Educational Segregation in STEM/Non-STEM Fields and Wage Gender Gap: Evidence from the U.S.A.

Rui Yang

Abstract

This paper uses the 2018 – 2020 U.S. Census and American Community Survey microdata to examine the extent to which gender differences in graduates' choice of degree fields - STEM or Non-STEM contribute to the gender pay gap in the United States. With the Blinder-Oaxaca wage decomposition method, this paper finds that among the highly-educated labor force in the U.S. for the period 2018-2020, the individual background characteristics (race, region), family characteristics (marital status, number of children), education-related characteristics (degree fields, the highest educational attainment, and school type), and job-related characteristics (hours worked, work experience, job sector, and occupation type) combined can explain 55.96% of the gender pay gap. Among them, whether an individual chooses a college major related to a STEM field (ie. whether or not to obtain a degree in a STEM field) contributes 2.4%-5.2% of the "explanatory part", and can only explain 1.35%-2.91% of the gender pay gap. Finally, the results of this study showed that, although the content of the field of study (STEM/Non-STEM) seems not to explain too much of the current gender wage gap in the U.S. highly-educated workforce, it is hard to deny that educational segregation is still a barrier to gender equality. In addition, this paper also found that the gender pay gap for STEM graduates is 3.77% smaller than the gender pay gap for Non-STEM graduates. What's more, the choice of degree subjects can explain 9.91% of the gender gap for STEM graduate samples.

Key Words: Labor market, STEM, Education, Field of degree, Segregation, Blinder-Oaxaca decomposition, Gender wage gap

JEL classifications: I26 J16 J24 J31 J71 **Advisor name: Jenifer Ruiz-Valenzuela**

1. Introduction

Education not only promotes social and economic development and progress (e.g., Gradstein, and Justman, 2002; Wang, Xu, and Guo, 2018; De Meulemeester and Rocha, 1995; Dauda, 2013; Gylfason, 2001); it is also considered to be the path to success for disadvantaged groups (Bobbitt-Zeher, 2007). Although many studies have shown that women have made great strides in education (e.g., Heath and Jayachandran, 2016; Cavaglia et al., 2020). However, with the transformation of many countries' economic bases (from industry-based economies to information-based economies) and the rapid development of science and technology, the educational system and the focus of research fields in the educational system must also change and adjust in response to this change (e.g., Hamidi et al., 2011; Griffin et al., 2012). This is particularly the case for the gradual shifts of the educational focus toward STEM subjects (e.g., Ramaley, 2007). According to Cavaglia et al. (2020), although women's educational attainment has improved significantly and more women than men are in higher education, women are much less likely to choose Science, Technology, Engineering, and Mathematics (STEM) in high school or higher education. In addition, they also mentioned that research on gender disparities and related policy issues in STEM fields has important implications for the gender pay gap.

With the development of science and technology, the public is no longer unfamiliar with STEM fields anymore. As early as the 1990s, the National Science Foundation (NSF) first introduced the acronym "STEM" to combine Science, Technology, Engineering, and Mathematics together. The term is often used to address educational policy or curriculum choices in schools, but it also has important implications for workforce development, national security issues, and immigration policy. According to Kanematsu and Barry (2016), STEM is critical to modern education. From the national level, they argue that the hard work of experts in Science, Technology, Engineering, and Mathematics (STEM careers) is part of the reason that makes a country a world leader. In terms of the benefits gained from the STEM field at the individual level, the most intuitive is that those who work in the STEM field usually obtain higher wages (e.g., Rothwell, 2013; Fayer, 2017; Noonan, 2017; Bol, and Heisig, 2021). Although pursuing STEM fields promotes social development, the advancement of national leadership, and the increase in personal income, the lack of diversity in STEM fields is still a stumbling block for future progress that cannot be ignored. The lack of diversity in STEM fields is mainly reflected in insufficient racial diversity (e.g., Miriti, 2020; Morris and Washington, 2017) and insufficient gender diversity (e.g., McDonald, 2016; Wang and Degol, 2017; Benavent et al., 2020). More importantly, people of color and women are greatly underrepresented both in STEM subjects in academia and in STEM occupations in the labor market (e.g., Alegria and Branch, 2015; Scott and Elliott, 2019).

It is not difficult to find that STEM seems like a product derived from the development of science and technology with the times, and gender segregation in STEM subjects has gradually become a true portrayal of contemporary educational segregation. Moreover, a huge number of educational segregation studies on various countries have gradually begun to emerge into our field of vision (van Langen, 2015; Sahoo and Klasen, 2018; Livanos and Pouliakas, 2012; Oblova et al., 2020).

Based on the aforementioned educational segregation and the importance of STEM, and the fact that there is insufficient gender diversity in STEM academia and STEM fields in the labor market, this paper uses 2018-2020 U.S. Census and American Community Survey microdata to analyze gender differences in the degree field (STEM or Non-STEM) choice of graduates who are active in the labor market (and with full-time wage jobs). I further examine the extent to which the different field choices contribute to the gender pay gap in the United States through the Blinder-Oaxaca wage decomposition method. The results show that among the highly-educated full-time workforce in the U.S. labor market from 2018-2020, the individual background characteristics(race, region), family characteristics (marital status, number of children), education-related characteristics (degree fields, the highest educational attainment, and school type), and job-related characteristics (hours worked, work experience, job sector, and occupation type) combined can explain 55.96% of the gender pay gap. Although the content of the field of study (STEM / Non-STEM) seems not to explain too much of the current gender wage gap between the U.S. highly-educated workforce with the percentage of only 1.35% - 2.91%, it is still hard to deny that educational segregation is a barrier to gender equality.

