The gender pay gap in Brazil: It starts with college students' choice of major

Gender and Diversity Division

Monserrat Bustelo
Suzanne Duryea
Claudia Piras
Breno Sampaio
Giuseppe Trevisan
Mariana Viollaz

TECHNICAL NOTE Nº IDB-TN-02099

January 2021

The gender pay gap in Brazil: It starts with college students' choice of major

Monserrat Bustelo
Suzanne Duryea
Claudia Piras
Breno Sampaio
Giuseppe Trevisan
Mariana Viollaz

January 2021

Inter-American Development Bank

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library
The gender pay gap in Brazil: it starts with college students' choice of major / Monserrat Bustelo, Suzanne Duryea, Claudia Piras, Breno Sampaio, Giuseppe Trevisan, Mariana Viollaz.
p. cm. — (IDB Technical Note ; 2099)
Includes bibliographic references.
1. Wages-Women-Brazil. 2. Wage differentials-Brazil. 3. Women science students-Brazil. 4. College majors-Brazil. 5. Pay equity-Brazil. 6. Equal pay for equal work-Brazil. I. Bustelo, Monserrat. II. Duryea, Suzanne. III. Piras, Claudia. IV. Sampaio, Breno. V. Trevisan, Giuseppe. VI. Viollaz, Mariana. VII. Inter-American Development Bank. Gender and Diversity Division. VIII. Series.
IDB-TN-2099

JEL codes: J16, J31, J38

Keywords: gender wage gap, major choices, STEM, decomposition analysis, unconditional quantile regression

http://www.iadb.org

The gender pay gap in Brazil: It starts with college students' choice of major

Monserrat Bustelo* Suzanne Duryea† Claudia Piras‡

Breno Sampaio§ Giuseppe Trevisan¶ Mariana Viollaz[1]

December 2, 2020

Abstract

We herein discuss how college major choice affects gender wage gaps by highlighting the role that STEM majors play in explaining the gender wage gap in a developing country. We focus on a Latin American country where a systematic analysis of the interaction between students' choice of college major and the gender wage gap is currently lacking. We take advantage of a very unique dataset of college students from the Universidade Federal de Pernambuco (UFPE), Brazil, to decompose the raw gender gap in hourly wages into one component that can be explained by differences in endowments between men and women as well as a second or residual component that reflects gender differences in the prices of market skills. We implement the commonly applied decomposition approach at the wage distribution's mean and a decomposition procedure that considers variations across the wage distribution. Our results reveal that the majors that women and men select explain 50% of the gender wage gap at the mean, and STEM majors contribute to 30% of this difference. When examining different percentiles of the wage distribution, we find that the selection of a major is more important at the middle of the distribution than at the bottom or top.

JEL codes: J16, J31, J38.

Keywords: gender wage gap, major choices, STEM, decomposition analysis, unconditional quantile regression

* Inter-American Development Bank, e-mail: MonserratB@iadb.org

† Inter-American Development Bank, e-mail: SuzanneD@iadb.org

‡ Inter-American Development Bank, e-mail: ClaudiaPI@iadb.org

§ Department of Economics, UFPE, Brazil, e-mail: Breno.Sampaio@ufpe.br

¶ Department of Economics, UFPE, Brazil, e-mail: Giuseppe.Trevisan@gmail.com

[1] Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), IIE-FCE, Universidad Nacional de La Plata, e-mail: mviollaz@cedlas.org

1 Introduction

Wage gaps between men and women characterize labor markets across the globe. The literature has linked these gaps to several factors that range from traditional drivers, such as gender differences in education levels and test scores, disparities in labor attachment and experience due to housework and family responsibilities borne by women, and wage penalties associated with motherhood (Blau & Khan, 2017), to a new set of factors related to norms, psychological attributes, and non-cognitive skills (Bertrand, 2011; Blau & Kahn, 2017).

The choice of one's field of study and major within that field can play an important role, as men often sort into majors that pay more than those chosen by women. The influence of one's choice of major on gender wage gaps has not been thoroughly studied, which is likely due to data scarcity in that most labor force surveys do not include direct measures of college majors. Further, the available evidence exclusively focuses on developed countries. Goldin and Katz (2008), Black, Haviland, Sanders, and Taylor (2008), and Carnevale, Smith, and Gulish (2018) report that college majors are important determinants of the wage gap between college-educated men and women in the U.S., while similar results are reported by Machin and Puhani (2003) for Germany and the U.K. and by Piazzalunga (2018) for Italy.

The study contributes to this scarce literature by exploring the role that college majors play in explaining the gender wage gap in a developing country. We focus on Brazil, where, similar to other Latin American countries, a systematic analysis of the interaction between majors and the gender wage gap is still lacking. Brazil is one of the most unequal countries in the region according to income inequality measures (e.g., the Gini coefficient), and differential earnings by gender—jointly with race—contribute to this aggregate inequality (Duryea et al., 2019). Despite their completion of a higher average number of years of schooling than their male counterparts, women in Brazil earn less than men, thereby exhibiting the largest gap in wages in the region after controlling for individual and family characteristics (Ñopo, 2012). Following the recent trends in labor markets, we highlight the role that science, technology, engineering, and mathematics (STEM) majors play in the analysis. STEM skills are considered crucial to thrive in the digital era, and the available evidence indicates that women are represented less than men in STEM disciplines and occupations (United Nations Educational, Scientific, and Cultural Organization [UNESCO], 2017a; UNESCO/UIS, 2017; ILOSTAT, 2018).[1]

[1] Data for the Latin American region reports that women represent 60% of all college graduates; however, merely 30%

We herein take advantage of a unique dataset of college students from the Federal University of Pernambuco (UFPE), a high-quality public flagship university in Brazil, to provide a comprehensive analysis by following 8,773 students from the 2002–2004 cohort. We perform a decomposition analysis to explore the raw gender gaps in wages along one component that can be explained by differences in endowments between men and women and a second or residual component that reflects gender differences in the prices of market skills. We implement the commonly applied approach at the wage distribution's mean following the Oaxaca–Blinder decomposition (Oaxaca, 1973; Blinder, 1973) as well as a decomposition procedure that considers variations across the distribution of wages. The latest approach is based on unconditional quantile (UQ) regressions proposed by Firpo, Fortin, and Lemieux (2009) and implemented by other authors, such as Chi and Li (2008) as well as Boaz Anglade and Deere (2017). This approach explains the different gender wage gaps we found at various points of the wage distributions in our data; for instance, men earn 20% more than women in the 10th percentile, 35% more in the 50th percentile, and 29% more in the 90th percentile.

Our findings point out that gender segregation in fields of study might perpetuate gender inequalities in the labor market. The varied distribution of females and males into college majors is a key determinant of the gender wage gap in Brazil. Majors that tend to lead to higher-paying jobs are dominated by male college students, whereas majors that feed into lower-paying jobs are dominated by female college students. The remainder of the paper is organized as follows. Section 2 presents the institutional background, and Section 3 discusses data sources. Section 4 discusses our empirical approach, Section 5 presents our main findings, and Section 6 concludes our report.

2 Institutional Background

2.1 UFPE: A flagship university

UFPE was founded in 1948 and is currently the major flagship university of North and Northeast Brazil. UFPE does not charge tuition fees and is regarded as employing the most qualified professors in the region; hence, the school is known for its focus on academic training. According to the Times Higher Education World University Ranking of 2018, UFPE is the fourteenth best university in the

graduate from a STEM major, and within STEM, merely 28% graduate from technology or engineering (UNESCO, 2017b). In the labor market, women are underrepresented in STEM jobs. For instance, only 1% of employed women work in the information and communication sector, while that figure is 1.9% for men (ILOSTAT, 2018).

country.[2] Among the top thirty universities in this ranking, merely four are private, and most are federal. UFPE presents itself as an interesting case study because it strongly represents the 36 public universities operating in the country in terms of admission policy, the quality of students admitted, academic environment, and economic prospects during adulthood.

Different from public universities in the U.S., a centralized entrance exam named the Unified Selection System (SISU) and organized by the Ministry of Education was only largely adopted in Brazil in 2015. Before that, most entrance exams were institution specific, thereby creating a major barrier that deterred students (especially those from lower socio-economic groups) from applying to other federal universities located in the country, as exams were administered locally.[3] For instance, a student living in Northeast Brazil had to travel to the State of Rio de Janeiro, which is located in the Southeast, to take the entrance exam for the Federal University of Rio de Janeiro. Further, local inhabitants' access to higher education in Brazil is not facilitated by federal, state, or local policies. As such, the students admitted to federal universities usually originate from wealthy families.

2.2 The admission process

Students are admitted to UFPE through an entrance exam called the Vestibular that is held once per year and is divided into two stages. The first stage (hereafter, "round 1 score") has a broad scope and covers a variety of subjects, including Portuguese, mathematics, physics, chemistry, history, geography, biology, and other foreign languages (e.g., English, Spanish, and French). The second stage comprises Portuguese, a foreign language, and three other subjects that directly relate to the major chosen by each student. The final entrance test score is a weighted average of the first- and second-stage scores.

