5. The dashed line is the lowess smoother applied to the distribution, allowing for a break at 0.5. Since the survey records do not separate Medicine from Health, both fields are embedded into the Health figure.

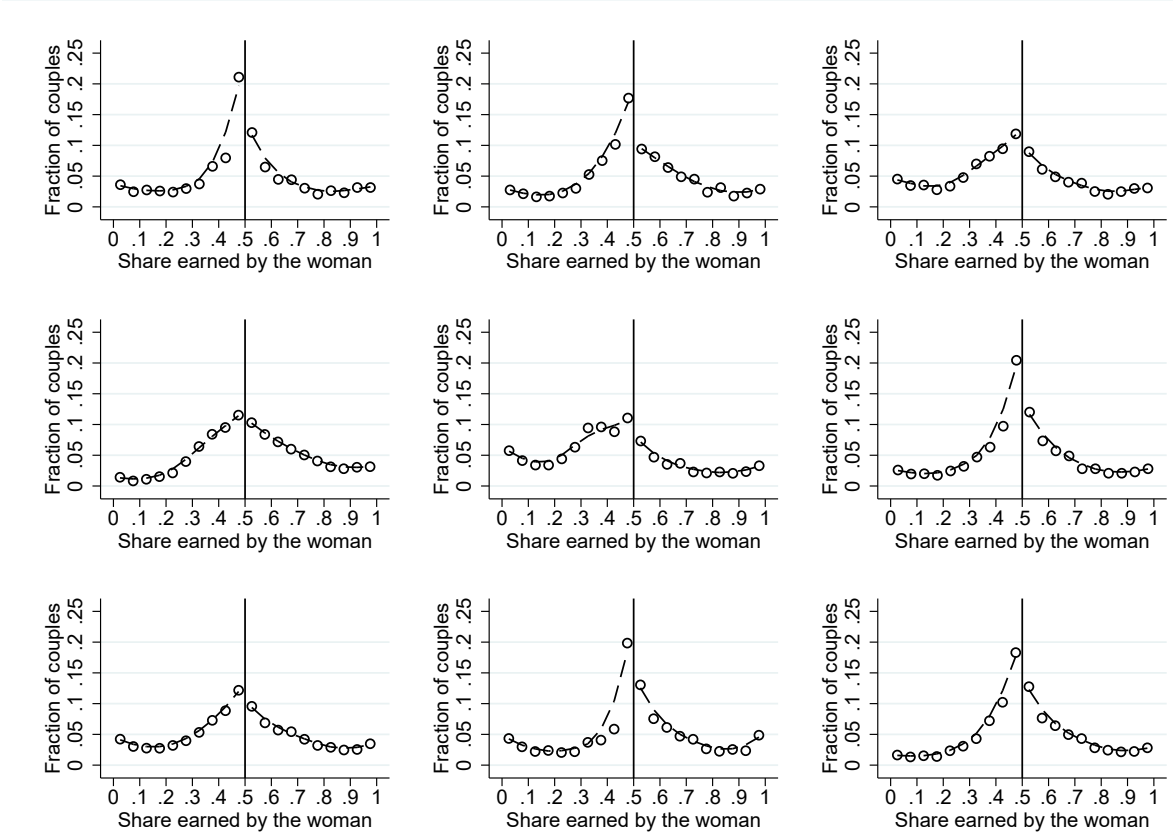
Figure C.4: Gender Norms: Removing bunching at 0.5 (pooled)



(a) All (b) No Bunching

Notes: These figures report the distribution of relative income. The sample includes couples of parents who both earn positive income. Panels (a) consider all parents, while Panel (b) removes cases in which parents report the same income. We use the observation from the third and fourth quarters before childbirth. Each dot is the fraction of couples in a 0.05 relative income bin. The vertical line indicates the relative income share=0.5. The dashed line is the lowess smoother applied to the distribution, allowing for a break at 0.5.

Figure C.5: Gender Norms: Removing bunching at 0.5



Notes: These figures report the distribution of relative income for females who graduated from different fields of study. The sample includes couples of parents who both earn positive income but excludes cases in which parents report the same income. We use the observation from the third and fourth quarters before childbirth. Each dot is the fraction of couples in a 0.05 relative income bin. The vertical line indicates the relative income share=0.5. The dashed line is the lowess smoother applied to the distribution, allowing for a break at 0.5.

Table C.2: Gender Norms

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel (a): Female graduates (and their partners)								
Break	-0.104	-0.095	-0.048	-0.023	-0.054	-0.099	-0.036	-0.080	-0.080
	(0.034)	(0.019)	(0.007)	(0.006)	(0.014)	(0.028)	(0.005)	(0.039)	(0.021)
	Panel (b): Male graduates (and their partners)								
Break	-0.109	-0.046	-0.033	-0.019	-0.012	-0.078	-0.008	-0.143	-0.111
	(0.019)	(0.013)	(0.011)	(0.011)	(0.012)	(0.011)	(0.010)	(0.026)	(0.019)

Notes: This table reports the estimated discontinuity in the distribution of relative earnings of individuals from different fields of study. The dependent variable is the fraction of couples. We calculate the fraction of couples in each of the twenty 0.05 relative income bins and project it on a discontinuity indicator while controlling for a quadratic polynomial at each side of the 0.5 cutoff. This table presents the OLS estimates of the discontinuity indicators obtained from separate regressions for each field of study. Panel (a) considers mothers who graduated from a field of study and their couples, and Panel (b) considers fathers who graduated from a field of study and their couples. We use the observations from the third and fourth quarters before childbirth and focus on couples where both members earn positive income.