The main contributions of this paper are shown as the following points: First, different from most research on the gender pay gap which has focused on only either educational segregation or occupational segregation (Bayard et al., 2003; Barón and Cobb-Clark, 2010;). This paper classifies both degrees and occupations by STEM fields and Non-STEM fields. By calculating and comparing the Educational Segregation Index and the Occupational Segregation Index, this paper found some differences between them and an interesting dynamic change with time in occupational segregation; Second, this paper not only limits the educational segregation and gender wage decomposition to the binary classification of degree fields (i.e., STEM and Non-STEM) but also further subdivides the subjects within STEM and Non-STEM separately, and found the fact that the gender pay gap for STEM graduates is 3.77% smaller than the gender pay gap for Non-STEM graduates; Third, despite the limitations of this paper, the study still can make some meaningful contributions to similar studies by explaining more than half of the gender pay gap for U.S. college graduates using the most recent data. Fourth, this study demonstrates through empirical analysis that the choice of research field (such as STEM or Non-STEM) seems to have a certain significant impact on income inequality. But the "gender composition" of the field with a much higher percentage of explaining the gender wage gap appears to be more related to the earnings inequality.

This paper consists of six parts which are organized as follows: The literature review is presented in Section 2; Section 3 introduces the data used in this study, which also includes a detailed descriptive statistical analysis of the data; Section 4 explains the methods and formulas used in this study; Section 5 presents the results of this study; The last section contains a brief conclusion and policy recommendations based on the research results. Besides, at the end of this section, I will also address the limitations of this paper and directions for future research.

2. Literature review

This paper connects to three strands of the literature. First, the paper is related to the literature on human capital and gender wage gaps over the world, and particularly, in the United States. The long-term trend of the gender pay gap is showing a narrowing trend in both of the United States and in other developed countries (e.g., Sicilian and Grossberg, 2001; Blau and Kahn, 2008; Kahn, 2015; Gharehgozli and Atal, 2020). It is noted that the role of human capital and education in influencing wages and narrowing the gender wage gap cannot be underestimated. According to Olson (2013), human capital models have made valuable contributions to the literature on the gender pay gap. Although, according to Blau and Kahn (2017), traditional human capital variables added together barely explain the gender pay gap in the united states from 1980 to 2010, and the more important factors that can explain the gender wage gap are mainly gender differences in occupations and industries. However, there is still a huge number literature that demonstrates the positive role of human capital in narrowing the gender pay gap. According to what was mentioned in Lim (2016), the observed gender pay gap among college graduates can be explained by gender differences in human capital characteristics such as college major choice, work experience, employment status, and cognitive skills. With the human capital model, Lim (2016) found that gender differences in observable (human capital) characteristics can explain more than half of the overall gender pay gap. Similarly, Livanos and Pouliakas (2012) found that through human capital models, up to 71% of wage differences could be "explained" by differences in demographic and job characteristics between the two genders, while an additional 8.4% of the gender pay gap can be attributed to the different choice of degree subjects by female and male.

Second, this study also relates to the extensive research literature on the lack of gender diversity in STEM fields. There is a lot of research that has discussed and examined the heterogeneity of gender differences across different subjects within STEM, and one of the most discussed STEM subjects is Mathematics (e.g., Benbow et al., 2000; Good et al., 2012; Keyserlingk et al., 2020). However, while most studies suggest that women are underrepresented in math-intensive fields, Ceci et al. (2009) argue that this is contradictory to claims of biological and sociocultural causality. In addition to Mathematics, there is extensive literature examining gender differences in Computer Science, Engineering, and Physics in STEM fields, in large part due to the widespread perception that women are severely underrepresented in these subjects (e.g., Cheryan et al., 2017; Brainard et al., 1998; Hill et al., 2010; Sattari et al., 2019). In this study, I found that highly-educated females accounted for only 36% of Mathematics and Statistics majors in STEM fields, and 30% of women in physical sciences. In addition, women make up even less in Computer and Information Sciences, about 23%, while women only make up 19% of the sample studying Engineering and Technologies. Therefore, the picture drawn from the data analysis in this paper is consistent with the facts described in the related literature.

Last but not least, the most important part of this study is the gender wage decomposition according to the different graduate degree fields (STEM fields or non-STEM fields). As mentioned earlier, many similar pieces of literature only focus on the study of a specific subject or occupation within the STEM field. For instance, Broyles (2009) studied and decomposed the gender wage gap for chemists in the United States, and ultimately found that 83% of the gender gap was caused by differences in production characteristics, while the remaining 17% was caused by discrimination or other unmeasured factors. Michelmore and Sassler (2016) extended the study to more broad STEM occupations and examined the gender wage gap by race among those working in Computer science, Life sciences, Physical sciences, and Engineering. They found that in fields where women are more represented (life sciences and physical sciences), the gender pay gap can be largely explained by differences in the characteristics observed between men and women working in those fields. However, in the fields with the lowest female concentrations (computer science and engineering), the gender pay gap persisted even after controlling for the observed characteristics.

The above literature mostly studies on occupations in STEM fields, and there are relative fewer paper on the gender wage gap in STEM subjects. Moreover, most studies that examine the wage gap for STEM college graduates but without decomposing wages. Differently, Lim (2016) examined the gender wage gap specific to STEM college graduates in the United States from 2008-2012 and also conducted a wage decomposition. Finally, Lim (2016) found that education-related experiences (i.e. college major, degree of college major and job relevance, and level of graduate degree earned) explained 42.8%-57.6% of the gender pay gap among STEM graduates. While the sample studied in Lim (2016) included only those workers with degrees in STEM fields, I studied a broader population that also included workers with degrees in Non-STEM fields. Also based on a human capital model, using similar variables as Lim (2016) and with the same wage decomposition method. In my study, all observable characteristics together explained 55.96% of the gender pay gap among U.S. highly-educated workers (including both who graduated in STEM fields and Non-STEM fields), with education-related variables explaining only about 20.62% - 20.7%. Variables related to degree fields play a smaller role in explaining the wage gap, accounting for only 1.35% - 2.91%.

However, to make up for the few empirical analyses that measure the contribution of female dominance in college subjects to the gender pay gap. Similar to Lim (2016), this study also includes a variable to measure the proportion of females in each subject. Although the representation of females in different subjects is not a statistically significant determinant for incomes in Lim (2016), this is an important variable in explaining the gender pay gap among all education-related variables in my research, which can explain 18.55% - 19.97% of the gender pay gap of American college graduates. In fact, this result is very close to Bobbitt-Zeher (2007). Similarly, in the research of Bobbitt-Zeher (2007), they found that the "gender dominance of the major" (ie, the percentage of females in the field of study) is one of the most significant differences between university majors, explaining about 14% of the overall gender pay gap. In general, gender segregation in education across degree fields or subjects is indeed a very tough issue for gender equality.