Similar to other countries in Latin America, all students in Brazil must decide their majors before entering college, which means all students applying for admission to UFPE must declare at the time of application to which major they choose to apply among those offered by the institution. One's chances of acceptance therefore not only depend on the entrance test score, but are also

[2] An evaluation performed by the Ministry of Education based on a vast range of inputs related to infrastructure, the quality of majors and teachers, management effectiveness, and students' academic performance ranks UFPE among the top twenty universities in the country. The university is currently in the second percentile on the distribution of institutional quality. More information about the evaluation process can be found at: http://portal.mec.gov.br

[3] UFPE, alongside most universities in the country, started fully adopting the new national centralized entrance process (i.e., SISU) in 2015. This centralized process caused many changes, the two most important being that institution-specific exams were largely abolished and that students were thereafter allowed to select up to two major choices per institution when applying to higher education. The system then assigns students to programs through a deferred acceptance algorithm that sorts candidates into their most preferred major for which they qualify.

conditional upon the major indicated while applying. Solely a small fraction of the original candidates per major are admitted due to the limited number of available seats. Of about 50,000 students who apply for UFPE every year, roughly 5,000 are accepted. Unlike the U.S. system, a student is not allowed to change majors unless they retake the exam as a candidate for the new major in the following year.

3 Data

Our data originates from three different sources. The first two datasets contain academic information concerning UFPE students. The Comissão de Processos Seletivos e Treinamentos (COVEST) admissions data provides detailed information about every UFPE applicant, including their entrance test scores and socioeconomic characteristics. The Sistema de Informações e Gestão Acadêmica (SIG@) data provides detailed information about every student enrolled at UFPE, including their chosen major program, enrollment status, and academic situation (i.e., active, graduated, or dismissed). The third dataset, the Annual Social Information Report (Relação Anual de Informações Sociais [RAIS]) from the Ministry of Labor, is an administrative dataset that contains information about the employees of every registered firm in Brazil.

Our sample comprises students who enrolled at UFPE during the years 2002, 2003, and 2004 and who were aged between 18 and 30 years.[4] Those three cohorts involve a total population of 14,093 students of all ages. After matching students with the tax records of those formally employed after 13 years of their matriculation at UFPE and after eliminating observations with missing values, we end up with a sample of 8,773 observations that represent 62% of all enrolled students on those three cohorts and about 70% of all enrolled students aged between 18 and 30 years.[5]

3.1 Academic information

COVEST data houses information about students at the time they apply to UFPE, including their round 1 scores, which we herein use as a measure of an individual's cognitive skills, and a full set of background information that includes an individual's gender, date of birth, enrollment in a public secondary school (if applicable), mother's education, and family's income bracket. We incorporate

[4] About 92% of all enrolled students are aged between 18 and 30 years, while the remaining 8% of students are aged between 31 and 70 years. The former range is being focused on to minimize the roles that work experience and other distortions play in the labor market returns associated with advanced age. The full age range does not affect our results which are available upon request.

[5] Thus, the other 30% of students are external to the labor market, self-employed, or employed in the informal sector.

all these predetermined background characteristics into our empirical analysis as control variables.

The round 1 score refers to an exam common to all students regardless of which major programs they applied for, thus rendering the scores comparable across students of the same cohort. Since regressions account for cohort-fixed effects, we do not standardize this variable. Date of birth is applied to firstly obtain each student's age and secondly divide the age groups into four dummies: 18–19 years, 20–21 years, 22–23 years, and 24–30 years. Regarding the type of secondary school attended, we generated a binary variable that indicates whether or not a student completed all grades at public schools. The student's mother's education level is divided into three categories: secondary incomplete or lower, completed secondary or incomplete college, and completed college or higher. Family income is organized into brackets raging from: 0–300 reais, 301–1,000 reais, 1,001–1,500 reais, 1,501–2,000 reais, and 2,001 reais or higher.[6]

The SIG@ data allows the identification of the major programs in which UFPE students have enrolled. UFPE offered 64 major programs during the 2002–2004 period,[7] and we grouped those majors into eight fields of knowledge:[8] STEM, social sciences, humanities, physical education, health, communication, music, and visual arts; STEM fields involve math, statistics, engineering, computation, biology, earth science, physics, and chemistry majors. Finally, we use information regarding each student's academic situation (i.e., active, graduated, or dismissed) to explore issues concerning the sample selection.

3.2 Labor market information

The RAIS data provides universal information for every employee working in registered firms in Brazil. Formal employees correspond with most of the workforce in Brazil and do not include firms' owners and domestic workers with no registration. All operating firms from public and private sectors are annually required to report their monthly payroll disbursals, hours of work (contract) per week, dates of admission (and dismissal), the standard classification code for each occupation, and some basic demographics, such as gender, age, and education level.

We have access to a restricted version of the RAIS data that enables us to identify individuals using unique identifiers—that is, their Social Security numbers—to match labor market data with students' full academic information provided by UFPE. A remarkable advantage of having an

[6] The sociodemographic questionnaire provided by COVEST uses those intervals by default in all cohorts.
[7] Currently, the university offers 99 programs. This number does not include those distance learning programs or courses targeted at high school teachers.

administrative dataset of this caliber regarding information about wages is that the latter is very unlikely to present measurement errors since RAIS does not depend on self-reported data. To achieve our goal of understanding the role that major choices shall play in the future gender wage gap, students' labor information was obtained from RAIS for the 2014–2016 rounds.

The RAIS data provides each individual's average hourly wage, which is calculated by considering all jobs the student held during the year and then deflated to the December 2006 level using the extended consumer price index. Average hourly wages were log-linearized to explore percentage gaps between women and men. This outcome was measured after thirteen years of each student cohort's matriculation at the university; for example, the academic profiles of the cohort enrolled in 2002 are matched with the RAIS data from 2014.

We additionally include the indicator variable of "one's status as a public worker" as a control. This status is relevant in the Brazilian context given that the public sector is known both for paying the highest salaries in the formal labor market and for implementing an extremely competitive selection process. As a consequence, the most skilled individuals tend to select public jobs. Regarding the workforce that possesses tertiary education, the spots in the public sector are more likely to require workers from specific fields, which implies that students from some major programs are more prone to filling these vacancies than others. The control variable aims to attenuate the influence of this selection type when investigating the role that college majors play in gender wage gaps.

We construct two dummies to measure the highest education level obtained by a student: graduation (or lack of graduation) from college or the attainment of a master degree or PhD. We obtained students' education statuses from RAIS since they may simultaneously be enrolled in other colleges or have dropped out from UFPE.

3.3 Summary statistics

Table 1 presents our sample's descriptive statistics by gender. About 74% of all enrolled applicants aged 18 to 30 years were working in a formal occupation. Overall, men and women slightly differ in their academic and labor market characteristics. On average, a gender gap exists in hourly wages in favor of men. Women are more likely than men to acquire positions in the public sector and to possess post-graduate qualifications.[8] In reference to their academic information, men on average achieve higher round 1 scores (i.e., university entrance scores) than women. Most students

[8] Women's graduation rate is higher than that of men, at 74% versus 55%.

completed at least part of their secondary-level education in non-public schools, wherein the share is larger among females. Students are predominantly young (up to 21 years of age), especially among females; young students represent about 72% of the total female sample.

Overall, males possess more desirable backgrounds than females in that their mothers are have higher levels of education and they originate from families that earn a higher average income. Regarding the major program options, the percentage of females who choose STEM careers compared to males is very pronounced—that is, 16% against 43%, to the females' detriment. Women primarily choose programs related to humanities, health, and communication, which on average embrace less competitive programs offered by UFPE.

Figure 1 presents some suggestive evidence concerning the importance of one's major choice to explain the gender wage gap in Brazil. The figure shows the proportion of male students by major and the average hourly wage of those students who were working in the formal labor market thirteen years after applying to UFPE. Two clear findings emerge from this figure. Firstly, higher-paying majors tend to be those in which the proportion of male students is larger, and secondly, STEM majors generally correspond with the highest-paying jobs. The next subsection explains how major choice and other characteristics can contribute to the gender wage gap.

4 Empirical Approach

To investigate wage differentials between male (m) and female (f) workers, we run Mincer's (1974) human capital earnings function using the following equation:

$$w_{ij} = X'_{ij}\beta_j + \varepsilon_{ij} \tag{1}$$

where $j \in (m, f)$, w_i is the wage observed for worker i, and X_i is a matrix containing a set of characteristics observed for each worker at the time they applied to college and some others related to each worker's job; the error term *Eij* captures other unobserved characteristics. By building upon these estimated wage regressions, the raw gender wage gap is then decomposed into one component that can be explained by differences in mean endowments between men and women as well as a second or residual component that reflects gender differences in the prices of market skills. This procedure, known as the Oaxaca–Blinder decomposition (Oaxaca, 1973; Blinder, 1973), can be obtained by differencing equation (1) across gender groups and taking their expected values such that:

$$E\left[\bar{w}_m - \bar{w}_f\right] = \left(\bar{X}'_m - \bar{X}'_f\right)\hat{\beta}_f - \bar{X}'_m(\hat{\beta}_m - \hat{\beta}_f) \tag{2}$$

where $\bar{X}'_j$ is the average level of earnings-related characteristics for group j. The first term on the right-hand side (i.e., the explained component) can be interpreted as the gap between males and females due to gender differences in their characteristics (weighted by price vector $\hat{\beta}_s$ for females), while the second term (i.e., the unexplained or residual component) can be interpreted as the difference in the prices of these characteristics weighted by men's mean characteristics ($\bar{X}'_m$).