Table C.3: Gender Norms: Removing bunching at 0.5

	Law	Science	Social Science	Teaching	Humanities	Engineering	Health	Medicine	Business
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel (a): Female graduates (and their partners)								
Break	-0.075	-0.070	-0.040	-0.022	-0.049	-0.070	-0.032	-0.041	-0.056
	(0.027)	(0.011)	(0.007)	(0.006)	(0.014)	(0.020)	(0.005)	(0.029)	(0.014)
	Panel (b): Male graduates (and their partners)								
Break	-0.087	-0.037	-0.027	-0.018	-0.007	-0.060	-0.001	-0.106	-0.091
	(0.013)	(0.012)	(0.009)	(0.011)	(0.012)	(0.010)	(0.009)	(0.017)	(0.014)

Notes: This table reports the estimated discontinuity in the distribution of relative earnings of individuals from different fields of study. The dependent variable is the fraction of couples. We calculate the fraction of couples in each of the twenty 0.05 relative income bins and project it on a discontinuity indicator while controlling for a quadratic polynomial at each side of the 0.5 cutoff. This table presents the OLS estimates of the discontinuity indicators obtained from separate regressions for each field of study. Panel (a) considers mothers who graduated from a field of study and their couples, and Panel (b) considers fathers who graduated from a field of study and their couples. We use the observations from the third and fourth quarters before childbirth and focus on couples where both members earn positive income but exclude cases in which both parents report the same income.

D Counterfactual Details

This section reports additional details related to the counterfactual policies we analyze in Section 6. We present the deferred acceptance algorithm, offer details on the machine learning model used to predict graduation probabilities, and describe the bootstrap procedure used in our counterfactual analysis.

D.1 Deferred Acceptance Algorithm

We use the following algorithm to produce an outcome of the student-proposing Deferred Acceptance algorithm.

Step 1) Each student proposes to her first choice. Each program tentatively assigns seats to its proposers one at a time, following their priority order based on the weighted score. Students are rejected if no seats are available at the time of consideration.

In general, in

Step k) Each student who was rejected in the previous step proposes to her next best choice. Each school considers the students it has tentatively assigned together with its new proposers and tentatively assigns its seats to these students one at a time following the program's priority order. The student is rejected if no seats are available when she is considered.

The algorithm terminates either when there are no new proposals or when all rejected students have exhausted their rank order list of preferences.

When introducing quotas, we apply a precedence order of processing first the open seats and then the gender specific seats. Following Dur et al. (2018), this order of precedence maximizes the impact of quotas.

D.2 Prediction Accuracy and Summary Statistics

We use statistical learning methods to predict graduation between the time of being accepted and up to twelve years after application. In particular, we use a Random Forest algorithm to generate predictions. See Hastie et al. (2009) for an introduction to the algorithm and Breiman (2001) for the original description. Predictors include the income quintile of the student at the time when they applied to college, type of high school and health insurance, PSU scores (Math, Language, History, and Science), parent's educational attainment, and the control functions from the choice model.

We predict graduation separately for each field, generating a predicted graduation probability at each major for each student. For each field, the prediction sample consists of students in the admissions process in 2007 who were accepted in a field of study. In this sample, we use a Random Forest algorithm that fits 500 trees separately, each of which draws three random variables from the included covariates. The models incorporate various features, including demographic characteristics, entrance exam scores, preferences revealed through rank-ordered lists and all other relevant attributes that are also part of Equation (9). This approach allows us to capture complex, non-linear relationships between student characteristics and their likelihood of graduating from a given major. To improve model performance, we fine-tune the parameters for each major separately, optimizing them to minimize the out-of-bag (OOB) Brier score. The Brier score, which measures the accuracy of probabilistic predictions, is particularly useful in this context as it reflects both the calibration and sharpness of our predictions. By minimizing the OOB-Brier score, we ensure that our random forest model provides reliable and well-calibrated predictions of graduation probabilities across different majors.

D.3 Counterfactual Bootstrap Details

The standard errors for Table 6 were estimated by a bootstrap procedure. To construct each of the b bootstrap samples we sample with replacement n students, where n is the number of students in the data for counterfactuals. For each student sampled, we take their preferences ($\mathbf{R}_i$) and their scores ($\mathbf{s}_i$). For each bootstrap sample of students $b \in \{1, \ldots, B\}$, we compute:

- Assignment ψ^b: We run the deferred acceptance algorithm for the counterfactual using the bootstrapped sample of students, their preferences and scores ($\mathbf{R}^b, \mathbf{s}^b$), fixed capacities ($\mathbf{q}$) and quota policy ($\omega$) for the corresponding counterfactual.
- Draw returns estimates: We use draws from a parametric bootstrap described in section 4. The bootstrap takes the asymptotic distribution of the estimates of preferences to draw control function and estimate the main returns model given by Equation (9).
- Compute returns and statistics: We use the bootstrapped parameters of the returns equation, the counterfactual assignment, and the predicted graduation probabilities to compute expected returns for each student.

Table 6 presents the median across 100 bootstrapped samples and the confidence interval corresponds to percentiles 2.5 and 97.5.