3. Data

This paper uses the 2018–2020 U.S. Census and American Community Survey microdata to examine the extent to which gender differences in graduates' choice of degree fields - STEM or Non-STEM contribute to the gender pay gap in the United States. Specifically, the data used in this paper are obtained from the IPUMS - USA, and each year includes a 1% sample of the U.S. population.

IPUMS is a very powerful micro-database that provides variables across multiple domains and a large sample of observations. It primarily focuses on collecting and storing U.S. Census microdata, including the decennial census from 1790 to 2010 and the American Community Survey (ACS) from 2000 to the present. In addition to basic demographic data such as age, race, gender, marital status, etc., the database also includes a series of economic data such as occupation, income, work status, education, etc. Furthermore, the ACS in 2009 began asking college graduates to report the major in which they earned their bachelor's degree.

This paper pooled three cross-sections from 2018 to 2020. A total of 1,708,359 individuals are in the age range of 25-65 and with higher education (a college degree or above). Among them, there are 779,330 males and 929,029 females. A total of 476,217 individuals aged 25-35 (211,801 males; 264,416 females), 420,772 individuals aged 36-45 (187,513 males; 233,259 females), 405,761 individuals aged 46-55 (male 187,202; female 218,559), and a total of 405,609 individuals aged 56-65 (male 192,814; female 212,795). This can reflect that more females than males are generally highly-educated across all four age groups in the United States. Besides, a total of 423,839 people graduated from STEM fields in all samples, accounting for only 24.81% of the sample, while the remaining 75.19% graduated with college degrees in Non-STEM fields. Far fewer people choose subjects in STEM fields than those who choose subjects in non-STEM fields[1].

According to Table 1, females are less likely than males to be in full-time employment and more likely to be part-time, regardless of whether they graduated from a STEM field or a Non-STEM field. In addition, females are less likely to be employed than males and less likely to participate in the labor market. However, it is not all appear negative for the performance in the labor market of females. Females who graduated in Non-STEM fields were less likely to be

1 It is important to note that some individuals have obtained two or more undergraduate degrees, but only different fields of the first degree were considered in this study. One of the main reasons is that the number of people who obtained two degrees accounted for 11.48% of the total sample, while the number of people who graduated from STEM fields accounted for only 1.43% of the total sample. Besides, to avoid confusion, especially for those with degrees in two STEM fields, this study chooses to consider only information from the first degree, which is also one of the limitations of this paper.

unemployed than males who also graduated in Non-STEM fields, at least according to the data used in this study.

Table 1. Comparison of employment status by gender among STEM/Non-STEM graduates, 2018- 2022, USA.

	STEM			Non-STEM		
Employment Status	Male (N=277219)	Female (N=146620)	Gender Gap (M-F)	Male (N=502111)	Female (N=782409)	Gender Gap (M-F)
Employed	89.03%	79.79%	9.25%	87.60%	79.19%	8.41%
Full-time job	76.27%	57.02%	19.26%	71.98%	52.90%	19.09%
Part-time job	12.76%	22.77%	-10.01%	15.61%	26.30%	-10.68%
Unemployment	2.08%	2.16%	-0.08%	2.50%	2.08%	0.41%
Out of labor force	8.88%	18.05%	-9.17%	9.91%	18.72%	-8.82%

Note:The raw data in the table are all from the IPUMS database.

Table 2 compared the different type of employment by gender among STEM and Non-STEM graduates. There are 156,039 self-employed workers in the sample, accounting for about 9.13% of the total sample, while there are 1,436,339 salaried workers, accounting for about 84.08% of the total sample. The remaining 115,981 individuals (6.79%) did not respond. As a result, only a minority of highly educated people in the United States are self-employed, and the vast majority are still salaried workers. Additionally, according to Georgellis et al. (2015), in countries such as the US and the UK, men are much more likely to be self-employed than women. Indeed, according to the data used in this paper, among those with higher education in the United States, 55.27% of the self-employed are men, and about 44.73% of women are self-employed. In addition, both in STEM and Non-STEM fields, males are more likely than females to be self-employed and to be hired for wages.

The main purpose of this paper is to study those full-time workers who earn wages (salaried workers). Therefore, this paper does not consider those individuals who are unemployed and those who are not active in the labor market and then filter out the self-employed and those who do not meet the definition of full-time. According to the definition of full-time full-year proposed by Winters (2014), this paper defines full-time as 40 or more hours worked per week and defines full-year as 50 or more weeks per year. It is also worth noting that some of the people in the sample are unpaid family workers, and most of them report an annual salary of 0, so this paper also deletes such individuals. Ultimately, the total sample size used in this study is of 987,036, of which 516,815 are male and 470,221 are female.

Table 2. Type of employment by gender among STEM/Non-STEM graduates, 2018-2022, USA.

Employment types	STEM			Non-STEM		
	Male (N=277219)	Female (N=146620)	Gender Gap (M-F)	Male (N=502111)	Female (N=782409)	Gender Gap (M-F)
Not observed	3.52%	9.31%	-5.79%	3.95%	9.30%	-5.35%
Self-employed	9.37%	6.91%	2.46%	12.00%	7.62%	4.38%
Work for wage	87.11%	83.78%	3.33%	84.05%	83.08%	0.97%

Note:The raw data in the table are all from the IPUMS database.

3.1 Main variables

The variables included in this study mainly fall into four categories:

(1) **Background variables** or **Demographic variables**: *sex* (male=0; female=1), *age* (age range from 25 to 65[2]), *race* (White people=0; Black people=1; American Indian or Alaska Native=2; Chinese & Japanese=3; Other Asian or Pacific Islander and other races=4; individual with two more major races=5), *region* (West Region=0; Northeast Region=1; Midwest Region=2; South Region=3).

(2) **Family variables**: marital status *married* (individual who is unmarried=0; the individual who is married=1; individual who is separated, divorced or widowed=2), number of children *nchild* (number from 0 to 9+).