Equation (2) delivers the wage decomposition at the mean. While understanding mean effects is important, this standard version of the Oaxaca–Blinder decomposition fails to consider variations across the wage distribution. This important shortcoming motivated the development of new decomposition methods for distributional statistics other than the mean (Fortin, Lemieux, & Firpo, 2010).[9] In this paper, we apply the decomposition procedure based on the UQ regression proposed by Firpo et al. (2009). The method is based on re-centered influence function (RIF) regressions that can be implemented within a quantile regression approach, thereby allowing a detailed decomposition at any quantile of the distribution. From a policy perspective, this procedure allows us to investigate whether or not the relative importance that one's major choice poses to the gender gap differs along the wage distribution.

The RIF for the τth quantile is defined as:

$$RIF(w, q_\tau) = q_\tau + \frac{[\tau - I(w < q_\tau)]}{f_w(q_\tau)} \tag{3}$$

where $I(w < q_\tau)$ is an indicator function that equals 1 if the value of the outcome is below quantile q_τ and 0 otherwise; $f_w(q_\tau)$ represents the marginal density of the outcome at point q_τ. An important property of the RIF regressions that makes them ideal for decomposition methods is that $E[RIF(w, q_\tau)|X] = X\beta_\tau$, which allows that the UQ estimates be applied to perform the standard Oaxaca–Blinder decomposition at any quantile τ. This decomposition is executed by setting $RIF(w, q_\tau)$ rather than w as the dependent variable:

$$E\left[\overline{RIF}(w_m, \hat{q}_{m\tau}) - \overline{RIF}\left(w_f, \hat{q}_{f\tau}\right)\right] = \left(\bar{X}'_m - \bar{X}'_f\right)\hat{\gamma}_\tau + [\bar{X}'_m\left(\hat{\beta}_{m\tau} - \hat{\gamma}_\tau\right) - \bar{X}'_f\left(\hat{\beta}_{f\tau} - \hat{\gamma}_\tau\right)] \tag{4}$$

[9] See, for instance, the works of Juhn, Murphy, and Pierce (1993), DiNardo, Fortin, and Lemieux (1996), Donald, Green, and Paarsch (2000), Machado and Mata (2005), Autor, Katz, and Kearney (2005), and, more recently, Chernozhukov, Fernandez-Val, and Melly (2012).

where $\hat{q}_{j\tau}$ are the τth quantiles of the marginal distributions of w_j; $\hat{\beta}_{jt}$ are the OLS estimates of regressions $RIF(w_i, q_\tau) = X_i'\beta_{j\tau} + \varepsilon_i$ for both genders; and $\hat{\gamma}_\tau$ is the nondiscriminatory wage structure at the τth quantile from the pooled RIF regression. The first term on the right-hand side of equation (4) captures the explained component of the decomposition, while the second and third terms represent the unexplained component at the τth quantile. This RIF-based procedure is capable of further dividing the explained and unexplained components into each explanatory variable's contribution to the gap at different areas of the unconditional distribution of earnings.

5 Results

We estimate wage regressions separately for men and women both with and without including major-fixed effects. We then decompose the gender gap in the explained and unexplained components following both the Oaxaca–Blinder procedure and the approach proposed by Fortin et al. (2010). The main objective is to learn which part of the wage gap between women and men can be attributed to differences in major selection and returns to majors.

We possess unique information for performing this exercise. Firstly, as highlighted above, we can follow the cohort of students enrolled at UFPE between 2002 and 2004 over thirteen years. Secondly, by incorporating information on standardized entrance test scores, we can control for a sufficient measure of ability. This measure is important because selection on unobserved characteristics can affect labor market participation both across genders and over time. Thirdly, we have obtained detailed information on the college's majors, thereby allowing us to test for career selection's contribution toward explaining the gender wage gap and to analyze the importance of specific majors, such as those related to STEM fields.

5.1 Wage regressions

5.1.A Regressions at the mean

Table 2, columns 1 and 2 report results without controlling for major effects, while columns 3 and 4 present the results when major-fixed effects are included. In general, the factors that are positively related to hourly wages are one's status as a public servant, a high education level (e.g., master degree or PhD), the proxy of ability (e.g., round 1 score), one's former study at a public secondary school, one's belonging to the youngest age group at enrollment, mother's completion of her college education, and family income. These factors hold among women and men both

with and without major controls.

When comparing estimates with and without major controls, our results reveal that sorting into majors biases the contribution of the variables included in the wage regressions. Controlling for major-fixed effects leads to an increase in the coefficient of one's status as a public servant and one's possession of a master degree or PhD as well as a decrease in the returns of the ability measure, one's former study at a public secondary school, one's belonging to the youngest cohort, and a high family income.

5.1.B Unconditional quantile regressions of hourly wages

Table 3 illustrates the UQ regression estimates at the 10th, 50th, and 90th percentiles. Panel A reports the results before controlling for major-fixed effects, while Panel B presents the estimates obtained when including these controls.

Our results indicate that the positive return of one's status as a public servant declines across quintiles of the wage distribution among women and men both with and without major controls. The effect of one's possession of a master degree or PhD portrays an inverted U-shaped pattern where the effect is larger for the 50th percentile, which holds for the samples of women and men both with and without major controls. The effect of our ability measure increases across the distribution before including major effects for men and women, and this pattern is mitigated with the inclusion of major controls: the effect of ability depicts an inverted U-shaped pattern among men after including major-fixed effects and a flattened pattern among women.

5.2 Decomposition results

5.2.A Predicted wages and gender gaps

In Table 4, we present the predicted wages for women and men as well as the gender gap at the mean and at different percentiles of the wage distribution. These are the differences we will decompose into explained and unexplained effects.

Our results indicate that, on average, men earn 29% more than women.[10] The gap in favor of men depicts an inverted U-shaped pattern along the distribution of wages such that men earn 20% more than women at the 10th percentile, 35% more at the 50th percentile, and 29% more at

[10] The size of the estimated gender wage gap is similar to the estimates obtained from the Brazilian National Household Sample Survey.

the 90th percentile.

5.2.B Decomposition at the mean

Table 5, Panel A presents the decomposition results at the mean before including major-fixed effects, while Panel B displays the results when including the major controls. Before including the major controls, the wage gap at the mean is primarily explained by gender differences in returns (or prices) to characteristics. The unexplained effect represents 81% of the gap, while the explained effect captures the remaining 19%. The opposite is observed when we include major controls; in this case, the unexplained effect captures 35% of the gap, while the explained effect represents 65%. This change in the decomposition results highlights the importance of selecting majors. Moreover, most of the total explained effect captures the contribution that majors make; precisely, majors represent 80% of the explained effect and 50% of the total gender wage gap. When comparing different major groups, STEM most contributes toward explaining the gender gap—45% of the explained effect and 29% of the gender wage gap—followed by communication, social sciences, and humanities. In Table A1 of the Appendix, we demonstrate the contribution of each individual major within each major group. Within STEM, engineering and computer science make the largest contribution to the explained component. The unexplained effect primarily captures the returns to ability (although this variable is not individually significant) and age effects.

This result highlights the importance of major selection in explaining the gender wage gap and the relevance of STEM majors, which contribute to a large portion of the wage difference between men and women. Revisiting the evidence presented in Figure 1, men sort into STEM and social sciences majors more frequently than women, and these majors represent the highest-remunerating jobs in the labor market. On the other hand, women are concentrated in communication majors, whose corresponding jobs are afforded lower hourly wages.

5.2.C Unconditional quantile decomposition

Panels A of Table 6 and Figure 2 present the decomposition results of the gender wage gap at the 10th, 50th, and 90th percentiles before including major-fixed effects, while Panels B of the same figure and table report the results when including major controls.

Before including major controls, the gap is primarily explained by gender differences in returns to characteristics across the entire distribution. The importance of the unexplained effect declines when considering higher percentiles of the distribution, whereas the importance of the

explained component increases. When we include major controls, the explained effect captures most of the gender wage gap, and its importance portrays an inverted U-shaped pattern (Figure 2). This pattern mainly captures the behaviors of various major programs. The contribution to the gender wage gap of the explained component of Communication and Health programs is higher at the 50th percentile; on the other hand, STEM careers contribute to the gap more so at the 10th and 50th percentiles than at the top. In total, college majors contribute toward explaining 38% of the gender wage gap at the 10th percentile, 53% at the 50th percentile, and 38% at the top. STEM majors specifically explain 57%, 29%, and 16% of the gender gap at the 10th, 50th, and 90th percentiles, respectively.

What explains the gender wage gap along the wage distribution percentiles? Students' sorting into majors is an important factor for explaining this gap, particularly in the middle of the wage distribution, where it captures more than half of the gap. The contribution of returns to majors also follows the inverted U-shaped pattern, although it represents a much smaller percentage of the gender wage gap (4%, 10%, and 9% of the gender gap at the 10th, 50th, and 90th percentiles, respectively). The remaining observed characteristics (age, status as a public servant, possession of a master degree or PhD, ability measure, former study at a public secondary school, mother's education, family income) increase their contribution to the explained component of the gender wage gap as we move along percentiles of the distribution, while the contribution of their unexplained component follows a U-shaped pattern.

5.3 Sample selection

Our findings fairly suggest that the disparate returns among women and men run in favor of the latter, and much of that difference is guided by their choice of major. It is worth pointing out that all results were obtained from a sample of students conditioned on their being assigned a formal job. Since we wish to understand the wage gap across gender, it is important to investigate potential biases that may arise from the sample selection and affect the labor market.

