CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Wage inequality among young college graduates: Can we find any evidence for reverse gender wage differential?

Bachelor's thesis

Author: Bc. Samuel Vrbovský

Study program: Economics and Finance

Supervisor: PhDr. Martina Mysíková, Ph.D.

Year of defense: 2020

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, May 7, 2020

Samuel Vrbovsky

Abstract

This thesis examines whether in the United States among young college graduates in male-dominated dominated job fields the the gender wage differential is reversed, i.e, women earn more than similar men. The thesis further adds two additional hypotheses that narrow the examination down to large employers and singles. To evaluate those hypotheses the thesis estimates linear regression models for each of the male-dominated job field and each hypothesis using data from 2017 National Survey of College Graduates (NSCG).

Although the results revealed that in IT and mathematics and in physical sciences women earned more than similar men, with the effects being more profound among those working for large employers and among singles, the results were not statistically significant. Those results are, however, still important in context of societal narrative and gender wage gap literature, since they do not hint any potential discrimination of women in male-dominated fields.

JEL Classification	J31, J38, J70
Keywords	United States, gender wage gap, gender wage differential, gender inequality, discrimination, college, men, women

Abstrakt

Tato práce zkoumá, zdali ve Spojených Státech mezi mladými, vysokoškolsky vzdělanými absolventy pracujícími ve většinově mužských oborech, ženy vydělávají více než srovnatelní muži. Dále tato práce zkoumá dvě další hypotézy, které zužují zkoumanou skupinu na pracující u velkých zaměstnavatelů a na svobodné. K otestování těchto hypotéz jsou využity lineární regrese pro každou z hypotéz a každý většinově mužský obor. Tato práce využívá data z National Study of College Graduatues z roku 2017.

Přestože výsledky ukázaly že v IT and matematických oborech a ve fyzikálních vědách ženy vydělávají vice než srovnatelní muži, s tím, že tyto rozdíly byly větší mezi pracujícími pro velké zaměstnavatele a mezi svobodnými, tak tyto rozdíly nebyly statisticky signifikantní. Výsledky této práce jsou nicméně důležité v kontextu literatury zabývající se mzdovými rozdíly v odměňování mužů a žen a pro společenský diskurz, jelikož neprokazují diskriminaci žen ve většinově mužských oborech.

Klasifikace JEL	J31, J38, J70
Klíčová slova	Spojené Státy, rozdíl v odměňování žen a mužů, genderová nerovnost, diskriminace, vysoká škola, muži, ženy

Acknowledgments

I would like to express my gratitude to PhDr. Martina Mysíková, Ph.D.. Her experience and guidance have helped me immensely in writing my thesis. I would also like to thank my family and friends for their emotional support, and I am further grateful to Marek Hlaváč for the amazing R packages *stargazer* and *oaxaca*.

Bibliographic Record
Vrbovsky, Samuel: *Wage inequality among young college graduates: Can we find any evidence for reverse gender wage differential?*. Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2020, pages 66. Advisor: PhDr. Martina Mysíková, Ph.D.

Bachelor's Thesis Proposal

Author	Bc. Samuel Vrbovský
Supervisor	PhDr. Martina Mysíková, Ph.D.
Proposed topic	Wage inequality among young college graduates: Can we find any evidence for reverse gender wage differential?

Research question and motivation

Since the 1960s, gender wage inequality has received extensive attention from researches across the world. The results of their work have overwhelmingly shown unexplained gender wage differential with the mean around 20%, as reported in meta-study done by Weichselbaumer and Winter-Ebmer (2005). The consistency of the results had rightfully left the imprint on the narrative in the society, and nowadays, it is taken as a fact that women earn less than men even if we adjust the gap for the important characteristics like work experience, hours worked, education, race, etc.

However, few studies have recently come out claiming that in some specific cases, the preferences for males and females reverse as the hiring experiment for STEM (Science, Technology, Engineering, and Mathematics) tenure positions done by Williams and Ceci (2015) has shown. Other studies went even further and showed that when women are in high demand, a firm might be willing to pay a premium to retain them (see, for instance, Leslie et al., 2017, or Hill et al., 2015).

In my thesis, I aim to find out how pervasive the phenomenon of reverse wage differential truly is, that is whether we can find it in less prestigious occupation (compared to CEOs) where women might be rare and thus possibly in demand such as engineering or computer science. I am going to focus on young college graduates as that is the group where employers might be willing to pay more to secure talented women in fields where there are not enough women due to their insufficient supply (commonly referred as a pipeline problem).

Contribution

The topic of reverse gender wage differential is relatively new and uncharted. To my best knowledge, it has only been found using specialized samples of people from the elite occupations. My contribution to the existing literature would be to test whether it is more common and whether we can find any evidence in lower occupational positions as well.

I believe that understanding the reverse gender wage gap phenomenon is crucial for the future employment policies as well as the society's narrative since it is likely that if the pipeline problem in some fields, i.e., STEM, won't change, this phenomenon will become more common due to the ever-increasing pressure on organizations to reduce gender inequality.

Methodology

I am going to use the U.S. data from 2017 National Survey of College Graduates. I will estimate the gender wage gap (GWG) using Oaxaca-Blinder decomposition. In doing so, I will run separate OLS wage regressions for men and women, controlling for race, education, region, industry, job characteristics, employer characteristics, experience, attitudes, and demographics. The Oaxaca-Blinder decomposition allows to separate the part of the GWG explained by measurable and observed characteristics, leaving us with the unexplained part, also called the adjusted GWG.

Outline

1. Introduction
2. Literature review
3. Methodology – overview of used methods
4. Data – information about the data used and selected variables
5. Empirical model
6. Conclusion

Refernces

BLACK, Dan A., et al. Gender wage disparities among the highly educated. Journal of human resources, 2008, 43.3: 630-659.

LESLIE, Lisa M.; MANCHESTER, Colleen Flaherty; DAHM, Patricia C. Why and when does the gender gap reverse? Diversity goals and the pay premium for high potential women. Academy of Management Journal, 2017, 60.2: 402-432.

WILLIAMS, Wendy M.; CECI, Stephen J. National hiring experiments reveal 2: 1 faculty preference for women on STEM tenure track. Proceedings of the National Academy of Sciences, 2015, 112.17: 5360-5365.

BLAU, Francine D.; KAHN, Lawrence M. The gender wage gap: Extent, trends, and explanations. Journal of Economic Literature, 2017, 55.3: 789-865.

XU, Yonghong. Focusing on women in STEM: A longitudinal examination of gender-based earning gap of college graduates. The Journal of Higher Education, 2015, 86.4: 489-523.

WEICHSELBAUMER, Doris; WINTER-EBMER, Rudolf. A meta-analysis of the international gender wage gap. Journal of Economic Surveys, 2005, 19.3: 479-511.

HILL, A. D., UPADHYAY, A. D., BEEKUN, R. I. 2015. Do female and ethnically diverse executives endure inequity in the CEO position or do they benefit from their minority status? An empirical examination. Strategic Management Journal, 36: 1115–1134.

Contents

List of Tables

List of Figures

Acronyms

BP Breusch-Pagan

CEO chief executive officer

GWG gender wage gap

HD highest degree

HH household

IT information technology

LR likelihood ratio

MBA master of business administration

NLSY National Longitudinal Survey of Youth

NSCG National Survey of College Graduates

OB Oaxaca-Blinder decomposition

OLS ordinary least squares

PSID Panel Study of Income Dynamics

SESTAT Scientists and Engineers Statistical Data System

STEM science, technology, engineering, and mathematics

US United States

Chapter 1

Introduction

Throughout the years, social sciences presented substantial evidence that gender differences in personal characteristics and job-related characteristics did not account fully for the gender wage gap (GWG), i.e., difference in average wages of men and women, and that the gender wage differential was positive[1] (Weichselbaumer & Winter-Ebmer 2005). For instance, according to Blau & Kahn (2017) gender wage differential in the United States (US), in 2010, was 8.4%. We define gender wage differential as the part of GWG that remains after adjusting for relevant differences between the populations of men and women.

Since the gender wage differential is essentially just an unexplained part of the GWG, it cannot be taken by itself as a proof of existence of discrimination. That being said, we find it reasonable to believe that gender wage differential is primarily a result of the uneven treatment of men and women in the workplace. Common examples off this uneven treatment include a lower rate of promotions for women (known as glass ceiling) or motherhood penalty (Maume Jr 1999; Budig & England 2001).

Naturally, as the GWGs are different for various groups of workers, so are the unexplained parts (gender wage differentials). In some groups, the gender wage differentials were shown to be negligible. Morgan (2008) found the gender wage differentials to be almost non-existent among young college graduates in computer science or engineering. Recently, Hill *et al.* (2015) shown using an extensive sample with detailed measures of present and past performance, that female chief executive officers (CEOs) receive a 6% premium all else being equal. They argued that it might result from mandatory disclosure of CEOs wages or companies willing to pay for the diversity of the CEO. The latter argument was

[1] Throughout the thesis, we refer to GWG and gender wage differential as "positive" if women receive lower wages to be consistent with previous literature.

expanded by Leslie *et al.* (2017). They obtained detailed data about wages, performance, and perceived potential of employees in one of the Fortune 500 companies. They found that women were paid significantly more than similar men, but only when the potential was high. They argued that it might be caused by the adoption of diversity goals as companies might be willing to pay a premium to high potential women to achieve their diversity goals in higher positions where women are underrepresented.

According to a survey of insiders conducted by Bartels *et al.* (2013), around 70% of organizations in the US were committed to diversity. Out of the organizations committed to diversity, almost half utilized questionable practices, i.e., diversity-related goals in performance management, incentives for diversity staffing goals, or targeted retention allowances. One of the commonly known examples includes Intel, where annual performance bonuses were linked to fulfilling diversity goals (Intel 2018). It is likely that those practices unfairly benefit minorities and, in turn, harm the majority. In fact, some anecdotal evidence, i.e., the infamous memo by former Google employee James Damore (2017), suggests that even more egregious practices might take place, e.g., restricting training and programs to minorities.

Based on the ubiquity of diversity goals in the US companies nowadays we believe that reverse gender wage differentials might be found in lower positions (not just among CEOs) where there is a shortage of women, and those positions are under high scrutiny from a diversity point of view, for instance, information technology (IT). Since, for most of those prestigious positions, college is required, we can utilize data from 2017 NSCG, which warrants large sample sizes and detailed information about respondents' occupations. To investigate the current state of affairs, which is not clogged by previous discrimination, we restrict our examination to young college graduates, working in male-dominated job fields only. We define male-dominated fields as IT and mathematics, physical sciences, engineering, and management.

As opposed to the sparse literature on reverse gender wage differentials, we will test our hypotheses on a broad sample, including various positions across many organizations and job fields. Compared to the literature examining GWG among college graduates, our examination is specifically designed to uncover the presence of reverse gender wage differential. Therefore we believe that our research is unique and has the potential to enrich contemporary research.

For our thesis, we specify three hypotheses, which we will be statistically tested using data from 2017 NSCG. In the first hypothesis, we test whether adjusting for all relevant factors women in male-dominated job fields earn more than similar men.

Hypothesis 1

H_0 : *In male-dominated job fields women do not earn more than similar men.*

H_A : *In male-dominated job fields women earn more than similar men.*

In the second hypothesis, we restrict the first hypothesis to men and women working for large employers, which we define as having more than 1000 employers (see Section 4.2). We believe that large employers are more likely to provide advantages to women as they are under more scrutiny due to their size and prestige. Therefore we want to examine this group separately.

Hypothesis 2

H_0 : *In male-dominated job fields, women working for large employer do not earn more than similar men.*

H_A : *In male-dominated job fields, women working for large employer earn more than similar men.*

In the third hypothesis, we further restrict the second hypothesis to singles. We suppose that married men and women might be ill-comparable because women tend to marry sooner than men and marry richer partners (according to 1979 National Longitudinal Survey of Youth (NLSY) database based on author's calculations). This might reduce women's incentives to earn more money.

Hypothesis 3

H_0 : *In male-dominated job fields, single women working for large employer do not earn more than similar men.*

H_A : *In male-dominated job fields, single women working for large employer earn more than similar men.*

The thesis is structured as follows: In Chapter 2 we provide extensive literature review that in the first part covers the factors responsible for the existence of GWG, namely differences in human capital, occupations, behavior, and discrimination. Second part summarize the existing literature on the reverse gender wage differential and the third part summarize the works that had examined the GWG among college graduates.

Chapter 3 provides an overview of the methods used for our analysis. There we also discuss the reasons for the usage of particular methods as well as their shortcomings.

Next, Chapter 4 focuses on data used for the analysis and variables applied in regression models. The first part elaborates on the survey the data come from and its properties, criteria have been applied out to obtain the used sample, and also provides information about GWG observed in the sample. The second part informs about used variables and comments on their descriptive statistics and expected effects.

Chapter 5 presents and discusses the results of the models used. Furthermore, it evaluates the hypotheses stated in Chapter 1. We summarize the results in Chapter 6 and discuss their implications.

Chapter 2

Literature Review

The gender wage gap, which is essentially a difference in mean earnings of men and women, is a complex phenomenon that cannot be explained entirely by one theory, rather a combination of several theories is needed to account for the GWG. In this chapter, we will summarize those theories which contributed to explaining the GWG in the last decades and refer to the most important studies evaluating those theories. While the theoretical and empirical literature on the GWG is extensive, there are only a few studies studying the cases when comparable men earn less than comparable women, phenomenon, known as reverse gender wage differential, that has been only recently observed. We thus provide an outline of studies that found a reverse gender wage differential. As our study focuses on the GWG among college graduates, we further provide a summary of studies concerned with GWG of college graduates as well.