(3) **Education-related variables**: school type *schltype* reports the school type of people who have enrolled in school in the last three months (individuals who have not enrolled in school in the last three months = 0; individuals who have enrolled in public schools in the last three months = 1; individuals who have enrolled in private schools in the last three months = 2), Educational attainment *educd* (Bachelor's degree=0; Master's degree=1; Professional degree beyond a bachelor’s degree=2; Doctoral degree=3), *degree2* (individual with no second bachelor degree=0; individual with second bachelor degree=1), Percentage of females in each

2 The upper age limit is based on Matthew S. Rutledge (2018). According to Matthew S. Rutledge (2018), the average retirement age in the United States has increased by about three years over the past three decades, reaching 64.6 for men and 62.3 for women. Therefore, this paper estimates that 65 years old is the retirement age.

subject *frate* (from 1% to 100%), and the categorical variables *stem*[3] (the bachelor degree of individual is not related to the STEM field=0; the bachelor degree of individual is in a STEM field=1) which cover the field information of the first bachelor degree only.

(4) **Job-related variables**: occupation variable *stem_occ*[4] (Individuals working in Non-STEM fields = 0, individuals working in STEM fields = 1), Employment sectors *sector* (Private sector=0; Non-profit sector=1; Government=2), Weeks worked last year *wkwork* (≥50 weeks per year), Usual hours worked per week *hwork* (≥40 hours per week), Working experience *wkexp*[5] (continuous variable: from 4 to 44), and each sample's yearly pre-tax wage and salary income *ywage*[6].

In addition to all those main variables mentioned above, this paper also subdivides STEM fields and Non-STEM fields in more detailed categories for degree fields, so that to further explore and compare the difference between STEM graduates and Non-STEM graduates. And also to find some important results that might be maskered within the STEM fields and Non-STEM fields[7].

3.2. Descriptive statistics

3.2.1. Means and gender differences of the main variables

Table 3 presents descriptive statistics for Income and Demographic characteristics for males and females. As expected, the average annual income of males is significantly higher than that of females. Besides, there are significantly more young females than males, which indicates that the average age of the male sample is likely to be higher than that of females so males might have longer work experience than females. And this is also an important factor when explaining the gender pay gap. Additionally, for both males and females, the proportion of highly-educated whites is far larger than other races. And it should be noted that the probability

3 STEM classifications for degree fields are based on the U.S. Department of Homeland Security (DHS) STEM Designated Degree Program List. Please check the Table A-1 in Appendix for the classification for STEM/Non-STEM degree fields.
4 The STEM classification of occupations is based on the document "STEM, STEM-related, and Non-STEM Occupation Code List 2010" which was updated and published by the U.S. Census Bureau on January 21, 2022. Please check the Table A-2 in Appendix for the classification for STEM/Non-STEM occupations.
5 Work experience is estimated based on a simple formula, explained in detail in Section 5.1.3.
6 All monetary units in this paper are US dollars ($).
7 Please check the Appendix for the complete classification of the subdivided degree fields by STEM/Non-STEM.

of the black female sample is significantly higher than that of males. With regard to this, we can refer to the huge number of previous literature on racial segregation that studied discrimination against blacks in the labor market (eg, Bergmann, 1974; Huffman, 2004). In Bergmann (1974) in particular, the authors mentioned two phenomena related to employment discrimination against blacks: First, the distribution of occupations among blacks is significantly different from that of whites, even after accounting for educational differences; Second, in occupations, whites earn more than blacks. Therefore, the differences in the racial distribution of males and females shown in Table 3 are also likely to be one of the important reasons for the gender wage gap between males and females. Finally, the higher education population is more distributed in the southern region of the United States, and there are also inevitably some gender differences between regions. According to Table 4, males are more likely to have more children than females, and males are more likely to be married than females, while females are more likely to be unmarried and single. Combined with the mean value of age shown in Table 3, it is not difficult to tell that this may be due to the fact that the males in the sample are generally older than the females.

Table 5 shows the mean of each Education-related variable, and more than half of both males and females have graduated from a bachelor's degree. Although males are significantly more likely to graduate with doctoral and professional degrees than females, females are significantly more likely to graduate with a master's degree than males. In addition, males are more likely to obtain a second bachelor's degrees and are more likely to major in a STEM field in their first degree. Last, the vast majority of males and females in the subsample have not attended school in the past three months, and females are more likely to enroll in school within the past three months maybe because the female populations are generally younger than males. What's more, In the sample of those who attended school in the past three months, both males and females were more likely to attend public schools than private schools.

Table 3. Means for Income and Demographic (Background) Variables, including Significance Tests of the Means for Females versus Males, 2018-2020, U.S.A.

Variables	Sample Mean (N=987,036)	Female (N=470,221)	Male (N=516,815)	t-Test of Mean Difference
Yearly wage	101042.70	82160.29	118222.8000	198.80***
Age Range				
age25-35	0.3065	0.3222	0.2921	-32.44***
age36-45	0.2686	0.2685	0.2688	0.34
age46-55	0.2486	0.2447	0.2522	8.61***
age56-65	0.1763	0.1646	0.1869	29.09***
Race				
White	0.7724	0.7583	0.7853	31.98***
Black	0.0645	0.0808	0.0497	-63.03***
Indian	0.0043	0.0053	0.0035	-13.38***
Asian	0.0321	0.0319	0.0323	1.08
Other	0.0890	0.0840	0.0935	16.54***
More-than-one	0.0377	0.0398	0.0358	-10.35***
Region				
West	0.2381	0.2285	0.2467	21.20***
Northeast	0.2032	0.2032	0.2032	-0.04
Midwest	0.2061	0.2083	0.2041	-5.16***
South	0.3527	0.3600	0.3460	-14.49***

Note: * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$ (two-tailed tests).

Table 4. Means for Family Variables, including Significance Tests of the Means for Females versus Males, 2018-2020, U.S.A.

Variables	Sample Mean (N=987,036)	Female (N=470,221)	Male (N=516,815)	t-Test of Mean Difference
Number of children				
No child	0.5149	0.5325	0.4990	-33.34***
1-2 Children	0.4011	0.4006	0.4015	0.91
3-4 Children	0.0789	0.0641	0.0924	52.30***
5-6 Children	0.0045	0.0026	0.0062	26.66***
7-8 Children	0.0005	0.0002	0.0008	12.92***
9+ Children	0.0001	0.0000	0.0002	6.06***
Marital status				
Unmarried	0.2252	0.2426	0.2093	-39.50***
Married	0.6711	0.6203	0.7173	103.06***
Single	0.1037	0.1372	0.0733	-100.00***

Note: * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$ (two-tailed tests).