Table 7 depicts the estimates of students' characteristics according to the probability of their being observed in RAIS.[11] The female dummy coefficient indicates that women are less likely to work a formal job than men, although no statistical difference was identified between genders among those who graduated. The interaction of a student's first round entrance exam score with

[11] We drop from the regression all traits related to the possession of a master's degree and the occupation of a public servant because that information is exclusively available for students who are formally employed.

the female dummy reveals that women with higher ability are more prone to be found in the RAIS than their male counterparts, which suggests that if participation in formal jobs did not differ by ability, then the wage gap would increase. Thus, our estimated earning gap may represent a lower bound.

6 Final Remarks

In this paper, we have analyzed how one's choice of major influences the gender wage gap in Brazil by highlighting the role that STEM majors play. We included longitudinal data of 8,773 students from a high-quality public flagship university (UFPE) who we observed thirteen years post-enrollment, during which time they were participating in the labor market as formal workers. We subsequently applied a decomposition analysis of the gender wage gap at the mean and at various points of the wage distribution.

Our results reveal that students' choice of major contributes toward explaining gender wage gaps, and STEM careers contribute to a large portion of this gender difference in the labor market. The majors that women and men select explain 50% of the gender wage gap at the mean, while STEM careers explain 30% of the difference. When looking at different percentiles of the wage distribution, we find that students' major selection is more important at the middle of the distribution than at the bottom or top. This selection explains 53% of the gender wage gap at the 50th percentile, while it explains 38% and 39% at the 10th and 90th percentiles, respectively. By specifically considering STEM majors, their corresponding careers explain 57%, 29%, and 16% of the gender wage gap at the 10th, 50th, and 90th percentiles, respectively.

Our study poses important policy implications that extend beyond the current gender wage gaps. Gender inequalities in the distribution of fields of study and STEM education may perpetuate the existing gender wage gap. STEM skills are relevant in a fast-changing world of work where technology is disrupting the labor market. Exposing girls to traditionally male subjects (e.g., STEM fields) early in their education encourages the hiring of women in male-dominated STEM professions. Furthermore, incentives and family–work balance policies to retain women in these fields may promote gender equality in the labor market.

References

Autor, D. H., Katz, L. B., & Kearney, M. S. (2005). *Rising wage inequality: The role of composition and prices*. NBER Working Paper No. 11628. Retrieved from https://www.nber.org/papers/w11628

Bertrand, M. (2011). New perspectives on gender. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 4b, pp. 1545–1592). Amsterdam, Netherlands: Elsevier, Ltd.

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

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

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, *8*, 436–455.

Boaz Anglade, P. U., & Deere, C. D. (2017). Decomposing the gender wealth gap in Ecuador. *World Development*, *96*, 19–31.

Carnevale, A., Smith, N., & Gulish, A. (2018). *Women can't win. Despite making educational gains and pursuing high-wage majors, women still earn less than men*. Center on Education and the Workforce, McCourt School of Public Policy, Georgetown University. Retrieved from https://cew.georgetown.edu/cew-reports/genderwagegap/

Chernozhukov, V., Fernández-Val, I., & Melly, B. (2012). *Inference on counterfactual distributions*. Working Paper CWP05/12. The Institute for Fiscal Studies Department of Economics, UCL. Retrieved from https://arxiv.org/abs/0904.0951

Chi, W., & Li, B. (2008). Glass ceiling or sticky floor? Examining the gender earnings differential across the earnings distribution in urban China, 1987–2004. *Journal of Comparative Economics*, *36*(2), 243–263.

DiNardo, J., Fortin, N. M., & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica*, *64*, 1001–1044.

Donald, S. G., Green, D. A., & Paarsch, H. J. (2000). Differences in wage distributions between Canada and the United States: An application of a flexible estimator of distribution functions in the presence of covariates source. *Review of Economic Studies*, *67*, 609–633.

Duryea, S., de Freitas, L. B., Garcia Ozemela, L. M. G., Sampaio, B., Sampaio, G. R., & Trevisan, G. (2019). Universities and intergenerational social mobility in Brazil: Examining patterns by race and gender. *Journal of Economics, Race, and Policy*, *2*(4), 240–256. https://doi.org/10.1007/s41996-019-00033-1

Firpo S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953–973.

Fortin, N. M., Lemieux, T., & Firpo, S. (2010). *Decomposition methods in economics*. NBER Working Paper, Series 16045. Retrieved from https://www.nber.org/papers/w16045

Goldin, C., Katz, L. F. (2008). Gender differences in careers, education, and games. Transitions: career and family life cycles of the educational elite. *American Economic Review: Papers & Proceedings*, *98*, 2.

ILOSTAT. (2018). *Employment statistics*. Retrieved from https://ilostat.ilo.org/topics/employment/

Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, *101*, 410–442.

Machado, J., & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, *20*, 445–465.

Machin S., & Puhani, P. A. (2003). Subject of degree and the gender wage differential: Evidence from the UK and Germany. *Economic Letters*, *79*(3), 393–400.

Ñopo, H. (2012). *New century, old disparities: Gender and ethnic earnings gaps in Latin America and the Caribbean*. Washington, DC: World Bank Publications.

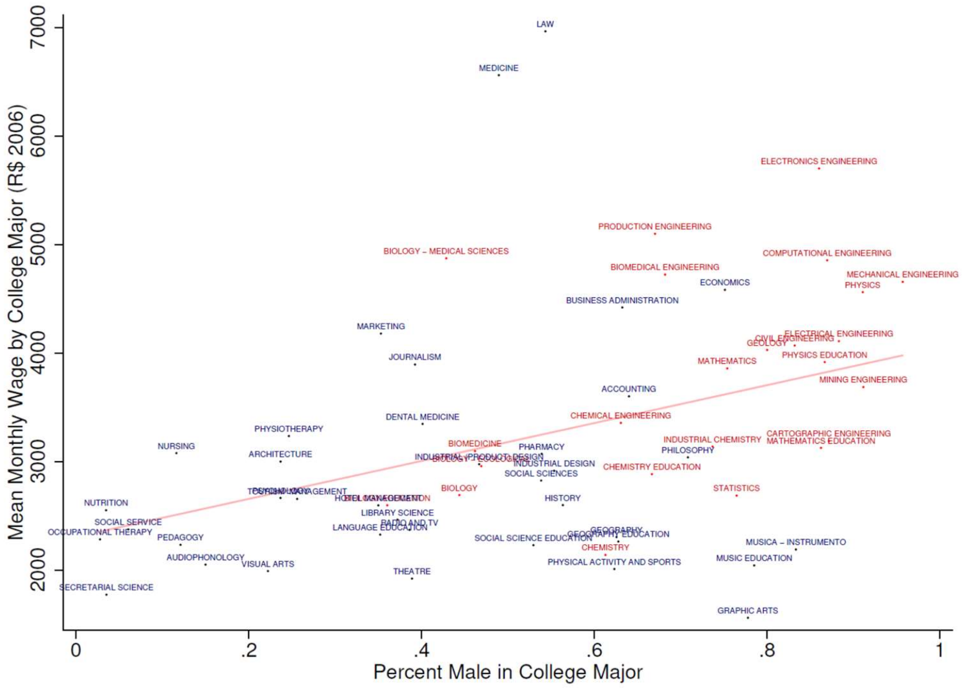
Oaxaca, R. (1973). Male–female wage differentials in urban labor markets. *International Economic Review*, *14*, 693–709.

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

UNESCO. (2017a). *Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM).* Paris, France: UNESCO.

UNESCO. (2017b). *Measuring gender equality in science and engineering: The SAGA toolkit*. SAGA Working Paper 2. Retrieved from http://repositorio.minedu.gob.pe/handle/MINEDU/5727

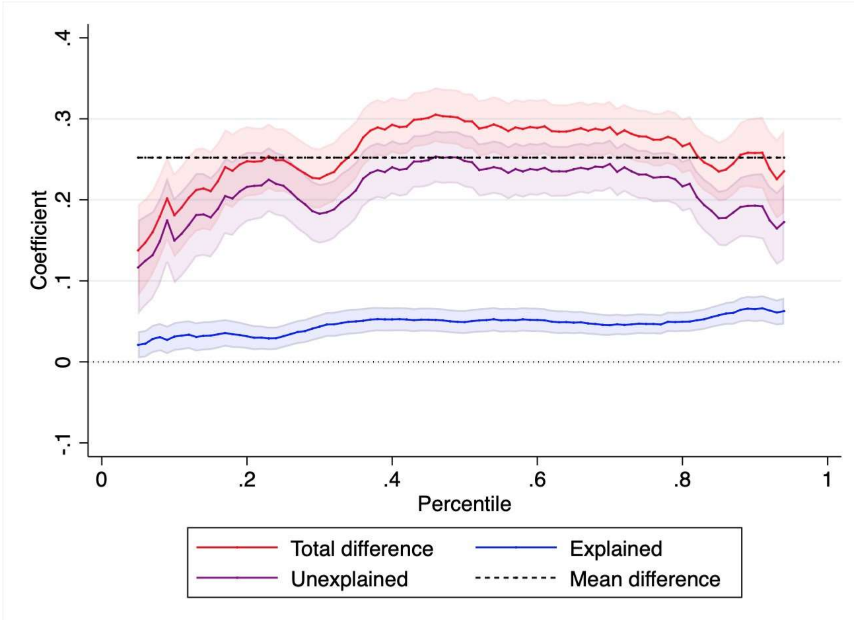
Figure 1. Correlation between average monthly wages and the proportion of males by major



Note: This figure illustrates the proportion of male students by major and the average monthly wage of those students who were working in the formal labor market thirteen years after applying to college. STEM majors are labeled in red font; the sample includes UFPE enrollees from the 2002–2004 cohort.