2.1 Gender Wage Gap Theory

The GWG started gaining attention in the 1960s. Since then, a large number of scholars studied the phenomenon to better understand the underlying reasons for its existence and especially to uncover how much of GWG is attributable to discrimination. Over the years, several dominant explanations for the existence of GWG emerged - human capital, job-related differences, behavioral differences, and discrimination. In the following sub-sections, we summarize those approaches by referring to studies from labor and experimental economics, sociology, and psychology. Where possible, we provide context by referring to official US statistics, and where appropriate, we provide historical comparisons to demonstrate changes in society over the years. Although there have been major changes in society in the US, namely an increase in labor force participation of females, the GWG did not disappear completely and only decreased from 38% in 1980 to 21% in 2010 as reported by Blau & Kahn (2017) and hence it remains very relevant topic today.

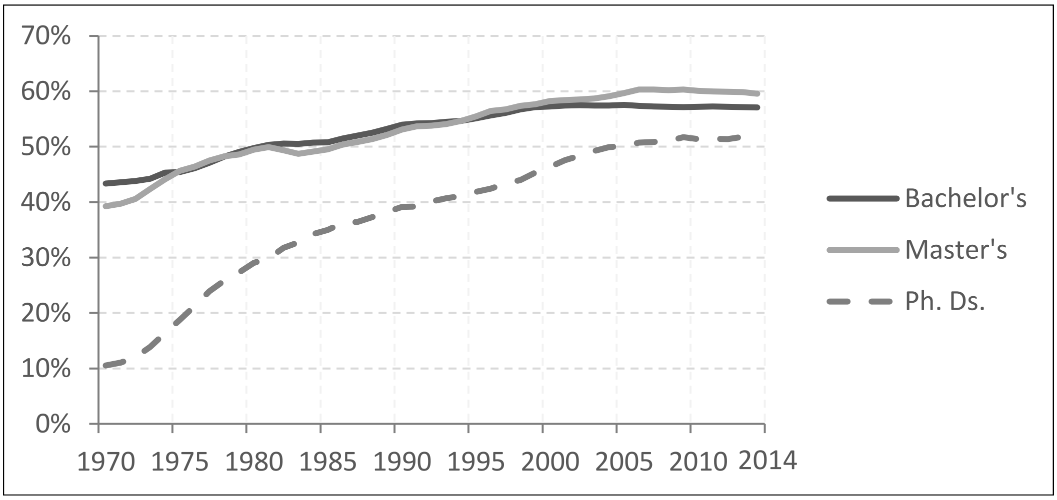
2.1.1 Human Capital

Human capital can be defined as the skills, knowledge, and experience possessed by an individual or population, used to produce economic value. Usually, human capital refers mainly to education, training, and experience. Necessarily, human capital acts as any other form of capital, in a way that it increases the productivity of labor. This increased productivity then leads to better remuneration of employees.

This theory has been notably examined by Mincer (1974), who showed that a large part of individual differences in earnings could be explained by the differences in the human capital of individuals. This clearly holds on the aggregate level as well, and as such it implies that differences in the average stock of human capital of men and women account for a part of GWG.

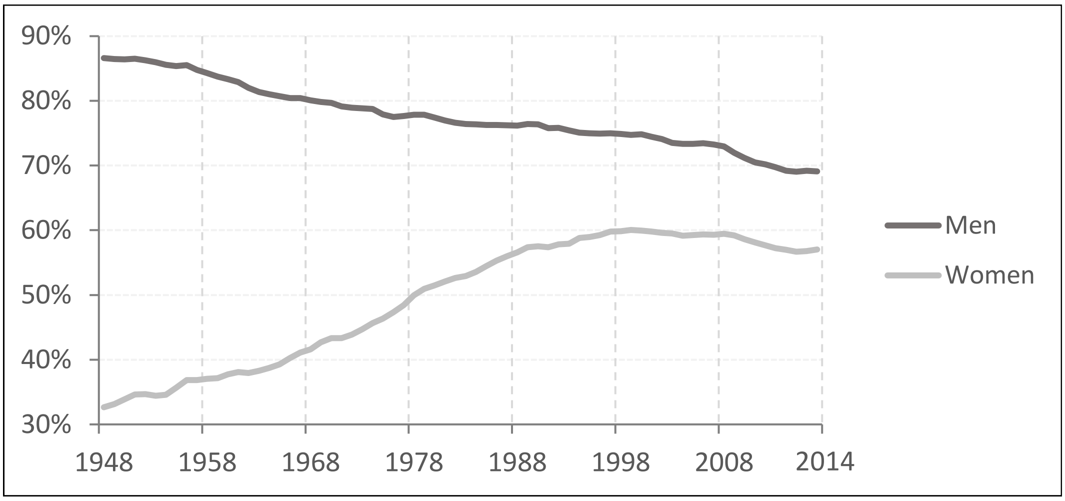
The portion of GWG explained by differences in human capital in the US has decreased significantly over the years. According to Blau & Kahn (2017), in 1980, they explained almost 23% of GWG. In 2010 they explained only 14% of much smaller GWG. The reason is that nowadays, women receive the majority of college degrees (Figure 2.1), and their labor force participation has greatly increased (Figure 2.2).

Figure 2.1: US Percentage of college degrees awarded to women by degree type (1970-2014)



Source: Bureau of Labor Statistics (2018)

Figure 2.2: US Labor force participation rate by gender (1948-2017)



Source: Bureau of Labor Statistics (2018)

On the other hand, some differences in human capital between men and women persist. Women are less represented in science, technology, engineering, and mathematics (STEM) related majors (Kena *et al.* 2015). For instance,in US, in 2015, women earned only 18%, 20%, 38%, and 42% of degrees in computer science, engineering, physical sciences, and mathematics and statistics, respectively. The differences in education then propagate into occupational segregation, which then accounts for some part of the GWG since STEM related fields pay more than average (Greenwood *et al.* 2011).

It should be noted that as Becker (1964) presumes, human capital is a result of previous investments. In essence, he posits that the employee will invest in his/her human capital as long as the net discounted returns to the investment are positive. This may, in turn, exacerbate the GWG if discrimination is present as it may decrease the willingness of women to invest in their human capital (since the discrimination decreases their returns to). It also implies that if women expect to spend less time working, i.e., due to parenting, they might lower their investment in human capital because their returns necessarily decrease.

2.1.2 Job-Related Differences

Men and women in the US clearly differ considerably in terms of jobs they tend to work in. While men are the majority in the primary and secondary sector, for instance in agriculture (74%), construction (90%), and manufacturing (70%), women are much more represented in services, especially in education (75%) and healthcare (78%) (U.S. Census Bureau 2019b). Within industries, men are clustered at better-paying firms and at better-paying positions (Petersen & Morgan 1995). According to Blau & Kahn (2017), job-related differences account for more than half of GWG.

In general, men work 5 hours more on average (Bureau of Labor Statistics 2018), which affects their wages directly, and it also means that they accumulate more experience, which might be associated with higher income growth (Gicheva 2013).

Overall, women tend to work in jobs that are more congruent with the role of primary caregiver, the role they disproportionately perform. For instance, in the US, 17.2% of employed women work in a part-time job compared to 8.4% percent of men (OECD 2019). Workers employed part-time in the US receive a 17% lower hourly wage than comparable full-time workers (Bardasi & Gornick 2008). Women are also more likely to receive family-friendly fringe benefits, i.e., childcare (Averett & Hotchkiss 1995; Lowen & Sicilian 2009). This might explain some portion of GWG as employers who pay for those benefits might be recouping some of the costs by lowering wages. However, there is mixed evidence whether employers recoup the costs associated with the benefits as suggested by economic theories or whether employers pay for those benefits with the hope of increasing productivity and decreased turnover (e.g., Lowen & Sicilian 2009; Baughman *et al.* 2003).

Lastly, women generally commute less, which might negatively affect women's chances of finding the best paying job or the job that offers the best career prospects. Unfortunately, most studies lack the data to control for the length of commuting, which might exacerbate the observed gender wage differential. According to Le Barbanchon *et al.* (2019), commuting may account for 10% of the observed GWG.

2.1.3 Behavioral Differences

Women and men, on average, differ in their behavior. In their meta-analysis, Croson & Gneezy (2009) found that women are more risk-averse and have more distaste for competition. Those are negatively related to choosing more difficult and prestigious academic tracks (Buser *et al.* 2014) and related to lower returns to human capital, i.e., education, tenure, and experience. Women are also more agreeable (Feingold 1994; Weisberg *et al.* 2011). This might manifest in their lower willingness to negotiate or in the results of their negotiations. (Leibbrandt & List 2015). Overall, Nyhus & Pons (2005) found the agreeableness to be negatively related to wages.

2.1.4 Discrimination

Ultimately, the most engrossing theory explaining GWG is discrimination. It is usually defined as behavior against an individual (or group) that is not based on objective consideration (Becker 1971). In case of wages, it means that the firm must be willing to pay more for the work done by members of one group as opposed to work done by members of the other group. In the context of GWG, the discrimination may also manifest itself by one gender having lower access to factors that have an impact on wages, e. g. education, on the job training, promotions.

The principal question of the majority of the first studies examining gender discrimination was whether women received unequal pay for the same work, practice which has been outlawed by the 1964 Civil Act. Sanborn (1964), using the 1950 US census data, found that GWG of 42% was largely caused by differences in hours worked, differences in experiences, and, most importantly, differences in the occupation. After adjusting, the resulting differential fell to 19%. Although the differential could have been interpreted as evidence for discrimination, Sanborn (1964) instead believed that it is a result of the very limited dataset.

Widely cited Oaxaca (1973) and Blinder (1973) went further to decompose the observed GWG into a part explained by the group differences in explanatory variables and a part explained by discrimination which consisted of the differences in returns to explanatory variables and the unexplained part. Using this approach, both estimated the discrimination coefficient to be around 30%. Unfortunately, their attempts to further decompose the discrimination part (in part attributable to returns to explanatory variables and the unexplained part) and to carry out the detailed decomposition on the level of explanatory variables, were later rejected (see, for instance Jones & Kelley 1984).

Malkiel & Malkiel (1973) examined a detailed dataset about 272 professional workers in the same corporation. They found that in the same positions and given the same individual characteristics, men and women were paid almost equally with gender wage differential only around 3%. However, they found that women with the same characteristics were assigned to lower positions. Hence they concluded that the mechanism through which discrimination arises was job assignments not paying unequally for the same work.

The most conclusive evidence that women and men indeed were paid equally for the same work was brought by Petersen & Morgan (1995). They analyzed wages in 16 industries obtained from surveys conducted between 1974 and 1983. Their detailed dataset allowed them to compare workers on industry, establishment, and occupation levels. They found that in the same establishment and in the same occupation, the differences in wages are negligible and that the GWGs were primarily result of women working in less paying establishments and in the lower-paid positions within the occupation itself.

The second principal topic examined by researchers was discrimination in job allocation among men and women. In Brown *et al.* (1980), the authors criticized the approach of Oaxaca (1973) and others, for not accounting for the differences in the occupational distribution of men and women because those differences might be itself result of discrimination rather than fair process. Using their model, they showed that there is a large difference between predicted and actual distribution of women.

Bielby & Baron (1986) argued that partially the occupational segregation might be explained by the theory of statistical discrimination. The theory posits that men and women are perceived to have different average productivity for jobs involving different activities. In their analysis, they estimated probabilities that women were excluded from a job, given a particular characteristic. They found several activities with a large impact, e.g., physical strength requirement increased the probability of exclusion of females by 32%, on the other hand, the need for finger dexterity decreased the probability by 26%. The results suggested that there was a pattern explaining the segregation. However, the authors were not able to prove that the behavior of employers would be rational and optimal.

Maume Jr (1999) presented ample evidence using the Panel Study of Income Dynamics (PSID) that women tend to receive promotions less often than similar men, a phenomenon known as a glass ceiling. Moreover, he found that men are more likely to be promoted, the higher the percentage of females in the occupation.

Field experiments, examining the likelihood of applicants being contacted, show that hiring processes often discriminate against one or other gender. Those experiments are beneficial because they provide an idea about discrimination in the workplace. Needless to say, the decisions of firms to reach back to an employee are very relevant for occupational segregation. While there are studies that found lower call back rates for women (e.g., Neumark *et al.* 1996; Petit 2007), other studies using different settings found lower call back rates for men (e.g., Riach & Rich 2006; Booth & Leigh 2010). There is some evidence that it may depend on the representation of the particular gender in the occupation (e.g., Booth & Leigh 2010; Carlsson & Eriksson 2017). Consequently, this may bolster already present occupational segregation.

There is also evidence that occupational segregation is at least partially voluntary. Most notably, Stoet & Geary (2018) showed that the segregation of women in STEM colleges was higher in countries with higher gender equality, i.e., Norway or Sweden. Authors suggested that the reason for this paradoxical finding was that more gender unequal countries were, on average, poorer. Women in those countries might thus pursue STEM degrees to financially secure their future while in richer countries, women can achieve it also by choosing non-STEM majors. Naturally, it can be expected that the segregation in college majors largely determines the observed occupational segregation as showed by Rosenfeld (1983).

Lastly, we should mention the different impacts children have on the earnings of women and men. It is not surprising that the earnings of women suffer as a result of having a child. After all, women are usually the primary caregivers and, as a result, work less and have less experience. However, Budig & England (2001) found that this accounts only for one-third of the 7% per child penalty. An economic experiment conducted by Correll *et al.* (2007) showed that women with children receive fewer callbacks from potential employers compared to men with children, which suggested that employers indeed discriminate against them. On the other hand, there is evidence that fatherhood has a large net positive effect on compensation. Hodges & Budig (2010) found, examining the 1976 NLSY, that men receive an 8% premium for fatherhood all things being equal.

2.2 Reverse Gender Wage Differential

Most of the GWG research found a large gender wage differential even after adjusting for relevant factors (Weichselbaumer & Winter-Ebmer 2005). It should thus not come as a surprise that the question of whether under specific conditions, women earn more than men, had laid dormant for years and only recently started getting little attention. However, the topic still remains uncharted territory. In this section, we shall summarize the development and the few studies that have been published so far.