Table 5. Means for Education-related Variables, including Significance Tests of the Means for Females versus Males, 2018-2020, U.S.A.

Variables	Sample Mean (N=987,036)	Female (N=470,221)	Male (N=516,815)	t-Test of Mean Difference
Female/Male ratio in different degree fields				
Female ratio	0.4760	0.5699	0.3906	-460.00***
Male ratio	0.5240	0.4301	0.6094	463.98***
Highest educational attainment				
Bachelor's degree	0.5992	0.5787	0.6178	39.64***
Master's degree	0.2899	0.3180	0.2642	-58.95***
Professional degree	0.0619	0.0588	0.0648	12.22***
Doctoral degree	0.0490	0.0444	0.0532	20.13***
Whether the individual has a second degree				
No	0.8841	0.8794	0.8883	13.68***
Yes	0.1159	0.1206	0.1117	-13.68***
Whether the individual university bachelor's degree belongs to the STEM field				
No	0.7222	0.8312	0.6231	-240.00***
Yes	0.2778	0.1688	0.3769	236.94***
Whether respondents attending school were enrolled in a public or a private school (any time in the past 3 months)				
Not enrolled	0.9450	0.9362	0.9530	36.59***
Public school	0.0348	0.0406	0.0295	-30.07***
Private school	0.0202	0.0232	0.0175	-20.10***

Note: * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$ (two-tailed tests).

Table 6 summarizes and compares the mean of different Job-related variables by gender. Similar to the mean value of males and females in STEM and Non-STEM degree fields, males are more likely than females to work in STEM fields. In general, most occupations in STEM fields have significantly higher average earnings than occupations in Non-STEM fields (e.g., Winters, 2014; Langdon et al., 2011). So this study would like to expect that this binary occupation variable can make a good contribution to explaining the current gender pay gap for highly-educated workers in the United States. Although both males and females are more likely to work in the private sector, males were shown to be greater. Furthermore, females are more likely than males to work in government and non-profit sectors. And according to Bender (1998), Lamo and Schuknecht (2012), Choudhury (1994), and many other related studies, the private sector usually has higher average income levels than the government sector and the non-profit sector. Therefore, this paper speculates that the existence of the gender wage gap might also be inevitably affected by this variable. Finally, males' average work experience and average weekly work hours are greater than females, although females' average workweeks per year are slightly greater than males.

Table 6. Means for Job-related Variables, including Significance Tests of the Means for Females versus Males, 2018-2020, U.S.A.

Variables	Sample Mean (N=987,036)	Female (N=470,221)	Male (N=516,815)	t-Test of Mean Difference
Whether the occupation of an individual belongs to the STEM field				
No	0.7308	0.7668	0.6981	-77.20***
Yes	0.2692	0.2332	0.3019	77.20***
Job sector				
Private	0.6213	0.5376	0.6974	165.73***
Government	0.2459	0.2924	0.2037	-100.00***
Non-profit	0.1328	0.1700	0.0989	-100.00***
Working experience	22.1890	21.7029	22.6313	41.47***
Usual hours worked per week	44.8449	43.9455	45.6633	110.32***
Weeks worked last year	51.6415	51.6437	51.6395	-4.27***

Note: * p < 0.05 ** p < 0.01 *** p < 0.001 (two-tailed tests).

3.2.2. Gender wage gap

- Average hourly wage by gender over time

Table 7 shows the average hourly wages by sex and wage ratios (women's wages/men's wages) for each year from 2018 to 2020 and the three-year period 2018-2020. It can be observed that the hourly wages of males and females have both increased slightly with a growth rate between 1.61% to 4.19%[8] over time, but the wage ratio still has not changed significantly. Although females' wages in 2020 increased by 3.77% compared with 2019, such an increase is still insufficient in reducing the gender wage gap.

8 The formula for calculating the wage growth rate is: (average hourly wage of the current year - average hourly wage of last year)/average hourly wage of last year. The results of the calculation show that the smallest increase was in men's average hourly wages in 2019, which increased by only 1.61% compared to 2018; the largest increase was in women's average hourly wages in 2020 (4.19%).

Table 7: Average hourly wages by gender and year, 2018-2020, U.S.A.

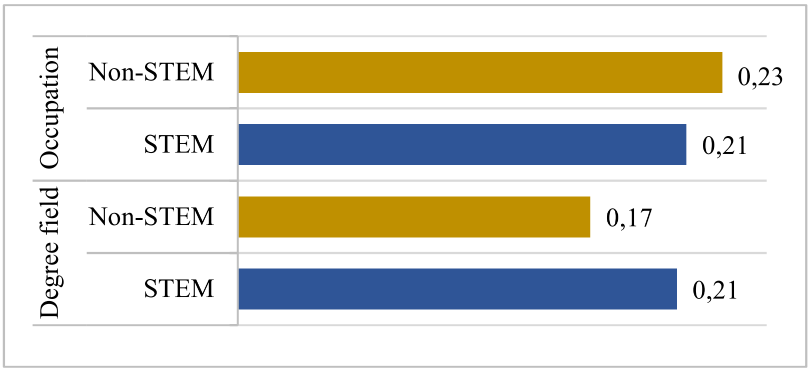
Year	All ($)	Male ($)	Female ($)	Wage ratio
2018	42.40	48.92	35.19	0.72
2019	43.08	49.70	35.81	0.72
2020	44.49	51.08	37.31	0.73
2018-2020	43.26	49.83	36.04	0.72

Note: The data all comes from the IPUMS database.

- Gender wage gap in STEM / Non-STEM degree fields

Referring to the OECD's definition of the gender pay gap, this paper also defines the gender pay gap as the difference between the median earnings of males and females relative to the median earnings of males. Figure 1 reports the gender pay gap for graduates in STEM degree fields and Non-STEM degree fields, respectively, as well as the gender pay gap for those who are employed in STEM occupations and Non-STEM occupations.