Figure 2. Decomposing the gender wage gap by percentile

(a) No major controls



(b) Major controls

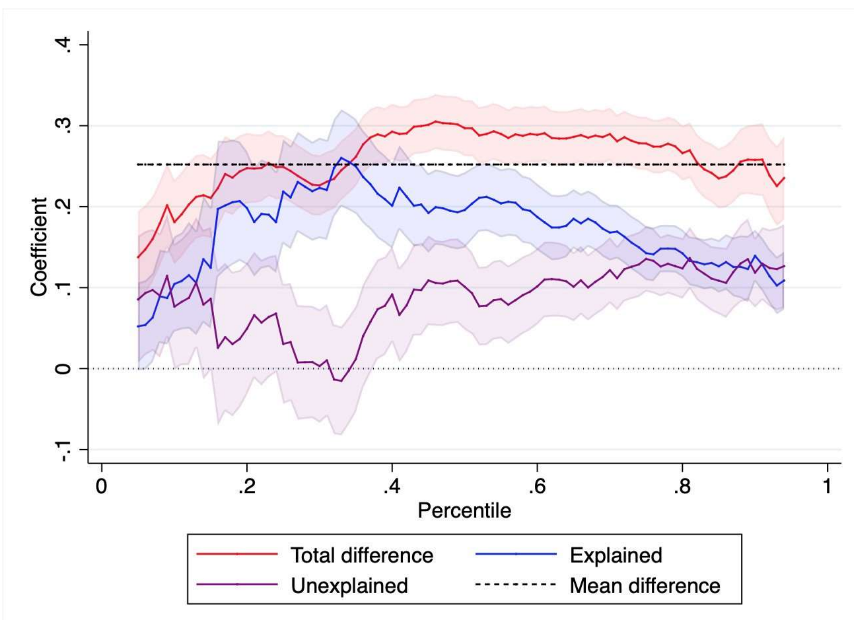


Table 1. Descriptive statistics

	Men			Women		
	Mean	SD	n	Mean	SD	n
(log) Average hourly wage	2.865	0.780	5,008	2.614	0.751	4,688
Public servant	0.546	0.498	5,008	0.569	0.495	4,688
Non-public servant	0.454	0.498	5,008	0.431	0.495	4,688
Master degree or PhD	0.089	0.285	5,008	0.103	0.305	4,688
Graduated or ungraduated	0.911	0.285	5,008	0.897	0.305	4,688
Round 1 score	6.209	1.162	4,599	6.086	1.153	4,294
Exclusive public schooling	0.277	0.448	4,547	0.255	0.436	4,273
Non-exclusive public schooling	0.723	0.448	4,547	0.745	0.436	4,273
Age brackets						
18–19 years	0.405	0.491	5,008	0.432	0.495	4,688
20–21 years	0.276	0.447	5,008	0.311	0.463	4,688
22–23 years	0.137	0.344	5,008	0.123	0.329	4,688
24–30 years	0.182	0.386	5,008	0.135	0.341	4,688
Mother's education						
Illiterate to incomplete secondary ed.	0.202	0.401	4,536	0.238	0.426	4,268
Completed secondary ed.	0.376	0.484	4,536	0.386	0.487	4,268
Completed college	0.422	0.494	4,536	0.376	0.485	4,268
Income brackets						
Family income: 0–300	0.254	0.436	4,535	0.303	0.460	4,267
Family income: 300–1,000	0.347	0.476	4,535	0.383	0.486	4,267
Family income: 1,000–1,500	0.161	0.368	4,535	0.131	0.337	4,267
Family income: 1,500–2,000	0.104	0.305	4,535	0.080	0.271	4,267
Family income: 2,000 or higher	0.134	0.341	4,535	0.103	0.304	4,267
Fields of knowledge						
Social sciences	0.164	0.370	5,008	0.096	0.294	4,688
Humanities	0.185	0.388	5,008	0.333	0.471	4,688
STEM	0.428	0.495	5,008	0.157	0.364	4,688
Physical education	0.033	0.180	5,008	0.022	0.145	4,688
Health	0.091	0.287	5,008	0.204	0.403	4,688
Communication	0.051	0.219	5,008	0.141	0.348	4,688
Music	0.020	0.139	5,008	0.006	0.074	4,688
Visual arts	0.028	0.166	5,008	0.042	0.201	4,688

Note: The sample comprises students who matriculated at UFPE, were aged eighteen to thirty years, and were working in the formal labor market thirteen years after applying; the 2002–2004 cohort is exclusively included.

Table 2. RIF regressions at the mean

	Without major control		With major control	
	Men	Women	Men	Women
	(1)	(2)	(3)	(4)
Public servant	0.139***	0.195***	0.186***	0.197***
	(0.022)	(0.021)	(0.022)	(0.021)
Master degree or PhD	0.203***	0.186***	0.300***	0.236***
	(0.038)	(0.034)	(0.037)	(0.034)
Round 1 score	0.215***	0.227***	0.171***	0.139***
	(0.010)	(0.011)	(0.015)	(0.016)
Exclusive public schooling	0.106***	0.050*	0.075***	0.029
	(0.027)	(0.027)	(0.025)	(0.026)
Aged 20–21 years	-0.119***	-0.051**	-0.063**	-0.032
	(0.026)	(0.024)	(0.025)	(0.024)
Aged 22–23 years	-0.158***	-0.058	-0.106***	-0.041
	(0.034)	(0.035)	(0.033)	(0.034)
Aged 24–30 years	-0.017	-0.083**	0.045	-0.081**
	(0.034)	(0.037)	(0.033)	(0.037)
Mother: illiterate to primary ed.	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Mother: completed secondary ed.	0.030	0.040	0.026	0.032
	(0.031)	(0.029)	(0.030)	(0.028)
Mother: completed college	0.089***	0.112***	0.078**	0.079**
	(0.034)	(0.032)	(0.032)	(0.031)
Family income: 300–1,000	0.079***	0.058**	0.070**	0.059**
	(0.029)	(0.027)	(0.028)	(0.026)
Family income: 1,000–1,500	0.164***	0.120***	0.128***	0.116***
	(0.036)	(0.038)	(0.035)	(0.037)
Family income: 1,500–2,000	0.242***	0.184***	0.186***	0.145***
	(0.042)	(0.045)	(0.041)	(0.044)
Family income: 2,000 or higher	0.324***	0.310***	0.230***	0.239***
	(0.041)	(0.043)	(0.039)	(0.042)
Constant	1.410***	1.089***	1.524**	1.050***
	(0.069)	(0.067)	(0.678)	(0.383)
Observations	4,519	4,254	4,519	4,254
R-squared	0.185	0.222	0.284	0.299

Note: All regressions include cohort-fixed effects; robust standard errors are listed in parentheses in regressions without major controls and are clustered at the major level in regressions with major controls. *** p¡0.01, ** p¡0.05, * p¡0.1; omitted categories include students aged eighteen years, at least one year in a non-public secondary school, family income of 0–300 R$, and mother's completion of primary school or lower.

Table 3. RIF regressions at the 10th, 50th, and 90th percentiles of the wage distribution