One of the pioneering works was written by Gayle *et al.* (2012). They examined how do executive compensation and job mobility differ by gender. Compared to earlier works on this topic, they used data about compensation of executives that were linked with data about their work history and demographics. They also created much more robust measures of executive compensation, also measuring indirect remuneration, i.e., stock options. Using this unusually large dataset (more than 16000 executives for the period 1991-2006), they found that adjusting for experience, position, returns, and education, the median total annual compensation of women executives was higher by $92000 which was 4.97% of the mean annual compensation of the executives in the sample.

Analogous study was conducted by Hill *et al.* (2015). In their paper, they examined whether female and ethnic minority CEOs were disadvantaged by their minority status or whether they were receiving a premium. To test this hypothesis, they used data from 1996-2006 about CEO compensation in 2225 firms. After controlling for various characteristics, i.e., firm size, return on assets, CEO tenure, the results showed that women CEOs received a 6% wage premium over similarly abled men.

Another work that showed that there exist cases of women preferred over men was written by Williams & Ceci (2015). They examined whether general knowledge regarding the discrimination of women in STEM academia was justified. In a randomized experiment, 873 tenure-track faculty members in

STEM compared various profiles of male and female candidates for an assistant professor position in two fields where women are underrepresented – economics and engineering. The experiment revealed 2:1 preference for female applicants regardless of the reviewer's gender. The results are in sharp contrast with the widely known, similar experiment ran by Moss-Racusin *et al.* (2012), which found bias against female students. The main difference between the papers is the prestigiousness of the job experimental candidates applied for – assistant professor vs. laboratory manager and the academic achievement of the candidates – Ph.D. vs. undergraduate degree.

Building on the works of Gayle *et al.* (2012), Hill *et al.* (2015), and Williams & Ceci (2015), Leslie *et al.* (2017) formulated a hypothesis that women receive premium when their potential is high. They further argued that the premium depends on the perceived diversity value of the employee. Since women are usually well represented in the lower echelons, they are perceived as high in diversity value only if they have the necessary potential to reach higher positions in the organization. This explains why not every woman receives the premium. To test the hypotheses, they conducted a survey, which was followed up by an experiment.

In the survey, a questionnaire was sent to the employees of one of the Fortune 500 organizations. The data collected on the respondents included pay, potential, performance, and range of control variables known to affect wages. The results showed that high potential women earned 8% more than comparable men, while for moderate-potential women, there was no difference, and low-potential women earned 8% less.

Following the results of the survey, the authors conducted an experiment aiming to evaluate competing theories claiming that high-potential women might earn a premium due to being perceived as more competent, agentic, warmer, or unique than high-potential men. In the experiment, 270 business school graduates evaluated made-up resumes where the gender of the applicant systematically varied. The results showed strong support for both hypotheses – women being perceived as higher in diversity value when their potential was high, and women receiving a premium for their diversity value rather than other factors.

2.3 Gender Wage Differential Among College Graduates

College graduates differ markedly from the rest of the population. Most importantly, their wages are, in general, much higher, and their unemployment rate lower. For instance, in 2017, the US college graduates with bachelor's degrees earned 64% more than high school graduates and more than double what those without high school earned. Only 2.5% of graduates with bachelor's degrees were unemployed compared to 4.6% of high school graduates and 6.5% of those without completed high school (Torpey 2018). Therefore college graduates should, in general, be able to stand up to discrimination more easily and have more freedom in choosing potential employment. It is thus meaningful to look at the gender wage differentials among college graduates separately. Herein we provide a summary of the studies concerned with gender wage differential among college graduates in the US.

Black *et al.* (2008) analyzed data from the 1993 NSCG. Using non-parametric matching, they found that the resulting gender wage differential among those with high labor market attachment and English speaking parents was 9% for white women. However, for Hispanic and Asian women, the differential was only 0.4% and 2.6%, respectively, comparing like for like.

The 1993 NSCG data was also analyzed by Morgan (2008). The study examined the GWG using a subset of 1993 NSCG data, selecting only respondents who graduated in the five years before 1989, worked more than 20 hours/week, and at least 26 weeks. It also excluded self-employed respondents. Using this sub-sample, she found that the choice of major had a large impact on gender wage differential. While for majors in humanities and business administration (excluding accounting) the gender wage differential was 10% and 8.6% respectively, for majors in engineering and mathematics the gender wage differential was insignificant and close to zero. The study also found a large difference in the size of the gender wage differential between graduates and undergraduates. Controlling for hours worked, work experience, and college major, the gender wage differential was just 1.9% among graduates, while among undergraduates, it was 5.1%. The results of Morgan's work suggested that it was advantageous for women to venture into traditionally male occupations and to invest in a higher degree and that there was little to no discrimination of college-educated women even if we allow for discrimination in job allocation.

Bertrand *et al.* (2010) studied career trajectories of master of business administration (MBA) graduates who graduated between 1990 and 2006. Controlling for pre-MBA characteristics, MBA performance, experience, hours worked, and reasons for choosing job characteristics, their analysis revealed that gender wage differential in the first year after finishing MBA was just 2.5% and insignificant. However, after the first year, the differential was much larger: 6.0% and significant. The results suggest that starting wages are very similar, but men progress faster.

Michelmore & Sassler (2016) analyzed the gender wage gap among college graduates working in STEM fields. They used data from Scientists and Engineers Statistical Data System (SESTAT) from 1995 to 2008. They found that adjusting for occupation, degree type, and family situation, the gender wage differential among Asians and Blacks was just 1% and insignificant. Among Whites and Hispanics, the gender wage differential was 4% and 8%, respectively.

Chapter 3

Methodology

In our analysis, we aim to explore the nature of GWG among young college graduates. Ultimately, we would like to find out whether, in the male-dominated fields, women earn more than comparable men. To find out, we need to adjust the GWG for the differences in individual and job-related characteristics among men and women that have an impact on wages and, consequently, GWG. In the end, we would like to obtain the estimates of the adjusted GWG (used interchangeably with gender wage differential) for the male-dominated fields. To test our research hypotheses (see Chapter 1), we need to obtain the gender wage differentials using the conditions specified by the hypotheses for each of the male-dominated job fields in a way that the differentials can be statistically tested.

It has become a standard practice in the GWG literature to estimate the gender wage differential using a pooled linear model, estimated using ordinary least squares (OLS). The pooled model can be conveniently extended to obtain the estimates of the within job field gender wage differentials, by including dummy variables for job fields and interaction terms of those variables with the *female* dummy. However, the pooled model has a few shortcomings that cause it not to suit our needs. Firstly, we believe that the effect of explanatory variables will likely differ among job fields. It is not hard to imagine scenarios in which the differences in the effects of particular variables among job fields might be striking. For instance, people working in IT sector are often paid an hourly wage while in other fields, e.g., management, fixed salaries might be much more common, therefore in IT, the return to hours worked should be much higher. To account for that in a pooled model, we would have to include interactions terms of the job fields with other explanatory variables, which would tremendously increase the number of explanatory variables and, in turn, made the interpretation cumbersome. Additionally, since our data include a detailed breakdown of job fields into smaller categories (we refer to them as detailed job categories), we would like to use them to account for differences in job allocation within a job field between genders, while still being able to obtain gender wage differentials for the male-dominated job fields. Accounting for detailed job categories in the pooled model would further increase the complexity of the pooled model and made the testing of our hypotheses unnecessarily difficult.

Therefore for our analysis, we are going to estimate the model separately for each male-dominated job field as it naturally allows the effect of variables to differ by job field and conveniently enables us to include detailed job categories as explanatory variables. Those separate models will then be used to draw inference about the validity of our hypotheses. The pooled model will only be used to provide some comparison with studies utilizing this model.

In this chapter, we provide a detailed specification of the pooled model, comment on the difference in specifications of the separate models. We further summarize likelihood ratio (LR) test that will be used to test whether separate models for each job field fit the data significantly better than one pooled model. Moreover, we provide a description of Breusch-Pagan (BP) test that will be used to test whether there is heteroskedasticity present. Lastly, we shall outline the Oaxaca-Blinder decomposition (OB), which will be used to decompose the within job field GWG into a part attributable to differences in endowments and a part attributable to differences in remuneration.

3.1 Pooled Model

The first model that is going to be estimated is pooled linear regression of the logarithm of weekly wage, i.e., we estimate the model using the whole sample for all job fields. We will run five specifications of the pooled model. In the specification (1), we include all explanatory variables concerning personal characteristics (and not job-related characteristics) as defined in Section 4.2. To be more specific, we include demographic variables, educational variables, and family-related variables. In the specification (2), we add the explanatory variables for job-related characteristics, specifically employer variables, job field variables, and variables concerned with the relatedness of the principal job to highest degree (HD) of the respondent. Specification (3) allows the estimated effects of working in a particular field to differ by gender through the addition of interaction terms of *female* dummy and job field variables. In specification (4), we allow the effects of employer size to vary by gender using the interaction terms with the female dummy and respective variables. Specification (5) further allows the effect of being in marriage-like relationship to differ by gender. Specifications (3)-(5) are designed to correspond to our hypotheses 1-3. Even though we shall not test the statistical significance of the pooled model's estimates, we want to give an idea about the differences in the predicted coefficient of the pooled model compared to the separate models even if the models are very different.

The model in the specification (5), specifications (1)-(4) are its more restrictive derivatives, can be written as

$$\begin{aligned} log(y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \beta_{k+1} x_{k+1i} \ldots \beta_l x_l + \\ + \beta_{l+1} x_{1i} * x_{k+1i} + \cdots + \beta_m x_{1i} * x_l + \beta_{m+1} x_{1i} * x_{2i} + \cdots + \beta_n x_{1i} * x_{si} + \varepsilon_i, \\ i = 1, \ldots, N, \quad (3.1) \end{aligned}$$

where $log(y_i)$ is natural logarithm of weekly wage, x_{1i} stand for *female* dummy, $x_{2i}, \ldots, x_{ki}$ represent explanatory variables as described in Section 4.2 except dummy variables for job fields, $x_{k+1i}, \ldots, x_{li}$, $x_{1i} * x_{k+1i} + \cdots + x_{1i} * x_{li}$ represents interaction terms between female and job fields, and $x_{1i} * x_{2i} + \cdots + x_{1i} * x_{si}$ stand for interaction terms between female and employer size, and female and being in marriage-like relationship. ε_i is i.i.d. error term with zero mean, β_0 stand for intercept, $\beta_1 + \cdots + \beta_n$ represent the effects of respective explanatory variables on explained variable, $i \in 1, \ldots, N$ represents an observation from sample of size N.

3.2 Separate Models for Job Fields

To test our hypotheses, we are going to estimate three specifications of the separate models for each of the male-dominated job fields, as defined in Chapter 1, separately. All of those three models are akin to the pooled models presented in the previous section with a few exceptions - they clearly do not include job fields dummy variables. However, they control for detailed job categories, which are not present in the pooled model. Further, the second and third specifications of the separate models include variables representing the groups of our interested directly (In the first specification, it is just *female* dummy.) instead of interaction terms to represent them. In the second specification of the separate models, those are all combinations of gender and employer size and in the third specification, all combinations of gender, employer size, and relationship status. This approach allows us to test our hypotheses directly, but the result presented this way might be less legible. The first specification of the model can be written as

$$log(y_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \cdots + \beta_k x_{ki} + \varepsilon_i \qquad i = 1, \ldots, N, \quad (3.2)$$

where $log(y_i)$ is the natural logarithm of weekly wage, x_{1i} stands for *female* dummy, x_{2i},x_{3i} stand for employer size dummy variables, x_{4i} is a dummy variable for being the marriage-like relationship, and $x_{5i} \ldots x_{ki}$ are the control variables including detailed job categories (as defined in Section 4.2). ε_i is i.i.d. error term with zero mean, β_0 stand for intercept, $\beta_1 + \cdots + \beta_k$ represent the effects of respective explanatory variables on explained variable, and $i \in 1, \ldots, N$ represents an observation from sample of size N. This model will be used to test the first hypothesis, i.e., in male-dominated job fields, women earn more than similar men.

In the second model, the *female* dummy and employer size dummies are replaced by variables representing groups created from the combination of *female* and employer size variables. To be specific, we include dummy variables for females working for small employers, mid employers, and large employers and males working for small employers, and mid employers. Males working for a large employer is the reference group against which we want to compare females working for large employers to test our second null hypothesis, i.e., in male-dominated job fields, women working for large employers earn more than similar men. For the sake of brevity, those variables are presented in the format *Gender* X *Employer size*; for instance, males working for a small employer will be represented by Male X Small dummy variable.

In the third model, the group variables we created in the second model are further differentiated by relationship status. Hence the dummy variable for being in a marriage-like relationship is removed. Those group variables follow the same naming convention as the variables in the second model - *Gender* X *Employer size* X *Relationship status*. The reference group is single males working for large employers. The third model should allow us to test the third hypothesis, i.e., in male-dominated job fields, single women working for large employers earn more than similar men.

3.3 Testing Models Assumptions

Likelihood Ratio Test

To test whether the separate models for job fields jointly fit the data better than the pooled model for all job fields, we shall use the LR test. LR statistics is defined as:

$$LR = 2(\mathcal{L}_{ur} - \mathcal{L}_{r}), \tag{3.3}$$

where $\mathcal{L}_{ur}$ is the log-likelihood of the unrestricted model, and $\mathcal{L}_{r}$ is the log-likelihood of the restricted model. LR statistics has an approximate chi-square distribution (Wooldridge 2016). In our case, the unrestricted model is represented by the first specification of the separate models, without the detailed job categories, estimated for each job field. The restricted model is a specification (3) of the pooled model.

Breusch-Pagan Test

We will use the BP test to test whether our models suffer from heteroskedasticity. Based on the result of this test, we shall decide whether we will have to mitigate the issue using heteroskedasticity robust standard errors.

Firstly, the squared model's residuals ($\hat{u}^2$) are regressed on the model's explanatory variables

$$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k. \tag{3.4}$$

Then the BP test statistics can be calculated. It is essentially just LM statistics defined as

$$LM = nR^2_{\hat{u}^2}, \tag{3.5}$$

where n is the sample size and $R^2_{\hat{u}^2}$ is the R^2 from the residuals regression. The LM statistics has an approximate chi-square distribution (Wooldridge 2016).