Figure 1: Gender wage gap (median hourly wage) in STEM/non-STEM fields, 2018-2020, U.S.A.



Note: The data all comes from the IPUMS database

According to Figure 1, gender pay gaps in degree fields and in occupations are both similar and different in STEM and Non-STEM fields. Similarly, the gender gap in STEM careers as well as STEM degrees both show a value of 0.21. However, what is very different is that the gender pay gap between workers in STEM fields is smaller than that of non-STEM workers. While the gender pay gap between graduates in STEM fields is larger than that of graduates in Non-STEM fields. This result is calculated from the (hourly) median wage, however, almost all the data analysis in this paper and the follow-up research are based on the analysis of the (hourly) mean wage, so I also provide the gender pay gap calculated by average (mean) hourly wage in the Appendix. Different from what has been concluded above, the final results (based on average wage calculation) show that the gender wage gap in Non-STEM fields is slightly larger than the gender gap in STEM fields, both in degree field choice and occupation.

3.2.3. Gender segregation and Duncan segregation index.

When it comes to measuring segregation, the most commonly used methodological analysis of segregation is the Duncan segregation index. It is considered the best measure of occupational segregation and the easiest measure of segregation to calculate. Therefore, this index is mostly used in the literature to measure occupational segregation between males and females. Differently, this paper focuses on both of exploring educational segregation and occupational segregation instead of paying attention only to studying occupational segregation as most previous studies did. Moreover, measuring educational segregation between males and females can help us obtain a general understanding of the current situation of educational segregation in the U.S. labor market, before moving on to measuring the gender wage gap later on.

The Duncan Segregation index is calculated using the following formula:

$$\mathrm{D} = \frac{1}{2}\sum_{s=1}^{N} \left| \frac{f_s}{F} - \frac{m_s}{M} \right| \qquad (1)$$

where "s" is the field of the bachelor's degree. In this paper "s" is either a STEM field or a non-STEM field. f_s is the female population with a college degree in either STEM fields or non-STEM fields; m_s is the male population with a college degree in either STEM or non-STEM; And F is the total population of females in the sample while M is the total number of males in the sample. The Duncan segregation index is mainly used to calculate whether there is a larger than expected presence of one gender over another in a given degree field by identifying the percentage of females (or males) who would have to change the bachelor's degree field for the distribution of males and females to be equal in this study.

Table 8 shows the calculated Duncan segregation index[9] for each year from 2018 to 2020. In addition to the educational segregation index shown in the second column, this paper also calculates the occupational segregation index according to the STEM classification. The U.S. educational segregation index classified by STEM and Non-STEM degree types remained

9 The calculation process and results are shown in the Appendix.

stable at 0.21 between 2018 and 2020, which means 21% of females need to change degree fields to equate the degree field distributions of males. When occupations are also classified into STEM types and Non-STEM types, the calculated segregation index is relatively small, which is about 0.07. In summary, two conclusions can be drawn from Table 8. First, the educational segregation index has been stable at 0.21 from 2018 to 2020, while the occupational segregation index has been increasing steadily and slightly by 0.01 per year. Second, the educational segregation index according to STEM classification is significantly larger than the occupational segregation index also calculated according to STEM classification. The reasons for this difference can be analyzed combined with Figure 2. Figure 2 shows the matching results of degrees and occupations in STEM fields and Non-STEM fields, and shown as percentages by gender. It can be observed that compared to males, females are more likely to graduate from Non-STEM fields and finally work in STEM fields. While males are more likely to graduate in a STEM field but work in a Non-STEM field.

Table 8: Educational Segregation and Occupational Segregation in the U.S.A., 2018-2020

Year	Fields of degree	Occupation
2018	0.21	0.06
2019	0.21	0.07
2020	0.21	0.08
2018-2020	0.21	0.07

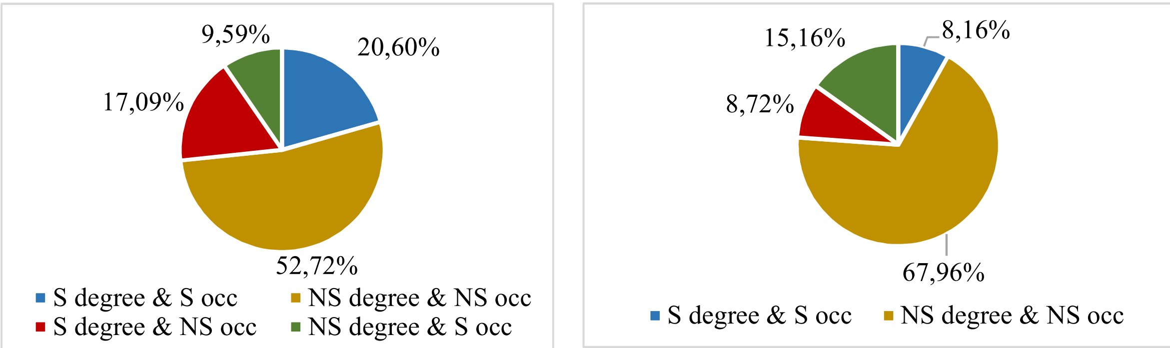
Note: The data all comes from the IPUMS database.

Based on consideration of the underrepresentation of females in certain subjects in STEM fields observed in existing research, this paper provides a more detailed classification of STEM fields as well as Non-STEM fields. Among them, there are 15 subjects in STEM fields, and 28 subjects in Non-STEM fields[10]. Figure 3 shows gender segregation across subjects in STEM

10 Two types of classifications were adopted for STEM fields in this study. The first classification comes directly from the IPUMS variable "DEGFIELD", which reports the general field in which the person received a Bachelor's degree if the person holds a Bachelor's degree. Another classification is based on Lim (2016) with some adjustments, and finally divides the STEM field disciplines into eight major categories:(1) Computer and Information Sciences; (2) Engineering; (3) Biology and Life Sciences; (4) Mathematics and Statistics; (5) Technologies; (6) Physical Sciences; (7)Medical and Health Sciences and Services; (8) Other STEM subjects. Since the classification of Non-STEM is derived only from the IPUMS variable "DEGFIELD", a consistent classification criterion (ie. "DEGFIELD") is adopted in this paper when calculating the gender segregation within the respective fields of STEM and Non-STEM.

fields and gender segregation across subjects in Non-STEM fields in the United States in 2018-2020, respectively. The results show that the gender segregation index is greater in STEM fields. That is, compared to the percentage of females who need to change fields of degree in Non-STEM fields, there are more than 4% of females who need to change fields of the degree to equate the degree fields distributions between males and females in STEM fields. This suggests that the unbalance in gender distribution across subjects in STEM fields is greater than that in Non-STEM fields.