	Panel A: without major controls						**Panel B: with major controls**					
	10th		**50th**		**90th**		**10th**		**50th**		**90th**	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Public servant	0.552***	0.424***	0.050*	0.272***	0.005	0.099***	0.659***	0.434***	0.136***	0.311***	-0.008	0.100**
	(0.049)	(0.044)	(0.028)	(0.027)	(0.034)	(0.038)	(0.052)	(0.047)	(0.029)	(0.028)	(0.033)	(0.040)
Master degree or PhD	0.087	0.126**	0.283***	0.252***	-0.018	-0.018	0.227***	0.139**	0.386***	0.291***	0.096	0.072
	(0.069)	(0.060)	(0.048)	(0.043)	(0.065)	(0.066)	(0.075)	(0.061)	(0.047)	(0.044)	(0.064)	(0.065)
Round 1 score	0.167***	0.186***	0.219***	0.191***	0.222***	0.279***	0.146***	0.138***	0.180***	0.138***	0.135***	0.130***
	(0.021)	(0.020)	(0.013)	(0.013)	(0.019)	(0.024)	(0.031)	(0.032)	(0.020)	(0.022)	(0.025)	(0.032)
Exclusive public schooling	0.152***	0.046	0.152***	0.044	0.059	0.006	0.108*	0.058	0.131***	0.017	0.064	-0.030
	(0.058)	(0.056)	(0.034)	(0.034)	(0.040)	(0.043)	(0.056)	(0.056)	(0.033)	(0.034)	(0.040)	(0.042)
Aged 20–21 years	-0.115**	-0.018	-0.169***	-0.072**	-0.129***	-0.015	-0.047	-0.008	-0.103***	-0.062**	-0.103**	0.000
	(0.056)	(0.047)	(0.034)	(0.031)	(0.042)	(0.045)	(0.055)	(0.048)	(0.033)	(0.031)	(0.041)	(0.044)
Aged 22–23 years	-0.106	-0.170**	-0.222***	-0.136***	-0.073	-0.014	-0.049	-0.127	-0.159***	-0.124***	-0.064	-0.027
	(0.075)	(0.080)	(0.043)	(0.045)	(0.052)	(0.057)	(0.075)	(0.079)	(0.043)	(0.044)	(0.051)	(0.057)
Aged 24–30 years	-0.158**	-0.219**	-0.093**	-0.137***	-0.049	-0.001	-0.069	-0.176**	0.004	-0.134***	-0.031	-0.032
	(0.076)	(0.087)	(0.044)	(0.048)	(0.050)	(0.056)	(0.079)	(0.089)	(0.044)	(0.048)	(0.050)	(0.056)
Mother: completed secondary ed.	-0.038	0.023	0.086**	0.029	0.043	0.018	-0.042	0.021	0.079**	0.028	0.035	0.005
	(0.070)	(0.065)	(0.040)	(0.037)	(0.042)	(0.040)	(0.068)	(0.065)	(0.039)	(0.037)	(0.041)	(0.040)
Mother: completed college	-0.027	-0.020	0.109**	0.135***	0.104**	0.111**	-0.023	-0.038	0.098**	0.113***	0.091*	0.059
	(0.076)	(0.072)	(0.044)	(0.042)	(0.048)	(0.052)	(0.074)	(0.071)	(0.043)	(0.042)	(0.047)	(0.051)
Family income: 300–1,000	0.188***	0.064	0.079**	0.064*	-0.030	0.037	0.146**	0.039	0.077**	0.065*	-0.020	0.055
	(0.068)	(0.060)	(0.037)	(0.035)	(0.039)	(0.039)	(0.067)	(0.060)	(0.036)	(0.034)	(0.038)	(0.038)
Family income: 1,000–1,500	0.212**	0.185**	0.161***	0.118**	0.109*	0.070	0.134*	0.170**	0.127***	0.110**	0.094*	0.072
	(0.082)	(0.072)	(0.047)	(0.049)	(0.057)	(0.068)	(0.081)	(0.072)	(0.046)	(0.048)	(0.056)	(0.067)
Family income: 1,500–2,000	0.364***	0.200**	0.274***	0.191***	0.145**	0.286***	0.279***	0.174**	0.221***	0.162***	0.094	0.212**
	(0.087)	(0.079)	(0.056)	(0.057)	(0.070)	(0.096)	(0.087)	(0.079)	(0.054)	(0.055)	(0.069)	(0.092)
Family income: 2,000 or higher	0.348***	0.152*	0.372***	0.219***	0.439***	0.548***	0.240***	0.112	0.285***	0.160***	0.344***	0.472***
	(0.083)	(0.079)	(0.052)	(0.054)	(0.078)	(0.101)	(0.083)	(0.082)	(0.051)	(0.054)	(0.076)	(0.099)
Constant	5.486***	5.352***	6.502***	6.302***	7.479***	6.753***	6.362***	3.040***	5.535***	6.420***	7.709***	7.352***
	(0.157)	(0.140)	(0.089)	(0.082)	(0.114)	(0.137)	(0.218)	(1.174)	(0.129)	(0.492)	(0.157)	(0.236)
Observations	4,519	4,254	4,519	4,254	4,519	4,254	4,519	4,254	4,519	4,254	4,519	4,254
R-squared	0.061	0.063	0.130	0.143	0.085	0.112	0.120	0.107	0.211	0.207	0.145	0.180

Note: All regressions include cohort-fixed effects; robust standard errors are listed in parentheses in regressions without major controls and are clustered at the major level in regressions with major controls. *** p<0.01, ** p<0.05, * p<0.1; omitted categories include students aged eighteen years, at least one year in a non-public secondary school, family income of 0–300 R$, and mother's completion of primary school or lower.

Table 4. Gender wage gaps at the mean and the 10th, 50th, and 90th percentiles of the wage distribution

	Mean	10th	50th	90th
Men	2.864***	1.796***	2.907***	3.806***
	(0.006)	(0.024)	(0.014)	(0.018)
Women	2.612***	1.615***	2.610***	3.548***
	(0.006)	(0.021)	(0.015)	(0.019)
Difference	0.252***	0.181***	0.297***	0.258***
	(0.009)	(0.032)	(0.021)	(0.026)

Note: *** , **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5. Decomposition effects at the mean

	Panel A: without major controls		Panel B: with major controls	
	Explained	Unexplained	Explained	Unexplained
Total	**0.047******	**0.205******	**0.164******	**0.088******
	(0.008)	**(0.015)**	**(0.010)**	**(0.005)**
	19%	**81%**	**65%**	**35%**
Public servant	-0.003**	-0.031*	-0.004*	-0.007
	(0.002)	(0.017)	(0.002)	(0.032)
Master degree or PhD	-0.003**	0.002	-0.004**	0.006
	(0.001)	(0.005)	(0.002)	(0.006)
Round 1 score	0.027***	-0.073	0.022***	0.193
	(0.005)	(0.097)	(0.005)	(0.146)
Exclusive public schooling	0.002*	0.014	0.002*	0.012
	(0.001)	(0.009)	(0.001)	(0.009)
Aged 20–21 years	0.004***	-0.022*	0.002**	-0.010
	(0.001)	(0.012)	(0.001)	(0.010)
Aged 22–23 years	-0.003**	-0.012**	-0.002**	-0.008
	(0.001)	(0.006)	(0.001)	(0.005)
Aged 24–30 years	-0.001	0.007	0.002	0.014***
	(0.002)	(0.005)	(0.001)	(0.004)
Mother: completed secondary ed.	-0.000	-0.004	-0.000	-0.002
	(0.000)	(0.016)	(0.000)	(0.016)
Mother: completed college	0.004**	-0.009	0.004**	-0.000
	(0.002)	(0.018)	(0.002)	(0.019)
Family income: 300–1,000	-0.003**	0.008	-0.002*	0.004
	(0.001)	(0.015)	(0.001)	(0.016)
Family income: 1,000–1,500	0.005***	0.006	0.004***	0.002
	(0.002)	(0.007)	(0.001)	(0.007)
Family income: 1,500–2,000	0.006***	0.005	0.005***	0.003
	(0.002)	(0.005)	(0.002)	(0.005)
Family income: 2,000 or higher	0.010***	0.001	0.007***	-0.001
	(0.003)	(0.006)	(0.002)	(0.007)
		Program contributions		
Social sciences			0.016***	-0.003***
			(0.002)	(0.001)
Humanities			0.013***	0.006
			(0.004)	(0.005)
STEM			0.074***	0.003***
			(0.003)	(0.001)
Health			0.011***	0.001
			(0.003)	(0.005)
Physical education			-0.004***	0.001**
			(0.001)	(0.000)
Communication			0.017***	-0.002
			(0.002)	(0.002)
Music			-0.004***	-0.001**
			(0.001)	(0.000)
Visual arts			0.006***	-0.010***
			(0.001)	(0.001)
Constant		0.321***		-0.090
		(0.101)		(0.152)
Observations	8,773	8,773	8,773	8,773

Note: All regressions include cohort-fixed effects; robust standard errors are listed in parentheses in regressions without major controls and are clustered at the major level in regressions with major controls. *** p<0.01, ** p<0.05, * p<0.1; omitted categories include students aged eighteen years, at least one year in a non-public secondary school, family income of 0–300 R$, and mother's completion of primary school or lower.