3.4 Oaxaca-Blinder Decomposition

We shall decompose the within job field GWG into a part that is explained by differences in characteristics between men and women (endowment effect) and a part that is explained by differences in returns to the characteristic between men and women (remuneration effect) using OB (Blinder 1973; Oaxaca 1973). Essentially, in OB, a model of logarithms of wages is estimated for both males and females separately, and the difference in mean wages can be then written as a difference in the mean level of explanatory variables and corresponding estimates.

The models for OB is defined as

$$lnW_i^J = X_i^J \beta^J + \varepsilon_i^J, \tag{3.6}$$

where lnW_i^J is the natural logarithm of wages, X_i^J is a vector of explanatory variables, β^J is the vector of the effects of explanatory variables on wages, ε_i^J is i.i.d. error term with zero mean, and $J \in \{M, F\}$ where M and F are males and females, respectively. The mean difference in wages of males and females is then defined as

$$\overline{lnW^M} - \overline{lnW^F} = \overline{X^M}\hat{\beta}^M - \overline{X^F}\hat{\beta}^F, \tag{3.7}$$

which can be rewritten into

$$\overline{lnW^M} - \overline{lnW^F} = \overline{X^M}\hat{\beta}^M - \overline{X^F}\hat{\beta}^M + \overline{X^F}\hat{\beta}^M - \overline{X^F}\hat{\beta}^F, \tag{3.8}$$

which can be rearranged into

$$\overline{lnW^M} - \overline{lnW^F} = (\overline{X^M} - \overline{X^F})\hat{\beta}^M + (\hat{\beta}^M - \hat{\beta}^F)\overline{X^F}, \tag{3.9}$$

where the term $(\overline{X^M} - \overline{X^F})\hat{\beta}^M$ represents the part of the difference in mean wages attributable to differences in explanatory variables (endowment effect). It could be interpreted as the predicted difference in mean wages had women had the same returns to the explanatory variables or, in other words, were remunerated equally.

The term $(\hat{\beta}^M - \hat{\beta}^F)\overline{X^F}$ represents the part of the differences in mean wages attributable to the difference in returns to explanatory variables (remuneration effect), and it also encompasses the differences in the unexplained part, i. e. intercepts and is referred to as remuneration effect. It could be interpreted as the predicted difference in mean wages had men had the same endowments as women. It is often interpreted as part attributable to discrimination. However, one should generally be judicious with this interpretation as some part of this is also attributable to differences in the unexplained part, and as such, it might be partially a result of the under-specified model.

Chapter 4

Data and Variables

In this chapter, we discuss the dataset used in our thesis and the criteria applied to obtain the sample for our analysis. We further provide an overview of the GWG, average wage, and female share by job fields in the analysis sample and comment on the used variables in our analysis and their summary statistics.

4.1 Data

To evaluate our hypotheses, we are going to use data from the 2017 National Survey of College Graduates (NSCG). It is a survey conducted by the National Science Foundation, which is an independent US federal agency financing around 24% of all research supported by the US government. The survey focuses solely on college graduates who earned at least a bachelor's degree before 2016.

The main advantage of this survey over other surveys is its unique focus on college graduates with special emphasis on young graduates, which warrants large sample sizes that will allow us to conduct a detailed examination of GWG across job fields.

Several sample selection criteria are applied to obtain the final sample. Since our focus is on working young college graduates, only employed people younger than 31 years old are included. Additionally, people who worked less than 20 hours a week, earned less than 10 dollars per hour, and their annual pay was based on less than 26 weeks, are excluded. The goal of these criteria is to exclude highly specific cases that would constitute only a small part of our sample but might be difficult to explain using traditional variables. Furthermore, people who are self-employed or did not graduate from US university are excluded. The income of the self-employed people tends to fluctuate in time and often confound with a return to capital. People who graduated from non-US university might be difficult to compare to US graduates as their education can differ considerably.

The resulting sample consists of 10548 observations, of which 5008 (47.5%) are women. Women are less represented in our sample compared to the population of college graduates because women are still primary care-givers, and they also study longer than men. Hence they are more often excluded from the sample given the criteria as mentioned earlier.

Table 4.1 shows that the overall unadjusted GWG in our sample is 18.2% which is close to values usually reported by research (e.g., Blau & Kahn 2017). Nevertheless, the gap might still seem surprisingly large, given that only young graduates are considered. However, it should be noted that we calculate the gap from weekly wages as opposed to hourly wages (see Section 4.2).

The GWG vary considerably among job fields, and those GWGs are lower than the overall GWG because GWG reflects primarily the differences in job distribution (see the discussion in Section 2.1.4) which is apparent as women are much more represented in fields that pay less than average, i.e., teachers or social workers and counselors, while men have the highest representation in the highest paying fields, i.e., IT and mathematics, management, engineering. The same pattern of women being more represented in worse paid positions will likely be observed within fields as well. This would explain why the GWG is comparatively high in social and related sciences (around 15%), among financial specialists, and all other, as those fields are not really well defined and encompass a large variety of jobs. Surprisingly, in physical sciences and in healthcare, the GWG is slightly negative -2.7% and -2.4%, respectively, which means that women in those fields earn 2.7% and 2.4% more on average, respectively. This rather surprising result might in healthcare be partially explained by having management as a separate field since the disparity between the overall representation of women and representation in the management positions is known to be striking there (Russell *et al.* 2019).

Table 4.1: Gender wage gap, female percentage, and sample distribution by job fields

Job field	Female GWG (%)	Average weekly wage (USD)	Female share (%)	Percent total (%)
All fields	18.2%	1311	47.5%	100.0%
IT and mathematics	4.2%	1703	25.2%	12.8%
Biological and other life sciences	5.6%	945	53.7%	3.9%
Physical sciences	-2.7%	1036	41.8%	2.5%
Social and related sciences	14.5%	1252	68.3%	2.5%
Engineering	6.1%	1475	22.7%	22.8%
Healthcare	-2.4%	1292	80.1%	11.7%
Management	6.5%	1692	33.5%	4.4%
Teachers	1.5%	973	64.2%	7.7%
Social workers and counselors	8.1%	888	82.9%	4.3%
Sales and marketing	12.3%	1220	50.0%	4.6%
Financial specialists	14.5%	1485	48.0%	3.4%
All other	14.5%	1122	56.4%	19.4%

Source: Sample of NSCG 2017. Author's computations.

4.2 Selected Variables

In this section, we discuss the explained variable and explanatory variables used in the analysis. Firstly we discuss the variables of our interest on whose effects we shall comment in Chapter 5. Then we discuss the control variables which are used in the analysis to account for an individual and job-related differences but whose effects shall not be presented in Chapter 5 but will be available in the Appendix A. We also discuss the summary statistics of all variables. Summary statistics of job field variables had already been discussed in the previous section and are summarized in Table 4.1. Summary statistics of continuous variables are presented in Table 4.2 and summary statistics of dummy variables are presented in Table 4.3.

Since the implicit aim of this thesis is to evaluate gender differences in wages, the explained variable is the logarithm of weekly wages (*log(weekly wage)*). The weekly wage is calculated from the reported annual salary and number of weeks the annual salary is based on. Compared to many previous works utilizing the NSCG (e.g, Black *et al.* 2008; Michelmore & Sassler 2016), we do not use hourly wages and instead include logarithm of hours worked per week as an explanatory variable (*log(hours worked)*). Men in our sample work 2.3 hours longer and earn $230 more weekly.

The choice of wage measure is crucial as it has a direct influence on GWG and, consequently, gender wage differential. It can be argued for both of those approaches. To do so fairly, we first need to specify the limitations and assumptions we work with. Firstly, NSCG does not collect information on whether a respondent is paid a fixed salary or hourly wage. Consequently, we also do not know the number of hours the fixed salary is based on, rather NSCG collects information about the average hours worked weekly. We can approximate the percentage of salaried employees using the Current Population Survey data (U.S. Census Bureau 2019a). From these data, we calculated that 68% of employed college graduates older than 25 years are paid fixed salaries.

The usage of hourly wage is clearly superior for explaining the wages of workers that are paid an hourly wage. It might also be argued that salaried employees who work more than they are obliged to by their contract are worse off, and this should then be directly taken into account. On the other hand, we argue that many, if not most employees, work longer hours for which they are not remunerated because of faster wage growth, higher likelihood of promotions, or simply because they enjoy the job.

Usage of weekly wage with hours worked as the explanatory variable, on the other hand, is superior in explaining the wages of employees paid a fixed salary because they are not remunerated for hours worked but are instead given fixed amount as per their contract. However, the main advantage of the weekly wage approach is that it does not assume the relationship between hours worked and wage outright but estimates it using OLS. We believe that this will reduce the size of residuals and consequently result in a better fit. This also means that we shall have different estimates of this relationship in the separate models for each of the male-dominated job fields. We suppose that there will be substantial differences in those estimates as, for instance, in IT and mathematics, people are often paid hourly wages while in management, they will likely be paid fixed salaries.

The fundamental explanatory variable is the gender of graduates represented by *female* dummy (in interaction terms represented by F). The estimated effect of this variable on the weekly wage will be used to draw inference about the differences in wages men and women receive. In the second and third specification of the separate models, the *female* dummy is morphed into group variables with employer size in the second model and *marriage* dummy in the third model (see Section 3.2).

Table 4.2: Summary statistics - continuous variables

Variable name	Mean	SD	Median
Males			
Weekly wage (USD)	1420.5	886.5	1346.2
Hours worked weekly	43.8	8.6	40.0
Years from graduation	4.6	2.3	5.0
Females			
Weekly wage (USD)	1189.7	1050.7	1057.7
Hours worked weekly	41.5	8.7	40.0
Years from graduation	4.4	2.3	4.0

Source: Sample of NSCG 2017. Author's computations.

Table 4.3: Summary statistics - dummy variables

Variable Name	Male (%)	Female (%)	Total (%)
Ethnicity			
White	58.8	51.5	55.3
Black	5.7	9.8	7.6
Asian	20.3	16.8	18.6
Hispanic	10.5	15.0	12.6
Other ethnicity	4.7	6.9	5.8
Highest degree			
Bachelor's	65.9	58.9	62.6
Master's	28.1	35.0	31.4
Ph. D.	3.7	2.6	3.2
Professional	2.3	3.5	2.8
Certificate or license	24.9	39.9	32.1
Still enrolled at university	10.9	13.1	12.0
Type of employer			
For-profit	70.2	47.5	59.4
Government	19.2	27.4	23.1
Non-profit	9.1	24.3	16.3
Other	1.4	0.8	1.2
Size of employer			
Small employer (<100)	18.1	17.3	17.7
Mid employer (100-1000)	20.9	22.4	21.6
Large employer (1000+)	61.0	60.3	60.7
Relation of job to HD			
Related closely	62.1	62.3	62.2
Related somewhat	26.8	24.7	25.8
Not related due pay	7.0	7.8	7.4
Not related due comfort	2.7	3.7	3.2
Not related other	1.4	1.4	1.4
Marriage			
In marriage like relationship	50.3	50.8	50.5
Not in marriage like relationship	49.7	49.2	49.5
Children	15.3	17.1	16.2

Source: Sample of NSCG 2017. Author's computations.

Employer size is represented by *mid employer* and *large employer* dummies with the *small employer* as reference group. For our analysis, we define *large employer* as having more than 1000 employees, *mid employer* more than 100 employees and less than 1000, and *small employer* less than 100. In the second and third specification of the separate models, those variables are morphed into group variables with *female* and *marriage* dummies. Employer size variables are crucial for evaluating whether women in male-dominated job fields in large firms earn more than men. We expect large firms to pay more than smaller firms and to remunerate women better as smaller firms do not usually face pressure to have a more diverse workforce nor to remunerate equally.

Being in the marriage-like relationship naturally influences the economic situation of an individual in a major way and, in turn, influences the decisions individual undertakes. Given that women tend to marry older and richer men (according to 1979 NLSY database based on author's calculations), we expect marriage to affect men and women unevenly. Because of that, we think that married men and married women might be ill-comparable, and it might thus make sense to examine the gender wage differential among single men and women. Being in a marriage-like relationship is represented by dummy variable *marriage*. In the third specification of the separate models, the *marriage* dummy is morphed into the group variables with employer size variables and *female* dummy (see Section 3.2 for more detail). This should allow us to examine our third hypothesis, i.e., in male-dominated job fields, single women working for large employers earn more than similar men.

Job field variables, as defined in 4.1, with *all other* as a reference group, are naturally present only in the pooled model. However, only the effects of dummies for male-dominated fields and their interaction terms with *female* are presented in Chapter 5.

To obtain estimates of gender wage differential, we need to further adjust for the difference in individual characteristics and job-related characteristics known to affect wages. We refer to those variables as control variables. Control variables used in our models can be divided into three categories: ethnicity variables, educational variables, and job-related variables.

The ethnicity variables consist of ethnicity dummies *Black, Asian, Hispanic,* and *Other ethnicity* with *White* as the reference group. We need to control for ethnicity because there are sizable differences in ethnicity representation among genders in our sample (Table 4.3), and research shows that Asians and Whites earn significantly more on average than Blacks and Hispanics (Patten 2016). We also do not want the estimates of gender wage differential to confound with any race discrimination that might be present.

With regards to education, the model controls for the type of HD, years from obtaining HD, whether the respondent is still enrolled in university, and whether the respondent has any professional certificate or license. To control for HD we include *master's*, *Ph. D.*, and *professional* degree dummies with *bachelor's* as the reference group. We control for years from obtaining the HD (*years from graduation*), which is our proxy for frequently used years of experience in other studies. The mean value of *years from graduation* is slightly higher for men since women in our sample have a master's degree more frequently (Table 4.3). Since women more frequently aim for master's degrees, they are more often still enrolled at university, which should negatively affect their wages as they might not get the best jobs or are not able to focus on the job fully. We control for this using the dummy variable *still enrolled at university*. Women also more often have some professional certification or license. This might be caused by a higher representation of women in job fields that require those certificates, i.e., healthcare. However, we still control for this difference using *certificate or license* dummy.