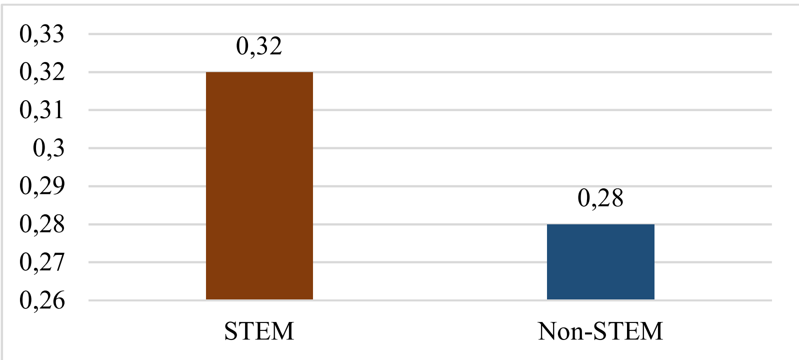
Figure 2: Matching of degrees and occupations in STEM fields and Non-STEM fields by gender, 2018-2020, U.S.A.



a. Male b. Females

Note: "S degree (occ)" represents the STEM field degree (occupation); "NS degree (occ)" represents the Non-STEM field degree (occupation).

Figure 3: Comparison of educational Segregation between STEM and Non-STEM degree field, 2018-2020, U.S.A.



Note: The data all comes from the IPUMS database.

4. Methodology

4.1. Introduction of the Blinder-Oaxaca decomposition method

The wage decomposition method has been greatly expanded and developed so far, and it can be divided into two different branches: the Mean Decomposition Method and the Distribution Decomposition Method. The Mean Decomposition Methods generally refer to the Blinder-Oaxaca decomposition and its improvement (Oaxaca, 1973; Blinder,1973), Brown decomposition and its improvement (Brown et al. 1980; Brown, 1993; O'Bryant, 2003), and JMP119 decomposition (Juhn et al., 1991). Among them, the Blinder-Oaxaca decomposition is the most basic and classic method. Blinder-Oaxaca decomposition refers to a method of grouping by gender, race, etc. and decomposing the average difference in logarithmic wages in a counterfactual way based on a linear regression model to finally calculate the portion of wage differences attributable to differences in skills and those that could be attributable to discrimination for different groups (e.g., males and females). Specifically, it divides the wage difference between two groups (such as males and females) into two parts, the first part being "explained" by group differences in productivity characteristics, such as education or work experience, while the second part, also is considered as the "unexplained" part which cannot be explained by differences in observable wage determinants and usually used to measure the discrimination between groups. In addition to this, the "unexplained" part also incorporates the effect of group differences in unobserved predictors.

4.2. Equations and the decomposition process

According to Oaxaca (1973) and Blinder (1973), if the wages of male groups *m* and female groups *f* in the labor market are written as w_m and w_f, respectively, the individual eigenvectors (or individual endowment) of these two groups as subsamples are $\boldsymbol{X}_m$ and $\boldsymbol{X}_f$ respectively. The corresponding regression coefficient vectors (or salary structure) are shown as $\boldsymbol{\beta}_m$ and $\boldsymbol{\beta}_f$, respectively. The semi-logarithmic wage estimation equations for these two groups (usually based on the Mincer wage determination equation) are:

$$lnw_m = \boldsymbol{X}_m \boldsymbol{\beta}_m + u_m \quad (2)$$

$$lnw_f = \boldsymbol{X}_f\, \boldsymbol{\beta}_f + u_f \qquad (3)$$

The means of the subsample individual eigenvectors for these two groups are $\overline{\boldsymbol{X}}_m$ and $\overline{\boldsymbol{X}}_f$, respectively. Then, according to the property that the mean of the least squares (OLS) residuals is zero, the difference between the mean wages of the two groups can be expressed as:

$$ln\overline{w}_m - ln\overline{w}_f = \overline{\boldsymbol{X}}_m\, \boldsymbol{\beta}_m - \overline{\boldsymbol{X}}_f\, \boldsymbol{\beta}_f \qquad (4)$$

Situation 1:

When the real wage structure of male group *m* is regarded as the non-discriminatory labor market wage structure, the Equation (4) can be decomposed into:

$$ln\overline{w}_m - ln\overline{w}_f = (\overline{\boldsymbol{X}}_m - \overline{\boldsymbol{X}}_f)\, \boldsymbol{\beta}_m + \overline{\boldsymbol{X}}_f\, (\boldsymbol{\beta}_m - \boldsymbol{\beta}_f) \qquad (4-1)$$

Situation 2:

When the real wage structure of female group *f* is regarded as the non-discriminatory labor market wage structure, the Equation (4) can be decomposed into:

$$ln\overline{w}_m - ln\overline{w}_f = (\overline{\boldsymbol{X}}_m - \overline{\boldsymbol{X}}_f)\, \boldsymbol{\beta}_f + \overline{\boldsymbol{X}}_m\, (\boldsymbol{\beta}_m - \boldsymbol{\beta}_f) \qquad (4-2)$$

Whether it is Equation (4-1) or Equation (4-2), the first term on the right side of the equation "$(\overline{\boldsymbol{X}}_m - \overline{\boldsymbol{X}}_f)\, \boldsymbol{\beta}_m$" or "$(\overline{\boldsymbol{X}}_m - \overline{\boldsymbol{X}}_f)\, \boldsymbol{\beta}_f$" represents the wage difference that exists between male group *m* and female group *f* even if there is no discrimination, that is, the difference in wages caused by the difference in individual characteristics between the male group *m* and the female group *f*; The second term "$\overline{\boldsymbol{X}}_f\, (\boldsymbol{\beta}_m - \boldsymbol{\beta}_f)$" or "$\overline{\boldsymbol{X}}_m\, (\boldsymbol{\beta}_m - \boldsymbol{\beta}_f)$" is the wage difference caused by the difference in wage structure between male group *m* and female group *f*. In other words, it is the difference between the wage difference between the male group *m* and the female group *f* in the presence of discrimination and the absence of discrimination, which Oaxaca calls discrimination.