Table 6. Decomposition effects at the 10th, 50th, and 90th percentiles

	Panel A: without major controls						Panel B: with major controls					
	10th		50th		90th		10th		50th		90th	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
Total	0.031***	0.150***	0.049***	0.248***	0.065***	0.193***	0.104***	0.076	0.196***	0.101***	0.139***	0.119***
	(0.011)	(0.032)	(0.009)	(0.020)	(0.010)	(0.024)	(0.036)	(0.047)	(0.028)	(0.032)	(0.020)	(0.026)
	17%	83%	16%	84%	25%	75%	58%	42%	66%	34%	54%	46%
Public servant	-0.011**	0.038	-0.002*	-0.037*	0.002	-0.049*	-0.013**	0.099**	-0.003*	-0.030	0.002*	-0.055*
	(0.005)	(0.037)	(0.001)	(0.022)	(0.001)	(0.028)	(0.006)	(0.039)	(0.002)	(0.023)	(0.001)	(0.029)
Master degree or PhD	-0.004**	-0.000	-0.004**	-0.005	-0.000	0.004	-0.006**	0.013	-0.005**	-0.001	-0.002	0.004
	(0.002)	(0.008)	(0.002)	(0.007)	(0.001)	(0.009)	(0.003)	(0.008)	(0.002)	(0.007)	(0.001)	(0.009)
Round 1 score	0.023***	-0.081	0.030***	0.066	0.031***	-0.284	0.022***	0.198	0.025***	0.317*	0.020***	0.144
	(0.005)	(0.172)	(0.006)	(0.116)	(0.007)	(0.191)	(0.006)	(0.262)	(0.005)	(0.188)	(0.005)	(0.261)
Exclusive public schooling	0.002	0.007	0.002*	0.017	0.002	0.028*	0.001	-0.011	0.002	0.017	0.002	0.034**
	(0.002)	(0.021)	(0.001)	(0.013)	(0.001)	(0.014)	(0.001)	(0.020)	(0.001)	(0.012)	(0.001)	(0.014)
Aged 20–21 years	0.004*	-0.044*	0.005***	-0.015	0.002	-0.008	0.001	-0.025	0.003**	-0.003	0.001	0.000
	(0.002)	(0.023)	(0.002)	(0.015)	(0.002)	(0.019)	(0.002)	(0.023)	(0.001)	(0.015)	(0.001)	(0.019)
Aged 22–23 years	-0.003	-0.008	-0.005***	-0.016**	-0.001	-0.016*	-0.001	-0.006	-0.004**	-0.012	-0.001	-0.012
	(0.002)	(0.012)	(0.002)	(0.007)	(0.001)	(0.009)	(0.002)	(0.012)	(0.002)	(0.007)	(0.001)	(0.009)
Aged 24–30 years	-0.002	0.012	-0.003	0.004	0.002	-0.004	0.002	0.019	0.001	0.010	0.003	0.003
	(0.003)	(0.012)	(0.002)	(0.007)	(0.002)	(0.008)	(0.003)	(0.012)	(0.002)	(0.007)	(0.002)	(0.008)
Mother: completed secondary ed.	0.000	-0.006	-0.000	-0.010	-0.000	0.002	0.000	0.003	-0.000	-0.006	-0.000	0.006
	(0.001)	(0.037)	(0.001)	(0.021)	(0.001)	(0.022)	(0.001)	(0.036)	(0.001)	(0.021)	(0.001)	(0.021)
Mother: completed college	0.001	0.006	0.005**	-0.029	0.008***	0.006	0.001	0.018	0.004*	-0.021	0.008***	0.021
	(0.004)	(0.039)	(0.002)	(0.023)	(0.003)	(0.026)	(0.003)	(0.038)	(0.002)	(0.023)	(0.003)	(0.026)
Family income: 300–1,000	-0.006**	0.042	-0.002	-0.020	0.000	-0.018	-0.005*	0.040	-0.002	-0.022	0.000	-0.022
	(0.003)	(0.035)	(0.001)	(0.020)	(0.001)	(0.021)	(0.003)	(0.034)	(0.001)	(0.019)	(0.001)	(0.020)
Family income: 1,000–1,500	0.006**	0.009	0.005**	-0.009	0.003*	0.019*	0.004	0.001	0.004**	-0.012	0.003	0.016
	(0.003)	(0.014)	(0.002)	(0.009)	(0.002)	(0.011)	(0.003)	(0.014)	(0.002)	(0.009)	(0.002)	(0.011)
Family income: 1,500–2,000	0.009***	0.016*	0.006***	-0.002	0.004*	0.001	0.006**	0.010	0.005***	-0.003	0.003	0.004
	(0.003)	(0.009)	(0.002)	(0.006)	(0.002)	(0.009)	(0.003)	(0.009)	(0.002)	(0.006)	(0.002)	(0.009)
Family income: 2,000 or higher	0.011***	0.014	0.010***	-0.002	0.010***	-0.016	0.007**	0.008	0.007***	-0.002	0.007**	-0.017
	(0.003)	(0.011)	(0.003)	(0.008)	(0.003)	(0.013)	(0.003)	(0.011)	(0.002)	(0.008)	(0.003)	(0.013)
Social sciences							0.023***	-0.008	0.017***	-0.003	0.008**	-0.013
							(0.005)	(0.007)	(0.003)	(0.005)	(0.003)	(0.008)
Humanities							0.012	-0.016	0004	0.040*	0.023***	0.022
							(0.026)	(0.034)	(0.018)	(0.022)	(0.008)	(0.017)
STEM							0.102***	-0.009	0.085***	0.008	0.042***	0.006
							(0.013)	(0.012)	(0.008)	(0.007)	(0.010)	(0.007)
Health							-0.024	-0.003	0.029**	0.000	0.002	0.013
							(0.013)	(0.020)	(0.014)	(0.017)	(0.011)	(0.019)
Physical education							-0.005**	0.002	-0.005***	-0.001	-0.001	0.004**
							(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.002)
Communication							-0.022**	0.045***	0.027***	-0.016	0.010***	-0.004
							(0.009)	(0.017)	(0.010)	(0.012)	(0.004)	(0.009)
Music							-0.023***	0.013**	-0.003	0.008	0.008***	0.004
							(0.006)	(0.006)	(0.012)	(0.013)	(0.002)	(0.003)
Visual arts							0.005	-0.016	0.003	-0.007	0.008***	-0.010**
							(0.007)	(0.010)	(0.003)	(0.005)	(0.002)	(0.004)
Constant		0.160		0.332***		0.516***		-0.260		-0.109		-0.011
		(0.206)		(0.123)		(0.179)		(0.285)		(0.190)		(0.252)
Observations	8,773	8,773	8,773	8,773	8,773	8,773	8,773	8,773	8,773	8,773	8,773	8,773

Note: All regressions include cohort-fixed effects; robust standard errors are listed in parentheses in regressions without major controls and are clustered at the major level in regressions with major controls. *** p<0.01, ** p<0.05, * p<0.1; omitted categories include students aged eighteen years, at least one year in a non-public secondary school, family income of 0–300 R$, and mother's completion of primary school or less.

Table 7. Correlation between students' characteristics and the probability of being employed

VARIABLES	in RAIS
Aged 20–21 years	-0.009
	(0.009)
Aged 22–23 years	-0.011
	(0.010)
Aged 24–30 years	0.004
	(0.014)
Mother: completed secondary ed.	-0.008
	(0.009)
Mother: completed college	-0.007
	(0.013)
Family income: 300–1,000	-0.007
	(0.009)
Family income: 1,000–1,500	-0.028*
	(0.014)
Family income: 1,500–2,000	-0.045***
	(0.017)
Family income: 2,000 or higher	-0.087***
	(0.014)
Exclusive public schooling	0.046***
	(0.008)
Round 1 score	0.018***
	(0.007)
Female * round 1 score	0.012*
	(0.006)
Female	-0.101**
	(0.039)
Graduated	0.022
	(0.014)
Female * graduated	0.004
	(0.019)
Constant	0.765***
	(0.044)
Observations	10,567

Note: The regression includes major-fixed effects. *** , **, and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix

Table A1. Decomposition effects: each major's contribution (on regressions of Table 5)

	Mean		10th		50th		90th	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained **Social**
				Sciences				
BUSINESS ADMINISTRATION	0.006***	0.001***	0.008***	-0.000	0.007***	0.001	0.001	-0.006
	(0.001)	(0.000)	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)	(0.005)
ACCOUNTING	0.005***	0.000	0.008***	0.000	0.006***	-0.001	0.004**	-0.001
	(0.001)	(0.000)	(0.002)	(0.005)	(0.002)	(0.004)	(0.002)	(0.005)
ECONOMICS	0.004***	-0.002***	0.005**	-0.001	0.004**	-0.002	0.002	-0.006*
	(0.001)	(0.000)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
HOTEL MANAGEMENT	0.001*	-0.002***	0.002	-0.007*	0.001	-0.001	0.001*	-0.000
	(0.001)	(0.000)	(0.002)	(0.004)	(0.001)	(0.002)	(0.001)	(0.001)
				Humanities				
SOCIAL SCIENCES	0.000	-0.001***	0.000	0.001	0.000	0.003	0.000	-0.002
	(0.000)	(0.000)	(0.002)	(0.005)	(0.000)	(0.002)	(0.000)	(0.002)
SOCIAL SCIENCES EDUCATION	-0.000	0.000***	-0.000	0.002	-0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.001)	(0.003)	(0.000)	(0.001)	(0.000)	(0.001)
PHILOSOPHY	0.000**	0.001***	-0.001	0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
LAW	0.000	-0.000	-0.000	-0.003	0.000	0.005	0.001	0.010
	(0.002)	(0.002)	(0.000)	(0.005)	(0.002)	(0.004)	(0.005)	(0.010)
GEOGRAPHY	-0.002***	0.001**	-0.003	0.004	-0.002*	-0.000	-0.001	-0.001
	(0.001)	(0.000)	(0.002)	(0.005)	(0.001)	(0.002)	(0.001)	(0.002)
GEOGRAPHY EDUCATION	-0.001***	0.003***	-0.000	0.006	-0.002*	0.003	-0.000	0.002
	(0.000)	(0.001)	(0.001)	(0.004)	(0.001)	(0.002)	(0.000)	(0.001)
HISTORY	-0.001*	0.000	-0.002	-0.001	-0.001	0.000	-0.001	-0.000
	(0.001)	(0.000)	(0.001)	(0.005)	(0.001)	(0.003)	(0.001)	(0.003)
SOCIAL SERVICES	-0.009***	0.015***	-0.031***	0.031***	-0.023	0.031*	0.007***	0.009**
	(0.001)	(0.002)	(0.005)	(0.008)	(0.015)	(0.016)	(0.003)	(0.004)
TOURISM MANAGEMENT	0.005***	-0.004***	0.007	-0.007	0.004*	-0.004	0.003	-0.000
	(0.001)	(0.000)	(0.005)	(0.008)	(0.002)	(0.004)	(0.002)	(0.003)
PEDAGOGY	0.015***	-0.006	0.035	-0.042	0.017*	0.009	0.007*	0.006
	(0.002)	(0.004)	(0.024)	(0.029)	(0.010)	(0.013)	(0.004)	(0.007)
PSYCHOLOGY	0.006***	-0.003***	0.007	-0.007	0.009***	-0.006	0.007***	-0.001
	(0.001)	(0.000)	(0.007)	(0.010)	(0.003)	(0.004)	(0.001)	(0.003)
				STEM				
BIOMEDICAL ENGINEERING	0.001**	0.000	0.001*	0.000	0.000	-0.001	0.001	0.002
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
CARTOGRAPHIC ENGINEERING	0.002***	0.001***	0.005**	0.001	0.002*	0.000	0.002	0.001**
	(0.000)	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.000)
CIVIL ENGINEERING	0.007***	-0.001***	0.011***	-0.001	0.013***	-0.002	-0.004**	-0.002
	(0.001)	(0.000)	(0.004)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)
COMPUTATIONAL ENGINEERING	0.002***	-0.001***	0.006***	0.001	0.003	-0.001	0.000	-0.002
	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.002)	(0.002)
MINING ENGINEERING	0.005***	0.000**	0.002	-0.001	0.008***	0.001	0.003*	0.001**
	(0.001)	(0.000)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)	(0.000)
PRODUCTION ENGINEERING	0.002***	-0.001***	0.001	-0.001	0.003**	-0.000	0.000	-0.003
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
ELECTRICAL ENGINEERING	0.014***	0.001***	0.023***	0.003*	0.013***	-0.001	0.003	0.001
	(0.001)	(0.000)	(0.004)	(0.002)	(0.003)	(0.001)	(0.003)	(0.001)
ELETRONIC ENGINEERING	0.011***	0.001***	0.010***	0.001	0.010***	0.001	0.015***	0.000
	(0.001)	(0.000)	(0.003)	(0.001)	(0.002)	(0.001)	(0.004)	(0.002)
MECHANICAL ENGINEERING	0.013***	0.001***	0.018***	0.001	0.013***	0.001	0.008**	0.000
	(0.001)	(0.000)	(0.004)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)
CHEMICAL ENGINEERING	0.001***	0.001***	0.001	-0.002	0.001	0.001	0.000	0.003
	(0.000)	(0.000)	(0.001)	(0.004)	(0.001)	(0.002)	(0.001)	(0.003)
STATISTICS	-0.000***	-0.000	0.000	0.001	0.000	0.000	-0.001***	0.000
	(0.000)	(0.000)	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)
PHYSICS	0.000	-0.000*	-0.000	-0.000	-0.000	-0.000	0.005**	0.000
	(0.000)	(0.000)	(0.002)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)
PHYSICS EDUCATION	0.002***	0.001***	0.003**	0.000	0.002	0.001	0.001	0.001**
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
GEOLOGY	0.003***	0.001***	0.004**	0.003	0.003**	0.001	0.004**	0.000
	(0.001)	(0.000)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
MATHEMATICS	0.001***	-0.001***	-0.001	-0.001	0.001	-0.001	-0.000	-0.002
	(0.000)	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
MATHEMATICS EDUCATION	0.000*	-0.001***	-0.002	-0.001*	0.001	-0.001	-0.000	0.000
	(0.000)	(0.000)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
CHEMISTRY	-0.000	-0.000**	-0.001	0.001	-0.000	-0.001	0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.001)	(0.000)	(0.001)
CHEMISTRY EDUCATION	0.001***	0.001***	-0.000	0.000	0.001	0.001	0.000	0.001*
	(0.000)	(0.000)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)