Next, we control for variables related to the principal job - a type of employer, the extent to which current job is related to HD, and the reasons for unrelatedness of the principal job to HD, and for detailed job categories. There are large disparities in the type of employers men and women work for. Women more often work for the government or non-profit employers (Table 4.3), which is to be expected given that they work much more frequently in job fields such as healthcare or education. However, this holds also for male-dominated fields such as engineering or IT and mathematics. We expect government and non-profit employers to pay significantly less compared to for-profit employers. We control for the type of employer using *government, non-profit,* and *other* employer dummy variables with *for-profit* employers as the reference group.

Although we expected women to be more often forced to work in a job not related to their HD, this is not the case in our sample. The share of women working outside of HD is very similar for men. The NSCG asks its respondents about the degree to which their principal job is related to their HD - *related closely*, *related somewhat*, and *not related*. Furthermore, NSCG asks respondents about the two most important reasons why they work in a job not related to HD. Based on the question we have created two dummy variables - *not related due pay* represent an answer that pay was one of the reasons for working outside of HD, *not related due comfort* represents an answer that either job location, job conditions, or family-related reasons were the reasons for working outside of HD and none of the reasons was pay. Other cases are represented by *not* *related other*. The final set of variables in the model is then *related somewhat*, *not related due pay*, *not related due comfort*, and *not related other* dummy variables with *related closely* dummy as the reference group.

We have split the *not related* answer into more variables to better account for cases when respondents have to work in job fields outside of the field of their HD, which we expect to have severe implications for their wages. We expected women to more often work outside of the field of their HD due to family-related reasons or convenience compared to men, but that is not the case. They might still opt for jobs with better locations or conditions in the field of their HD, which we, unfortunately, cannot control for using the data in NSCG.

Lastly, NSCG uniquely allows us to control for really detailed job categories, which should further increase the explaining power of estimated models and also account for allocation differences within job fields. We control for these only in the separate models for male-dominated field (see Section 3.2). Since there are many of those categories for each job field, we do not present their summary statistics, and their effects are not discussed in the Chapter 5 nor displayed in the Appendix A. To give an idea, how does the split of the job field to detailed job categories looks like we provide an example of a breakdown of IT and mathematics job field in Table 4.4.

Table 4.4: Example: Detailed job categories - IT and mathematics

Detailed job categories for IT and Mathematics
Computer & information scientists, research
Computer network architect
Computer support specialists
Computer system analysts
Database administrators
Information security analysts
Network and computer systems administrators
Software developers - applications and systems software
Web developers
Other computer information science occupations
Computer engineers - software
Mathematicians
Operations research analysts, including modeling
Statisticians
Other mathematical scientists

Chapter 5

Results

In this chapter, we first present the results of BP test and LR test to see whether our assumptions about heteroskedasticity still being present and estimating models separately for each job field have been correct. Then we estimate five specifications of the pooled model and comment on their results. Since we believe that the pooled model is inferior to a separate model for each job field, we shall not use the pooled model to evaluate our hypotheses. We include it to allow comparison with works utilizing this more traditional approach. Next, we estimate three specifications of the separate models for each job field and use the results to evaluate our hypotheses. We will conclude with the Oaxaca-Blinder decomposition of the GWGs in male-dominated fields.

Due to a large number of explanatory variables used in our models, we shall not display the effects of all variables in this chapter. Effects of all variables (except *detailed job categories*) are presented in Appendix A.

5.1 Testing Models

Heteroskedasticity Testing

We suspect that heteroskedasticity might still be present even though we use the logarithmic transformation of wages. We test for it using the Breusch-Pagan test. The results are displayed in Table 5.1. The results imply that there is heteroskedasticity present on 0.05 level in the pooled model and all specification of the separate models for IT and mathematics and engineering. We were not able to disprove the homoskedasticity null hypothesis for physical sciences and management. For consistency, we use robust standard errors for inference in all our models.

Table 5.1: Breusch-Pagan test

Model	Breusch-Pagan p-value
Pooled - all fields	<0.001
Separate - first specification	
Separate - IT and mathematics	0.003
Separate - physical sciences	0.527
Separate - engineering	<0.001
Separate - management	0.963
Separate - second specification	
Separate - IT and mathematics	0.005
Separate - physical sciences	0.439
Separate - engineering	<0.001
Separate - management	0.921
Separate - third specification	
Separate - IT and mathematics	0.004
Separate - physical sciences	0.115
Separate - engineering	<0.001
Separate - management	0.528

Source: Subsample of NSCG 2017. Author's computations.
Note: Pooled-all fields is specification (5) of the pooled model.

Comparisson of Pooled Model vs. Separate Models

We believe that separate models for each of the job categories jointly fit the data more than the pooled model (see the discussion in section 3.1). To test this, we use the likelihood-ratio test.

LR	DF	P(>Chisq)
1434.396	247	<0.0001

The results of the likelihood-ratio test showed that the separate models jointly fit the data better than the pooled model on the 0.0001 level of significance. Therefore to evaluate our hypotheses, we shall use the separate models. However, we will still provide the pooled model as well to ensure comparability of the estimates with other studies.

5.2 Pooled Model

In order to provide results that can be compared to previous studies (see Section 2.3), we first estimate the pooled OLS model using the whole sample for all job fields. The results of five different specifications of the pooled model are presented in Table 5.2. The pooled model displaying the effects of all variables is shown in Appendix A.1. We discuss the specifications of the model and its theoretical background in Section 3.1.

The results of the specification (1) of the pooled model imply that personal characteristics account for slightly less than a third of the 18.2% GWG with the resulting gender wage differential being 13.8%[1]. The addition of job-related variables in the specification (2) further reduces the gender wage differential to 3.7%. Both of those results are in line with the present research showing that nowadays, most of the GWG is explained by the differences in job allocation between men and women, while differences in personal characteristics account for only a small part of the GWG (e.g., Blau & Kahn 2017).

As was expected, the estimated relationship between *log(hours worked)* and *log(weekly wage)* is lower than 1, only around 0.7 in all specifications of our model, which could be interpreted that an employee receives only about 0.7% wage more for a 1% increase in his/her hours worked. Therefore we believe that it has been a correct approach to use the logarithm of weekly wage controlling for the logarithm of hours worked, rather than calculating hourly wages and thus assuming the relationship to be 1.

Being in a marriage-like relationship has a positive, statistically significant effect on wages in all specifications of the pooled model ranging from 6.1% in the specification (3) to 8.6% in the specification (1). Allowing the effect on wages to vary by gender in the specification (5) shows that married women earn 3% less than married men.

The size of the employer has a large and significant estimated effect on wages in all specifications of the model that include those variables (2-5). Individuals working for *mid employer* earn around 9.4% more than those working for *small employers*. Individuals working for *large employers* earn even more, around 17.3% more than those working for *small employers*. Allowing the effect of employer size to differ by gender in specifications (4) and (5) shows that women earn slightly less working for those employers than men (around 2%). However, those coefficients are not significant. We would generally expect that women face fewer hardships working for large employers and be remunerated more equally. Hence we would expect the differences between the estimated effects of *F*Mid employer* and *F*Large employer* to be more profound.

[1] We calculate the differences in percentages using the following formula: $\delta y = exp(\delta x \beta)$.

Table 5.2: Pooled model - all job fields

	Dependent variable:				
	log(Weekly wage)				
Specifications	(1)	(2)	(3)	(4)	(5)
Female	-0.149^{***}	-0.038^{***}	-0.070^{***}	-0.053^{*}	-0.039^{+}
	(0.008)	(0.008)	(0.017)	(0.022)	(0.023)
log(Hours worked)	0.758^{***}	0.680^{***}	0.685^{***}	0.685^{***}	0.683^{***}
	(0.024)	(0.022)	(0.022)	(0.022)	(0.022)
Marriage	0.083^{***}	0.060^{***}	0.059^{***}	0.060^{***}	0.074^{***}
	(0.009)	(0.008)	(0.008)	(0.008)	(0.010)
Mid employer		0.090^{***}	0.089^{***}	0.101^{***}	0.102^{***}
		(0.012)	(0.012)	(0.016)	(0.016)
Large employer		0.160^{***}	0.159^{***}	0.168^{***}	0.168^{***}
		(0.010)	(0.010)	(0.013)	(0.013)
IT and mathematics		0.289^{***}	0.269^{***}	0.269^{***}	0.268^{***}
		(0.014)	(0.018)	(0.018)	(0.018)
Physical sciences		-0.101^{***}	-0.161^{***}	-0.161^{***}	-0.161^{***}
		(0.023)	(0.031)	(0.031)	(0.031)
engineering		0.173^{***}	0.157^{***}	0.156^{***}	0.156^{***}
		(0.012)	(0.015)	(0.015)	(0.016)
management		0.233^{***}	0.208^{***}	0.208^{***}	0.207^{***}
		(0.019)	(0.023)	(0.023)	(0.023)
F*Marriage					-0.030^{*}
					(0.014)
F*Mid employer				-0.026	-0.025
				(0.023)	(0.023)
F*Large employer				-0.021	-0.020
				(0.020)	(0.020)
F*IT and mathematics			0.044	0.046	0.047
			(0.031)	(0.031)	(0.031)
F*Physical sciences			0.134^{**}	0.135^{**}	0.136^{**}
			(0.046)	(0.046)	(0.046)
F*Engineering			0.031	0.034	0.036
			(0.023)	(0.023)	(0.023)
F*Management			0.057	0.058	0.061
			(0.041)	(0.041)	(0.041)
Personal control vars.	Yes	Yes	Yes	Yes	Yes
Job control vars.	No	Yes	Yes	Yes	Yes
Detailed job cat.	No	No	No	No	No
Observations	10,548	10,548	10,548	10,548	10,548
DF	10,533	10,513	10,502	10,500	10,499
Adjusted R^2	0.289	0.451	0.453	0.453	0.454

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
See Section 3.1 for methodology and 4.2 for variables definition.

All of the interaction terms between *female* and male-dominated job-fields are positive, with only the effects of *F*Physical sciences* being statistically significant. However, to obtain the estimated gender wage differential for the particular job field, we ought to sum the estimated coefficient on *female* and the *female*job field interaction* (and possibly other interactions with *female* in other specifications). Therefore in the specification (3), the estimated gender wage differential is positive[2] for all male-dominated fields except physical sciences. The estimated gender wage differentials are then 2.5%, -6.6%, 4.0%, and 1.3% for IT and mathematics, physical sciences, engineering, and management, respectively. Those estimates of the gender wage differentials are then further differentiated in the specifications (4) and (5) via adding *female***employer size* and *female***marriage*.

5.3 Separate Models

Next, we estimate three specifications of separate models for each of the male-dominated fields (see Section 3.1 for details on the methodology). Estimating a separate model for each job field allows all the estimated effects to differ by job field, and allow us to conveniently include detailed job categories as explanatory variables. Furthermore, the way the models are specified allows us to test our hypotheses easily. The results of the models are specified in Tables 5.3, 5.4, 5.5. The full models can be found in Appendix A. Table 5.6 summarize the hypothesis testing for all hypotheses and all male-dominated job fields.

To evaluate our first hypothesis that in male-dominated job fields, women earn more than similar men, we estimate the first specification of the separate models. The results are presented in Table 5.3. In the model for physical sciences we had no observations on *not related other* and *professional*, and in the model for engineering we had no observation on *professional* (See Table A.2). This naturally applies to all specifications of the separate models.

Although the results show that women earn marginally more in all fields except management, the size of coefficients on *female* is negligible and insignificant. Therefore we cannot reject the null hypothesis that women do not earn more than comparable men for any of the male-dominated job fields. Further, we see that there are indeed vast differences in returns to hours worked across job fields. In IT and mathematics, the estimated coefficient is 0.932, while in engineering and physical sciences, it is 0.494 and 0.404, respectively, and in management the coefficient is even lower, just 0.237. The size of the coefficient in IT and mathematics which is close to 1 suggest that majority of employees there are paid hourly salary. The particularly small return to hours worked in management is in-line with our expectations because managers are typically remunerated for results of their subordinates and their working hours are not

[2] GWG, and gender wage differentials are reported as positive when women are estimated to earn less then men.

under such scrutiny compared to regular workers. In all job fields, the estimated effects of *marriage* are positive. The size of the employer (represented by *mid employer* and *large employer* dummies) is clearly positively related to earnings with the estimated effects being more profound in IT and mathematics and engineering.

Table 5.3: Separate models - first specification

	Dependent variable:			
	log(Weekly wage)			
Job fields	(1)	(2)	(3)	(4)
Female	0.011	0.027	0.004	−0.027
	(0.025)	(0.048)	(0.015)	(0.040)
log(Hours worked)	0.932^{***}	0.404^{***}	0.494^{***}	0.237^{**}
	(0.080)	(0.097)	(0.054)	(0.089)
Marriage	0.061^{**}	0.086^{+}	0.039^{**}	0.008
	(0.020)	(0.046)	(0.013)	(0.036)
Mid employer	0.090^{**}	0.056	0.071^{***}	0.023
	(0.031)	(0.084)	(0.020)	(0.049)
Large employer	0.177^{***}	0.098	0.144^{***}	0.098^{*}
	(0.027)	(0.083)	(0.017)	(0.046)
Control variables	Yes	Yes	Yes	Yes
Detailed job cat.	Yes	Yes	Yes	Yes
Observations	1,351	268	2,410	460
DF	1,313	240	2,363	430
Adjusted R^2	0.441	0.432	0.387	0.131

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
(1) IT and mathematics, (2) physical sciences, (3) engineering, (4) management
See Section 3.2 for methodology and 4.2 for variables definition.

To evaluate our second hypothesis that in male-dominated job fields, women working for large employers earn more than similar men, we estimate the second specification of the separate models. From the results presented in Table 5.4, we can calculate that women working for large employers (represented by *Female X Large*) earn 3.7%, 4.1%, and 1.2% more than men working for large employers in IT and mathematics, physical sciences and management, respectively. In engineering, they earned -1.6% less. As shown in Table 5.6, the second hypothesis involves around 60% of men and women in male-dominated fields (since around 60% of men and women work for large employers). However, none of the corresponding coefficients is statistically significant. Therefore, we cannot reject the null hypothesis that women working for large employers do not earn more than similar men for any of the male-dominated fields.

Table 5.4: Separate models - second specification

	Dependent variable:			
	log(Weekly wage)			
Job fields	(1)	(2)	(3)	(4)
log(Hours worked)	0.935^{***}	0.406^{***}	0.492^{***}	0.226^{*}
	(0.081)	(0.097)	(0.054)	(0.090)
Marriage	0.064^{**}	0.086^{+}	0.039^{**}	0.008
	(0.021)	(0.046)	(0.013)	(0.037)
Gender X Employer size				
Male X Small	-0.180^{***}	-0.094	-0.148^{***}	-0.067
	(0.029)	(0.080)	(0.018)	(0.044)
Female X Small	-0.133^{*}	-0.065	-0.148^{***}	-0.157
	(0.052)	(0.145)	(0.042)	(0.101)
Male X Mid	-0.057^{*}	-0.011	-0.094^{***}	-0.044
	(0.028)	(0.091)	(0.016)	(0.047)
Female X Mid	-0.135^{**}	-0.029	-0.008	-0.119^{*}
	(0.042)	(0.072)	(0.038)	(0.054)
Female X Large	0.036	0.040	-0.016	0.012
	(0.033)	(0.053)	(0.016)	(0.053)
Control variables	Yes	Yes	Yes	Yes
Detailed job cat.	Yes	Yes	Yes	Yes
Observations	1,351	268	2,410	460
DF	1,311	238	2,361	428
Adjusted R^2	0.442	0.428	0.389	0.130

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
(1) IT and mathematics, (2) physical sciences, (3) engineering, (4) management
See Section 3.2 for methodology and 4.2 for variables definition.

To evaluate our third hypothesis that in male-dominated job fields, single women working for large employers earn more than similar men, we estimate the third specification of the separate models. The results are presented in Table 5.5. We can calculate that compared to single men working for a large employer, single women working for large employers (represented by *Female X Large X Single*) earn 5.8%, 12.4%, and 9.4% more in IT and mathematics, physical sciences and management, respectively, and 2% less engineering. As shown in Table 5.6, the third hypothesis involves around 30% of women and men in male-dominated fields (since around 30% of employees work for large employers and are single). However, none of the corresponding coefficients is statistically significant at predefined 0.05 level (the corresponding coefficient for physical sciences is significant at 0.1 level). Therefore we cannot reject the null that single women working for large employers do not earn more than similar men for any of the male-dominated fields.

Table 5.5: Separate models - third specification

	Dependent variable:			
	log(Weekly wage)			
Job fields	(1)	(2)	(3)	(4)
log(Hours worked)	0.930^{***}	0.421^{***}	0.493^{***}	0.227^{*}
	(0.080)	(0.101)	(0.054)	(0.091)
Gender X Employer size X Married				
Male X Small X Married	-0.117^{**}	0.026	-0.112^{***}	-0.026
	(0.039)	(0.102)	(0.025)	(0.065)
Male X Small X Single	-0.173^{***}	-0.050	-0.150^{***}	-0.077
	(0.042)	(0.109)	(0.026)	(0.066)
Female X Small X Married	-0.070	0.012	-0.129^{**}	-0.055
	(0.075)	(0.105)	(0.049)	(0.257)
Female X Small X Single	-0.125^{*}	0.019	-0.130^{+}	-0.180^{+}
	(0.063)	(0.250)	(0.070)	(0.095)
Male X Mid X Married	0.041	0.165^{+}	-0.051^{*}	-0.028
	(0.044)	(0.093)	(0.026)	(0.069)
Male X Mid X Single	-0.075^{*}	0.001	-0.103^{***}	-0.024
	(0.037)	(0.125)	(0.022)	(0.069)
Female X Mid X Married	-0.095^{*}	0.044	0.030	-0.081
	(0.042)	(0.099)	(0.067)	(0.070)
Female X Mid X Single	-0.095	0.076	-0.011	-0.140^{+}
	(0.085)	(0.098)	(0.030)	(0.084)
Male X Large X Married	0.070^{**}	0.162^{*}	0.035^{*}	0.031
	(0.026)	(0.069)	(0.017)	(0.051)
Female X Large X Married	0.086^{+}	0.131^{+}	0.021	-0.015
	(0.050)	(0.072)	(0.021)	(0.060)
Female X Large X Single	0.056	0.117^{+}	-0.020	0.090
	(0.044)	(0.069)	(0.026)	(0.090)
Control variables	Yes	Yes	Yes	Yes
Detailed job cat.	Yes	Yes	Yes	Yes
Observations	1,351	268	2,410	460
DF	1,306	233	2,356	423
Adjusted R^2	0.441	0.424	0.388	0.128

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
(1) IT and mathematics, (2) physical sciences, (3) engineering, (4) management
See Section 3.2 for methodology and 4.2 for variables definition.

Table 5.6 summarizes the results of the separate models. For each male-dominated job field and for each hypothesis, it presents a corresponding estimate of the gender wage differential along with the confidence interval, p-value, and the percentage of women and men the hypothesis involves in the particular job-field. We were not able to reject the null hypothesis for any of the hypotheses for any of the job-fields. Had we settled for a 0.1 significance level, we would have rejected the null hypothesis of the third hypothesis for physical sciences, which would have implied that single women working for large employers in physical sciences earn more than similar men.

Table 5.6: Hypothesis testing summary

Job fields	(1)	(2)	(3)	(4)
Hypothesis 1				
Coefficient	0.011	0.027	0.004	-0.027
CI 95%	(-0,037 ; 0,059)	(-0,066 ; 0,121)	(-0,025 ; 0,034)	(-0,107 ; 0,052)
P-value	0.66	0.57	0.77	0.50
Female %/Male %	100%/100%	100%/100%	100%/100%	100%/100%
Hypothesis 2				
Coefficient	0.036	0.040	-0.016	0.012
CI 95%	(-0,028 ; 0,100)	(-0,064 ; 0,145)	(-0,048 ; 0,015)	(-0,092 ; 0,115)
P-value	0.27	0.45	0.31	0.82
Female %/Male %	58%/60%	58%/71%	72%/63%	62%/51%
Hypothesis 3				
Coefficient	0.056	0.117^{+}	-0.020	0.090
CI 95%	(-0,031 ; 0,143)	(-0,020 ; 0,254)	(-0,071 ; 0,032)	(-0,087 ; 0,266)
P-value	0.20	0.09	0.46	0.32
Female %/Male %	30%/30%	27%/37%	32%/29%	27%/22%

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
(1) IT and mathematics, (2) physical sciences, (3) engineering, (4) management
See Section 3.2 for methodology and Chapter 1 for hypotheses definition.

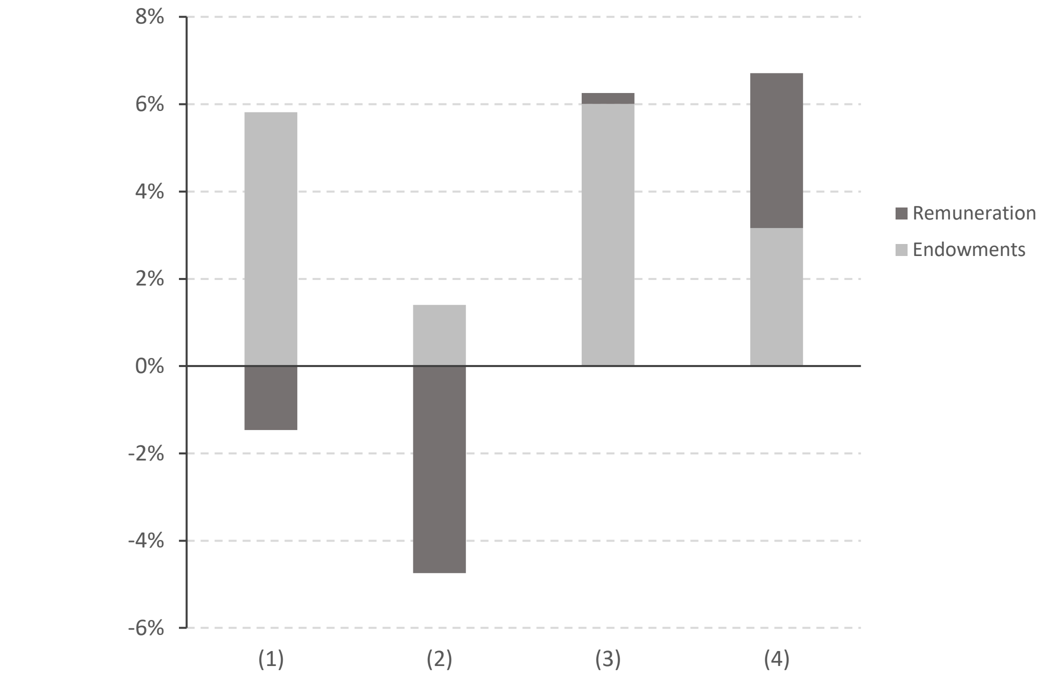
5.4 Oaxaca-Blinder Decomposition

To provide more insight into the composition of the within job GWG, we estimated how much of the within GWG is explained by remuneration effect and endowment effect using OB. The models for OB are the same as the first specification of the separate models (naturally without *female* dummy), except we had to eliminate a few covariates for each model for which one of the genders did not have any observation. The effect on sample size and within GWG was negligible, with the only exception being physical sciences where the GWG decreased from -2.7% to -3.3%.

The results are presented in Figure 5.1. The endowment effect is positive in all job categories, which means that in all of the categories, men had on average better characteristics. In IT and mathematics, the endowment effect explains more than 100% of the GWG (131%) because the remuneration effect is negative. In physical sciences, the overall GWG is negative, and therefore endowment effect only dampens the negative effect of remuneration. In engineering, the endowment effect accounts for more than 95% percent of the GWG, and in management, it accounts for 48% of the GWG.

In IT and mathematics and physical sciences, the remuneration effect was

Figure 5.1: Results of Oaxaca-Blinder decomposition for male-dominated job fields



Remuneration	-0.015	-0.047	0.003	0.035
	(0.028)	(0.053)	(0.016)	(0.051)
Endowments	0.058	0.014	0.060	0.032
	(0.023)	(0.053)	(0.013)	(0.038)
GWG	0.044	-0.033	0.063	0.067

Source: Sample of NSCG 2017. Author's computations.
Note: (1) IT and mathematics, (2) physical sciences, (3) engineering, (4) management

negative, which means that women were, on average, remunerated better for the same characteristics. In engineering and management, men were, on average, remunerated better. In management, the higher remuneration effect is partially caused by the low explanation power of the underlying model.

Chapter 6

Conclusion

In this thesis, we examined gender wage gaps (GWGs) among young college graduates in the US working in male-dominated job fields using data from 2017 National Survey of College Graduates (NSCG). Our aim was to asses whether adjusting for personal and job-related characteristics women earned more than men. In other words, we wanted to obtain estimates of gender wage differentials for male-dominated fields and test their statistical significance. For that we have formulated three hypotheses: 1) *In male-dominated job fields women earn more than similar men*, 2) *In male-dominated job fields, women working for large employer earn more than similar men,* and 3) *In male-dominated job fields, single women working for large employer earn more than similar men.*

Although the estimated coefficients of gender wage differentials would support hypotheses 1-3 for IT and mathematics and physical sciences, and hypotheses 2 and 3 for management, and therefore would suggest that in some cases women earn more than similar men, those coefficients are not statistically significant, and therefore we cannot reject the null hypothesis that women in male-dominated job fields do not earn more than similar men. We believe that the main limitation of our inference is the lack of variables that could act as proxies for employee's potential, e.g., academic performance or college ranking, as Leslie *et al.* (2017) have shown that the women earn more than similar men when their potential to reach managerial positions is high.

Comparing our results to results of Morgan (2008) suggests that in the US among young college graduates, in male-dominated fields, the gender wage differentials did not change much and women are not more/less favored nowadays compared to 20 years ago. However, we should be judicious in making this comparison because of methodological differences in our and Morgan's works.

Even though our thesis did not prove the existence of reverse gender wage differentials in male-dominated fields, our results also do not hint at any potential discrimination of women in those fields. This is also a valuable finding as it implies that contrary to the societal narrative, women can thrive in male-dominated fields.

Bibliography

AVERETT, S. L. & J. L. HOTCHKISS (1995): "The probability of receiving benefits at different hours of work." *The American economic review* **85(2)**: pp. 276–280.

BARDASI, E. & J. C. GORNICK (2008): "Working for less? women's part-time wage penalties across countries." *Feminist economics* **14(1)**: pp. 37–72.

BARTELS, L. K., J. T. NADLER, K. KUFAHL, & J. PYATT (2013): "Fifty years after the civil rights act: Diversity-management practices in the field." *Industrial and Organizational Psychology* **6(4)**: pp. 450–457.

BAUGHMAN, R., D. DINARDI, & D. HOLTZ-EAKIN (2003): "Productivity and wage effects of "family-friendly" fringe benefits." *International Journal of Manpower* **24(3)**: pp. 247–259.

BECKER, G. S. (1964): *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: National Bureau of Economic Research.

BECKER, G. S. (1971): *The Economics of Discrimination.* Chicago: University of Chicago Press.

BERTRAND, M., C. GOLDIN, & L. F. KATZ (2010): "Dynamics of the gender gap for young professionals in the financial and corporate sectors." *American economic journal: applied economics* **2(3)**: pp. 228–255.