Continued on next page

Table A1 – continued								
	Mean		10th		50th		90th	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
INDUSTRIAL CHEMISTRY	0.001***	0.000	0.005**	0.000	0.000	0.001	0.000	0.001
	(0.000)	(0.000)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
BIOMEDICINE	0.001	-0.003***	0.002	-0.010*	0.001	-0.000	0.001	0.001
	(0.001)	(0.001)	(0.002)	(0.005)	(0.001)	(0.003)	(0.001)	(0.003)
COMPUTATIONAL SCIENCE	0.006***	-0.001**	0.007***	0.000	0.009***	-0.000	0.002	-0.000
	(0.001)	(0.000)	(0.003)	(0.000)	(0.003)	(0.001)	(0.004)	(0.002)
BIOLOGY	0.001*	0.003***	0.002	0.001	0.001	0.003	0.001	-0.003
	(0.001)	(0.000)	(0.001)	(0.006)	(0.001)	(0.003)	(0.001)	(0.003)
BIOLOGY EDUCATION	0.001***	0.002***	0.005	-0.003	0.001	0.003	-0.000	0.004
	(0.000)	(0.000)	(0.003)	(0.007)	(0.001)	(0.003)	(0.001)	(0.003)
BIOLOGY – MEDICAL SICENCE	-0.000	-0.000**	-0.000	-0.000	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
BIOLOGY – ECOLOGICAL	0.000	0.001***	0.001	-0.003	-0.000	0.003	0.001	0.001*
	(0.000)	(0.000)	(0.001)	(0.003)	(0.000)	(0.002)	(0.001)	(0.001)
Physical Education								
PHYSICAL ACT. AND SPORTS	-0.004***	0.001**	-0.005**	0.002	-0.005***	-0.001	-0.001	0.004**
	(0.001)	(0.000)	(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.002)
Health								
NURSING	0.007***	-0.004***	0.003	-0.010	0.013*	-0.012	-0.009	0.022*
	(0.001)	(0.001)	(0.012)	(0.013)	(0.007)	(0.008)	(0.010)	(0.012)
PHARMACY	-0.000	-0.003***	-0.000	-0.016**	-0.000	0.001	-0.000	0.001
	(0.002)	(0.000)	(0.003)	(0.006)	(0.001)	(0.003)	(0.001)	(0.003)
PHYSIOTERAPHY	0.001***	0.004***	-0.002	0.001	0.003	0.005	0.002	0.002
	(0.000)	(0.001)	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.006)
AUDIOPHONOLOGY	0.002***	0.001***	-0.003**	0.004*	0.004*	0.002	0.002***	0.001*
	(0.001)	(0.000)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
NUTRITION	0.002***	0.005***	-0.016***	0.016***	0.000	0.008	0.008***	0.003
	(0.001)	(0.001)	(0.004)	(0.005)	(0.012)	(0.013)	(0.002)	(0.003)
DENTAL MEDICINE	0.001**	-0.000	0.002	-0.008*	0.001	-0.001	-0.001	0.008
	(0.000)	(0.001)	(0.002)	(0.005)	(0.001)	(0.004)	(0.001)	(0.005)
MEDICINE	-0.001	-0.006**	-0.000	0.002	-0.002	0.001	-0.002	-0.026***
	(0.001)	(0.002)	(0.000)	(0.004)	(0.001)	(0.004)	(0.002)	(0.010)
OCCUPATIONAL THERAPY	-0.000	0.005***	-0.007***	0.009***	0.011***	-0.004**	0.002**	0.003**
	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
Communication								
LIBRARY SCIENCE	-0.001**	0.003***	-0.002	0.009**	0.000	0.000	-0.000	0.002
	(0.000)	(0.001)	(0.002)	(0.005)	(0.001)	(0.003)	(0.001)	(0.002)
VISUAL ARTS	-0.001**	-0.000**	-0.003*	-0.002	-0.001	0.000	-0.000*	0.000
	(0.001)	(0.000)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
JOURNALISM	-0.000	0.001***	0.000	0.001	-0.000	0.003	0.001	-0.004
	(0.000)	(0.000)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.004)
LANGAUGE EDUCATION	0.007***	-0.008***	0.014***	-0.028***	0.006***	-0.005	0.005***	-0.005
	(0.001)	(0.001)	(0.005)	(0.010)	(0.002)	(0.004)	(0.001)	(0.003)
MARKETING	-0.000*	-0.002***	-0.002**	0.002	0.000	-0.005**	-0.000	-0.002
	(0.000)	(0.000)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)	(0.004)
RADIO AND TV	0.002**	-0.003***	0.005	-0.012**	0.002*	-0.002	-0.001	0.003
	(0.001)	(0.001)	(0.003)	(0.005)	(0.001)	(0.002)	(0.001)	(0.003)
SECRETARIAL SCIENCE	0.007***	0.008***	-0.035***	0.071***	0.017*	-0.006	0.003	0.007*
	(0.001)	(0.002)	(0.006)	(0.010)	(0.010)	(0.010)	(0.003)	(0.004)
THEATRE	0.001*	-0.002***	0.002	-0.004	0.001	0.001	0.001*	-0.002
	(0.001)	(0.000)	(0.002)	(0.004)	(0.001)	(0.002)	(0.000)	(0.002)
Music								
MUSIC	0.000	0.000*	0.000	0.001	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
MUSIC EDUCATION	-0.004***	-0.001***	-0.007**	-0.004**	-0.003**	-0.001	-0.000	0.001*
	(0.001)	(0.000)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Visual Arts								
ARCHITECTURE	0.006***	-0.008***	0.007	-0.012	0.004	-0.007	0.008***	-0.006*
	(0.001)	(0.001)	(0.007)	(0.009)	(0.003)	(0.005)	(0.001)	(0.003)
INDUSTRIAL DESIGN	-0.000	0.001***	-0.000	0.001	-0.000	0.002	-0.000	-0.002
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.001)	(0.000)	(0.002)
INDUSTRIAL (PRODUCT) DESIGN	0.001	-0.002***	0.001	-0.003	0.001	-0.003**	0.001	-0.003
	(0.001)	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)
DESIGN	-0.000	0.000	0.000	0.000	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GRAPHIC ARTS	0.001***	0.001***	-0.004***	0.008***	0.001	-0.001	0.002***	-0.001
	(0.000)	(0.000)	(0.001)	(0.003)	(0.001)	(0.002)	(0.000)	(0.001)