BIELBY, W. T. & J. N. BARON (1986): "Men and women at work: Sex segregation and statistical discrimination." *American journal of sociology* **91(4)**: pp. 759–799.

BLACK, D. A., A. M. HAVILAND, S. G. SANDERS, & L. J. TAYLOR (2008): "Gender wage disparities among the highly educated." *Journal of human resources* **43(3)**: pp. 630–659.

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

BLINDER, A. S. (1973): "Wage discrimination: reduced form and structural estimates." *Journal of Human resources* **8(4)**: pp. 436–455.

BOOTH, A. & A. LEIGH (2010): "Do employers discriminate by gender? a field experiment in female-dominated occupations." *Economics Letters* **107(2)**: pp. 236–238.

BROWN, R. S., M. MOON, & B. S. ZOLOTH (1980): "Incorporating occupational attainment in studies of male-female earnings differentials." *Journal of Human Resources* **15(1)**: pp. 3–28.

BUDIG, M. J. & P. ENGLAND (2001): "The wage penalty for motherhood." *American sociological review* **66(2)**: pp. 204–225.

BUREAU OF LABOR STATISTICS (2018): *Women in the labor force: a databook.* U.S. Department of Labor.

BUSER, T., M. NIEDERLE, & H. OOSTERBEEK (2014): "Gender, competitiveness, and career choices." *The Quarterly Journal of Economics* **129(3)**: pp. 1409–1447.

CARLSSON, M. & S. ERIKSSON (2017): "The effect of age and gender on labor demand-evidence from a field experiment." *Econstor* .

CORRELL, S. J., S. BENARD, & I. PAIK (2007): "Getting a job: Is there a motherhood penalty?" *American journal of sociology* **112(5)**: pp. 1297–1338.

CROSON, R. & U. GNEEZY (2009): "Gender differences in preferences." *Journal of Economic literature* **47(2)**: pp. 448–74.

FEINGOLD, A. (1994): "Gender differences in personality: a meta-analysis." *Psychological bulletin* **116(3)**: p. 429.

GAYLE, G.-L., L. GOLAN, & R. A. MILLER (2012): "Gender differences in executive compensation and job mobility." *Journal of Labor Economics* **30(4)**: pp. 829–872.

GICHEVA, D. (2013): "Working long hours and early career outcomes in the high-end labor market." *Journal of Labor Economics* **31(4)**: pp. 785–824.

GREENWOOD, C., M. HARRISON, & A. VIGNOLES (2011): "The labour market value of stem qualifications and occupations." *London: Institute of Education* .

HILL, A. D., A. D. UPADHYAY, & R. I. BEEKUN (2015): "Do female and ethnically diverse executives endure inequity in the ceo position or do they benefit from their minority status? an empirical examination." *Strategic Management Journal* **36(8)**: pp. 1115–1134.

HLAVAC, M. (2013): "stargazer: Latex code and ascii text for well-formatted regression and summary statistics tables." *URL: http://CRAN. R-project. org/package= stargazer* .

HLAVAC, M. (2014): "oaxaca: Blinder-oaxaca decomposition in r." *Available at SSRN 2528391* .

HODGES, M. J. & M. J. BUDIG (2010): "Who gets the daddy bonus? organizational hegemonic masculinity and the impact of fatherhood on earnings." *Gender & Society* **24(6)**: pp. 717–745.

INTEL (2018): "Corporate responsibility at intel." .

JAMES DAMORE (2017): "Google's ideological echo chamber."

JONES, F. L. & J. KELLEY (1984): "Decomposing differences between groups: A cautionary note on measuring discrimination." *Sociological Methods & Research* **12(3)**: pp. 323–343.

KENA, G., L. MUSU-GILLETTE, J. ROBINSON, X. WANG, A. RATHBUN, J. ZHANG, S. WILKINSON-FLICKER, A. BARMER, & E. D. V. VELEZ (2015): "The condition of education 2015." *National Center for Education Statistics* .

LE BARBANCHON, T., R. RATHELOT, & A. ROULET (2019): "Gender differences in job search: Trading off commute against wage." *Available at SSRN 3467750* .

LEIBBRANDT, A. & J. A. LIST (2015): "Do women avoid salary negotiations? evidence from a large-scale natural field experiment." *Management Science* **61(9)**: pp. 2016–2024.

LESLIE, L. M., C. F. MANCHESTER, & P. C. DAHM (2017): "Why and when does the gender gap reverse? diversity goals and the pay premium for high potential women." *Academy of Management Journal* **60(2)**: pp. 402–432.

LOWEN, A. & P. SICILIAN (2009): "Family-friendly fringe benefits and the gender wage gap." *Journal of Labor Research* **30(2)**: pp. 101–119.

MALKIEL, B. G. & J. A. MALKIEL (1973): "Male-female pay differentials in professional employment." *The American Economic Review* **63(4)**: pp. 693–705.

MAUME JR, D. J. (1999): "Glass ceilings and glass escalators: Occupational segregation and race and sex differences in managerial promotions." *Work and occupations* **26(4)**: pp. 483–509.

MICHELMORE, K. & S. SASSLER (2016): "Explaining the gender wage gap in stem: Does field sex composition matter?" *RSF: The Russell Sage Foundation Journal of the Social Sciences* **2(4)**: pp. 194–215.

MINCER, J. (1974): *Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.* Education Resources Information Center.

MORGAN, L. A. (2008): "Major matters: A comparison of the within-major gender pay gap across college majors for early-career graduates." *Industrial Relations: A Journal of Economy and Society* **47(4)**: pp. 625–650.

MOSS-RACUSIN, C. A., J. F. DOVIDIO, V. L. BRESCOLL, M. J. GRAHAM, & J. HANDELSMAN (2012): "Science faculty's subtle gender biases favor male students." *Proceedings of the national academy of sciences* **109(41)**: pp. 16474–16479.

NEUMARK, D., R. J. BANK, & K. D. VAN NORT (1996): "Sex discrimination in restaurant hiring: An audit study." *The Quarterly journal of economics* **111(3)**: pp. 915–941.

NYHUS, E. K. & E. PONS (2005): "The effects of personality on earnings." *Journal of Economic Psychology* **26(3)**: pp. 363–384.

OAXACA, R. (1973): "Male-female wage differentials in urban labor markets." *International economic review* **14(3)**: pp. 693–709.

OECD (2019): *OECD Labour Force Statistics 2019.* OECD.

Patten, E. (2016): "Racial, gender wage gaps persist in us despite some progress." *Pew Research Center* **1**.

Petersen, T. & L. A. Morgan (1995): "Separate and unequal: Occupation-establishment sex segregation and the gender wage gap." *American Journal of Sociology* **101(2)**: pp. 329–365.

Petit, P. (2007): "The effects of age and family constraints on gender hiring discrimination: A field experiment in the french financial sector." *Labour Economics* **14(3)**: pp. 371–391.

Riach, P. A. & J. Rich (2006): "An experimental investigation of sexual discrimination in hiring in the english labor market." *Advances in Economic Analysis & Policy* **5(2)**.

Rosenfeld, R. A. (1983): "Sex segregation and sectors: An analysis of gender differences in returns from employer changes." *American Sociological Review* **48(5)**: pp. 637–655.

Russell, M. S., M. Krentz, K. Abouzahr, & M. Doyle (2019): "Women dominate health care—just not in the executive suite." BCG.

Sanborn, H. (1964): "Pay differences between men and women." *ILR Review* **17(4)**: pp. 534–550.

Stoet, G. & D. C. Geary (2018): "The gender-equality paradox in science, technology, engineering, and mathematics education." *Psychological science* **29(4)**: pp. 581–593.

Torpey, E. (2018): "Measuring the value of education. career outlook, us bureau of labor statistics." U.S. Bureau of Labor Statistics.

U.S. Census Bureau (2019a): "Curent population survey." U.S. Bureau of Labor Statistics.

U.S. Census Bureau (2019b): "Employed persons by detailed industry, sex, race, and hispanic or latino ethnicity." U.S. Bureau of Labor Statistics.

Weichselbaumer, D. & R. Winter-Ebmer (2005): "A meta-analysis of the international gender wage gap." *Journal of Economic Surveys* **19(3)**: pp. 479–511.

WEISBERG, Y. J., C. G. DEYOUNG, & J. B. HIRSH (2011): "Gender differences in personality across the ten aspects of the big five." *Frontiers in psychology* **2**: p. 178.

WILLIAMS, W. M. & S. J. CECI (2015): "National hiring experiments reveal 2: 1 faculty preference for women on stem tenure track." *Proceedings of the National Academy of Sciences* **112(17)**: pp. 5360–5365.

WOOLDRIDGE, J. M. (2016): *Introductory econometrics: A modern approach.* Mason: Nelson Education.

Appendix A

Appendix - Uncensored Models

Table A.1: Pooled model - all job fields [Full]

	Dependent variable:				
	log(Weekly Wage)				
Specifications	(1)	(2)	(3)	(4)	(5)
Female	-0.149^{***}	-0.038^{***}	-0.070^{***}	-0.053^{*}	-0.039^{+}
	(0.008)	(0.008)	(0.017)	(0.022)	(0.023)
log(Hours worked)	0.758^{***}	0.680^{***}	0.685^{***}	0.685^{***}	0.683^{***}
	(0.024)	(0.022)	(0.022)	(0.022)	(0.022)
Years from graduation	0.049^{***}	0.047^{***}	0.047^{***}	0.047^{***}	0.047^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Related somewhat		-0.074^{***}	-0.074^{***}	-0.075^{***}	-0.074^{***}
		(0.008)	(0.008)	(0.008)	(0.008)
Not related due pay		-0.168^{***}	-0.166^{***}	-0.166^{***}	-0.166^{***}
		(0.017)	(0.017)	(0.017)	(0.017)
Not related due comfort		-0.324^{***}	-0.321^{***}	-0.321^{***}	-0.321^{***}
		(0.022)	(0.022)	(0.022)	(0.022)
Not related other		-0.398^{***}	-0.396^{***}	-0.397^{***}	-0.396^{***}
		(0.034)	(0.034)	(0.034)	(0.034)
Master's	0.194^{***}	0.188^{***}	0.187^{***}	0.187^{***}	0.187^{***}
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
Ph. D.	0.304^{***}	0.392^{***}	0.397^{***}	0.397^{***}	0.396^{***}
	(0.026)	(0.024)	(0.024)	(0.024)	(0.024)
Professional	0.245^{***}	0.276^{***}	0.294^{***}	0.293^{***}	0.293^{***}
	(0.032)	(0.029)	(0.029)	(0.029)	(0.029)
Certificate or license	0.034^{***}	0.062^{***}	0.059^{***}	0.059^{***}	0.059^{***}
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Still enrolled at university	-0.222^{***}	-0.149^{***}	-0.150^{***}	-0.150^{***}	-0.150^{***}
	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)
Black	-0.102^{***}	-0.043^{**}	-0.042^{**}	-0.042^{**}	-0.043^{**}
	(0.016)	(0.014)	(0.014)	(0.014)	(0.014)
Asian	0.118^{***}	0.068^{***}	0.071^{***}	0.071^{***}	0.071^{***}
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)
Hispanic	-0.049^{***}	0.001	0.001	0.001	0.001
	(0.013)	(0.011)	(0.011)	(0.011)	(0.011)
Other ethnicity	-0.039^{*}	0.007	0.008	0.007	0.007
	(0.019)	(0.017)	(0.017)	(0.017)	(0.017)
Marriage	0.083^{***}	0.060^{***}	0.059^{***}	0.060^{***}	0.074^{***}
	(0.009)	(0.008)	(0.008)	(0.008)	(0.010)
Children	-0.006	0.004	0.005	0.005	0.004
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)
Mid employer		0.090^{***}	0.089^{***}	0.101^{***}	0.102^{***}
		(0.012)	(0.012)	(0.016)	(0.016)
Large employer		0.160^{***}	0.159^{***}	0.168^{***}	0.168^{***}
		(0.010)	(0.010)	(0.013)	(0.013)

Continues on the next page.

Table A.1: Results (cont.)

Government		−0.210***	−0.207***	−0.207***	−0.207***
		(0.011)	(0.011)	(0.011)	(0.011)
Non-profit		−0.166***	−0.164***	−0.163***	−0.164***
		(0.012)	(0.012)	(0.012)	(0.012)
Other employer		−0.105**	−0.106**	−0.107**	−0.107**
		(0.035)	(0.034)	(0.034)	(0.034)
IT and mathematics		0.289***	0.269***	0.269***	0.268***
		(0.014)	(0.018)	(0.018)	(0.018)
Biological and other life scientists		−0.213***	−0.229***	−0.230***	−0.230***
		(0.019)	(0.028)	(0.028)	(0.028)
Physical scientists		−0.101***	−0.161***	−0.161***	−0.161***
		(0.023)	(0.031)	(0.031)	(0.031)
Social and related scientists		0.034	0.104*	0.103*	0.103*
		(0.029)	(0.050)	(0.050)	(0.050)
Engineers and technicians		0.173***	0.157***	0.156***	0.156***
		(0.012)	(0.015)	(0.015)	(0.016)
Healthcare		0.057***	−0.077*	−0.078*	−0.078*
		(0.017)	(0.034)	(0.034)	(0.034)
Managers		0.233***	0.208***	0.208***	0.207***
		(0.019)	(0.023)	(0.023)	(0.023)
Teachers		−0.100***	−0.119***	−0.120***	−0.120***
		(0.017)	(0.028)	(0.028)	(0.028)
Social workers and councelors		−0.125***	−0.058	−0.059	−0.058
		(0.018)	(0.041)	(0.041)	(0.041)
Sales and marketing		−0.015	−0.013	−0.013	−0.013
		(0.022)	(0.033)	(0.032)	(0.032)
Financial specialists		0.131***	0.147***	0.146***	0.147***
		(0.022)	(0.033)	(0.033)	(0.033)
F*Marriage					−0.030*
					(0.014)
F*Mid employer				−0.026	−0.025
				(0.023)	(0.023)
F*Large employers				−0.021	−0.020
				(0.020)	(0.020)
F*IT and mathematics			0.044	0.046	0.047
			(0.031)	(0.031)	(0.031)
F*Biological and other life scientists			0.028	0.030	0.032
			(0.036)	(0.036)	(0.036)
F*Physical sciences			0.134**	0.135**	0.136**
			(0.046)	(0.046)	(0.046)
F*Social and related scientists			−0.097	−0.095	−0.093
			(0.060)	(0.060)	(0.060)
F*Engineers			0.031	0.034	0.036
			(0.023)	(0.023)	(0.023)
F*Healthcare			0.177***	0.179***	0.181***
			(0.038)	(0.039)	(0.039)
F*Managers			0.057	0.058	0.061
			(0.041)	(0.041)	(0.041)
F*Teachers			0.036	0.038	0.039
			(0.032)	(0.032)	(0.032)
F*Social workers and councelors			−0.069	−0.067	−0.068
			(0.044)	(0.045)	(0.045)
F*Sales and marketing			−0.005	−0.005	−0.005
			(0.044)	(0.044)	(0.044)
F*Financial specialists			−0.035	−0.034	−0.035
			(0.043)	(0.043)	(0.043)
Constant	3.960***	4.125***	4.127***	4.117***	4.118***
	(0.092)	(0.086)	(0.086)	(0.087)	(0.087)
Detailed job cat.	No	No	No	No	No
Observations	10,548	10,548	10,548	10,548	10,548
DF	10,533	10,513	10,502	10,500	10,499
Adjusted R^2	0.289	0.451	0.453	0.453	0.454

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
See Section 3.1 for methodology and 4.2 for variables definition.

Table A.2: Separate models - first specification [Full]

	Dependent variable:			
	log(Weekly Wage)			
Job fields	(1)	(2)	(3)	(4)
Female	0.011 (0.025)	0.027 (0.048)	0.004 (0.015)	−0.027 (0.040)
log(Hours worked)	0.932^{***} (0.080)	0.404^{***} (0.097)	0.494^{***} (0.054)	0.237^{**} (0.089)
Years from graduation	0.046^{***} (0.005)	0.046^{***} (0.012)	0.044^{***} (0.003)	0.040^{***} (0.007)
Related somewhat	-0.041^{+} (0.023)	−0.055 (0.052)	-0.029^{*} (0.013)	−0.014 (0.042)
Not related due pay	−0.045 (0.045)	0.174 (0.247)	-0.111^{**} (0.041)	−0.042 (0.062)
Not related due comfort	-0.341^{***} (0.070)	0.393^{***} (0.072)	-0.105^{*} (0.048)	-0.305^{***} (0.092)
Not related other	-0.185^{**} (0.069)		−0.123 (0.573)	-0.256^{***} (0.066)
Master's	0.169^{***} (0.022)	0.129^{*} (0.062)	0.118^{***} (0.014)	0.172^{***} (0.039)
Ph. D.	0.427^{***} (0.055)	0.521^{***} (0.100)	0.234^{***} (0.051)	0.264^{*} (0.113)
Professional	0.181^{*} (0.074)			0.358^{*} (0.178)
Certificate or license	−0.020 (0.027)	−0.034 (0.068)	0.039^{*} (0.016)	−0.055 (0.041)
Still enrolled at university	-0.160^{***} (0.043)	-0.239^{**} (0.075)	-0.132^{***} (0.026)	−0.003 (0.049)
Black	−0.025 (0.063)	−0.111 (0.091)	−0.036 (0.041)	−0.066 (0.057)
Asian	0.069^{**} (0.023)	−0.004 (0.075)	0.001 (0.016)	0.002 (0.047)
Hispanic	−0.020 (0.034)	−0.069 (0.049)	−0.005 (0.021)	−0.008 (0.048)
Other ethnicity	0.036 (0.045)	0.051 (0.114)	0.007 (0.029)	0.028 (0.079)
Marriage	0.061^{**} (0.020)	0.086^{+} (0.046)	0.039^{**} (0.013)	0.008 (0.036)
Children	0.013 (0.030)	0.146^{*} (0.069)	−0.003 (0.016)	0.024 (0.038)
Mid employer	0.090^{**} (0.031)	0.056 (0.084)	0.071^{***} (0.020)	0.023 (0.049)
Large employer	0.177^{***} (0.027)	0.098 (0.083)	0.144^{***} (0.017)	0.098^{*} (0.046)
Government	-0.209^{***} (0.033)	-0.218^{***} (0.056)	-0.188^{***} (0.022)	-0.132^{*} (0.057)
Non-profit	-0.259^{***} (0.042)	-0.250^{**} (0.087)	-0.185^{***} (0.046)	-0.146^{+} (0.077)
Other employer	-0.398^{**} (0.151)	0.064 (0.100)	−0.146 (0.097)	−0.080 (0.058)
Constant	3.351^{***} (0.308)	5.118^{***} (0.362)	5.147^{***} (0.208)	6.238^{***} (0.365)
Detailed job cat.	Yes	Yes	Yes	Yes
Observations	1,351	268	2,410	460
DF	1,313	240	2,363	430
Adjusted R^2	0.441	0.432	0.387	0.131

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
(1) IT and mathematics, (2) Physical sciences, (3) Engineers, (4) Managers
See Section 3.2 for methodology and 4.2 for variables definition.

Table A.3: Separate models - second specification [Full]

	Dependent variable:			
	log(Weekly Wage)			
Job fields	(1)	(2)	(3)	(4)
Employer size X Gender				
Small X Male	-0.180^{***}	-0.094	-0.148^{***}	-0.067
	(0.029)	(0.080)	(0.018)	(0.044)
Small X Female	-0.133^{*}	-0.065	-0.148^{***}	-0.157
	(0.052)	(0.145)	(0.042)	(0.101)
Mid X Male	-0.057^{*}	-0.011	-0.094^{***}	-0.044
	(0.028)	(0.091)	(0.016)	(0.047)
Mid X Female	-0.135^{**}	-0.029	-0.008	-0.119^{*}
	(0.042)	(0.072)	(0.038)	(0.054)
Large X Female	0.036	0.040	-0.016	0.012
	(0.033)	(0.053)	(0.016)	(0.053)
log(Hours worked)	0.935^{***}	0.406^{***}	0.492^{***}	0.226^{*}
	(0.081)	(0.097)	(0.054)	(0.090)
Years from graduation	0.045^{***}	0.047^{***}	0.044^{***}	0.040^{***}
	(0.005)	(0.012)	(0.003)	(0.007)
Related somewhat	-0.043^{+}	-0.054	-0.029^{*}	-0.013
	(0.023)	(0.052)	(0.013)	(0.041)
Not related due pay	-0.048	0.180	-0.110^{**}	-0.044
	(0.045)	(0.241)	(0.041)	(0.063)
Not related due comfort	-0.333^{***}	0.399^{***}	-0.111^{*}	-0.293^{**}
	(0.070)	(0.072)	(0.047)	(0.095)
Not related other	-0.186^{**}		-0.110	-0.246^{***}
	(0.070)		(0.574)	(0.071)
Master's	0.168^{***}	0.131^{*}	0.120^{***}	0.175^{***}
	(0.022)	(0.062)	(0.014)	(0.038)
Ph. D.	0.423^{***}	0.524^{***}	0.235^{***}	0.282^{*}
	(0.056)	(0.101)	(0.052)	(0.116)
Professional	0.192^{**}			0.375^{*}
	(0.074)			(0.178)
Certificate or license	-0.022	-0.038	0.040^{**}	-0.054
	(0.027)	(0.069)	(0.016)	(0.042)
Still enrolled at university	-0.161^{***}	-0.239^{**}	-0.132^{***}	-0.002
	(0.043)	(0.074)	(0.026)	(0.049)
Black	-0.030	-0.113	-0.036	-0.068
	(0.064)	(0.092)	(0.041)	(0.057)
Asian	0.067^{**}	0.002	0.001	0.005
	(0.022)	(0.074)	(0.016)	(0.047)
Hispanic	-0.023	-0.069	-0.007	-0.009
	(0.034)	(0.049)	(0.021)	(0.048)
Other ethnicity	0.038	0.055	0.007	0.030
	(0.045)	(0.113)	(0.029)	(0.078)
Marriage	0.064^{**}	0.086^{+}	0.039^{**}	0.008
	(0.021)	(0.046)	(0.013)	(0.037)
Children	0.010	0.147^{*}	-0.003	0.018
	(0.030)	(0.069)	(0.016)	(0.038)
Government	-0.211^{***}	-0.218^{***}	-0.188^{***}	-0.130^{*}
	(0.033)	(0.058)	(0.022)	(0.057)
Non-profit	-0.255^{***}	-0.248^{**}	-0.183^{***}	-0.140^{+}
	(0.041)	(0.089)	(0.046)	(0.078)
Other employer	-0.405^{**}	0.070	-0.146	-0.073
	(0.149)	(0.100)	(0.097)	(0.058)
Constant	3.513^{***}	5.196^{***}	5.300^{***}	6.360^{***}
	(0.310)	(0.372)	(0.208)	(0.369)
Detailed job cat.	Yes	Yes	Yes	Yes
Observations	1,351	268	2,410	460
DF	1,311	238	2,361	428
Adjusted R^2	0.442	0.428	0.389	0.130

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
(1) IT and mathematics, (2) Physical sciences, (3) Engineers, (4) Managers
See Section 3.2 for methodology and 4.2 for variables definition.

Table A.4: Separate models - third specification [Full]

	Dependent variable:			
	log(Weekly Wage)			
	(1)	(2)	(3)	(4)
Employer size X Gender X Marriage				
Small X Male X Married	−0.117**	0.026	−0.112***	−0.026
	(0.039)	(0.102)	(0.025)	(0.065)
Small X Male X Single	−0.173***	−0.050	−0.150***	−0.077
	(0.042)	(0.109)	(0.026)	(0.066)
Small X Female X Married	−0.070	0.012	−0.129**	−0.055
	(0.075)	(0.105)	(0.049)	(0.257)
Small X Female X Single	−0.125*	0.019	−0.130+	−0.180+
	(0.063)	(0.250)	(0.070)	(0.095)
Mid X Male X Married	0.041	0.165+	−0.051*	−0.028
	(0.044)	(0.093)	(0.026)	(0.069)
Mid X Male X Single	−0.075*	0.001	−0.103***	−0.024
	(0.037)	(0.125)	(0.022)	(0.069)
Mid X Female X Married	−0.095*	0.044	0.030	−0.081
	(0.042)	(0.099)	(0.067)	(0.070)
Mid X Female X Single	−0.095	0.076	−0.011	−0.140+
	(0.085)	(0.098)	(0.030)	(0.084)
Large X Male X Married	0.070**	0.162*	0.035*	0.031
	(0.026)	(0.069)	(0.017)	(0.051)
Large X Female X Married	0.086+	0.131+	0.021	−0.015
	(0.050)	(0.072)	(0.021)	(0.060)
Large X Female X Single	0.056	0.117+	−0.020	0.090
	(0.044)	(0.069)	(0.026)	(0.090)
log(Hours worked)	0.930***	0.421***	0.493***	0.227*
	(0.080)	(0.101)	(0.054)	(0.091)
Years from graduation	0.045***	0.046***	0.044***	0.041***
	(0.005)	(0.012)	(0.003)	(0.008)
Related somewhat	−0.042+	−0.053	−0.029*	−0.011
	(0.023)	(0.056)	(0.013)	(0.042)
Not related due pay	−0.050	0.169	−0.110**	−0.043
	(0.046)	(0.272)	(0.042)	(0.063)
Not related due comfort	−0.332***	0.392***	−0.110*	−0.291**
	(0.070)	(0.079)	(0.047)	(0.099)
Not related other	−0.186**		−0.106	−0.241*
	(0.071)		(0.576)	(0.096)
Master's	0.170***	0.139*	0.120***	0.178***
	(0.022)	(0.060)	(0.014)	(0.038)
Ph. D.	0.423***	0.508***	0.235***	0.276*
	(0.055)	(0.103)	(0.052)	(0.122)
Professional	0.191**			0.390*
	(0.074)			(0.184)
Certificate or license	−0.022	−0.047	0.040**	−0.055
	(0.027)	(0.072)	(0.016)	(0.039)
Still enrolled at university	−0.161***	−0.248**	−0.132***	0.006
	(0.043)	(0.076)	(0.026)	(0.047)
Black	−0.034	−0.126	−0.036	−0.071
	(0.063)	(0.104)	(0.041)	(0.056)
Asian	0.068**	0.009	0.0004	0.004
	(0.022)	(0.077)	(0.016)	(0.048)
Hispanic	−0.025	−0.074	−0.007	−0.003
	(0.034)	(0.055)	(0.021)	(0.048)
Other ethnicity	0.038	0.037	0.007	0.044
	(0.045)	(0.114)	(0.029)	(0.084)
Children	0.009	0.146*	−0.003	0.017
	(0.030)	(0.068)	(0.016)	(0.039)
Government	−0.212***	−0.212***	−0.188***	−0.122*
	(0.034)	(0.056)	(0.022)	(0.057)
Non-profit	−0.255***	−0.246**	−0.183***	−0.142+
	(0.042)	(0.090)	(0.046)	(0.075)
Other employer	−0.403**	0.094	−0.147	−0.087
	(0.150)	(0.100)	(0.097)	(0.059)
Constant	3.531***	5.104***	5.299***	6.338***
	(0.307)	(0.395)	(0.208)	(0.376)
Detailed job cat.	Yes	Yes	Yes	Yes
Observations	1,351	268	2,410	460
DF	1,306	233	2,356	423
Adjusted R^2	0.441	0.424	0.388	0.128

Source: Sample of NSCG 2017. Author's computations.
Notes: +p<0.1; *p<0.05; **p<0.01; ***p<0.001
(1) IT and mathematics, (2) Physical sciences, (3) Engineers, (4) Managers
See Section 3.2 for methodology and 4.2 for variables definition.