4.3. Potential problem of Blinder-Oaxaca decomposition and challenges of this paper

Apparently, Oaxaca attributes all the "unexplained" part of wage differences to discrimination and this may lead to a series of potential problems when we use this method to decompose wages. There are three main types of problems in the Blinder-Oaxaca decomposition process. The first problem is the "index number problem", the second problem is related to "sample selection problems", and the third problem is the "dummy variable coefficient identification problems" (Jann, 2008).

(1) The so-called ***index benchmark problem*** refers to such a problem: In the process of mean wage difference decomposition, different decomposition results are obtained due to the different decomposition benchmarks selected. It is impossible to objectively and accurately measure the degree of influence of various factors or characteristics on wage differences, nor to uniquely infer the true degree of discrimination.

(2) The ***sample selection problem***, also known as sample selection bias, mainly refers to the fact that the sample selected in the study is not completely random. That is the multiple possibilities of decomposition result due to a certain degree of subjectivity in the attribution of sample selection bias correction term differences.

(3) The ***dummy variable coefficient identification problem*** means that when decomposing wage differences into the effects of every single covariate, the characteristic effect and coefficient effect of each dummy variable will change due to the choice of the benchmark group, resulting in the ambiguity of the decomposition results.

Since this study mainly focuses on the "explained" part of the wage decomposition results, and the “dummy variable coefficient identification problem” mainly affects the "unexplained" part. Thus, the main challenge of this study is how to address the first and second questions: that is, the “index benchmark problem” and the “sample selection bias problem”. For the first problem, since the results of the Blinder-Oaxaca decomposition completely depend on which group of regression coefficients are used as a reference, we need to carefully consider which group of regression coefficients to use as a reference before delving into the results of the Blinder-Oaxaca decomposition analysis. In order to avoid this ambiguity, this study chooses to perform the wage decomposition with the group indicator (female) in the pooled regression model as suggested by Jann (2008) and Elder et al. (2010) for wage decomposition.

In addition, this paper studies the highly-educated population in the US labor market, that is, those with a bachelor's degree or higher. In addition, the subsample used in this study restricts the population to those full-time workers who are active in the labor market and are not self-employed. According to descriptive statistical analysis, women are more likely than men to work part-time rather than full-time, and men are more likely than women to work as self-employed rather than working for wages. Thus, this paper great likely excludes a disproportionately large sample of part-time women and self-employed men and the potential existence of sample selection bias is likely to affect the results of the study. However, due to the limitation of data and many problems encountered in finding an effective instrumental variable, it is hard for this study to correct the selection bias problem according to the Heckman model (Heckman two-step method) effectively. I will explain this problem again in the limitations at the end of the paper.

5. Results

5.1. The semi-logarithmic income equation regression results

In addition to focusing on the binary degree field: the gender gap between STEM and Non-STEM, I am also concerned about the gender gap across subjects within STEM fields. Therefore, in this study, a binary dummy variable and a STEM field categorical variable (including eight different categories of STEM subjects) are respectively included in the semi-logarithmic income equation for regression, and the omitted variable for two types of variables are all the "Non-STEM degree fields" variable. Table 10 presents the regression results of the eight STEM subjects variable and the binary subjects variable are jointly reported and displayed. Both Tables 9 and 12, even all other variables except the "degree field variable" in Table 10 all show the part of the regression results when the binary degree field variable is incorporated into the equation regression.

5.1.1. The influence of Background and Family factors on average annual income

Table 9 shows Demographic and Family characteristics affect annual wages for males and females, respectively, and gender differences in these characteristics. The annual income of males gradually increases with age, and the group of male samples with the most income advantage is the oldest group, while the most significant income advantage of females is the youngest group (age range: 25-36).

Table 9: Regression results of income equation by gender (Demographic characteristics and Family characteristics), 2018-2020, U.S.A.

Variables	Male (N=516,815)	Female (N=470,221)
	Background and demographic characteristics	
Age Range		
age36-45	0.019*** (0.004)	0.000 (0.004)
age46-55	0.030*** (0.007)	-0.042*** (0.006)
age56-65	0.050*** (0.009)	-0.022** (0.009)
Race		
Black	-0.231*** (0.004)	-0.095*** (0.003)
Indian	-0.254*** (0.015)	-0.179*** (0.011)
Asian	-0.020*** (0.005)	0.052*** (0.005)
Other	-0.120*** (0.003)	-0.077*** (0.003)
More-than-one	-0.066*** (0.005)	-0.028*** (0.004)
Region		
Northeast	0.006** (0.003)	0.009*** (0.003)
Midwest	-0.155*** (0.003)	-0.146*** (0.003)
South	-0.106*** (0.002)	-0.141*** (0.002)
	Family formation characteristics	
Number of children		
1-2 Children	0.064*** (0.002)	0.018*** (0.002)
3-4 Children	0.092*** (0.003)	-0.009** (0.003)
5-6 Children	-0.004 (0.011)	-0.078*** (0.016)
7-8 Children	-0.036 (0.031)	-0.226*** (0.058)
9+ Children	0.009 (0.066)	-0.205* (0.124)
Marital status		
Married	0.197*** (0.003)	0.082*** (0.002)
Single	0.034*** (0.004)	-0.011*** (0.003)

Note: Significant at: *p , 0.1, * *p , 0.05 and ***p , 0.01; robust standard errors are in parentheses. For the complete regression table, please check the Appendix.

Almost all races have significantly lower wages than whites. The only exception is the Asian females, whose annual wage is significantly greater than that of white females. In addition, the differences between regions are also significant. The annual wage of males and females in the Northeast of the United States is higher than that of other regions. For the Family characteristics, both married males and females all have greater annual wages than unmarried males and females, but being single (divorced, separated, or widowed) only seems to have a negative effect on females' annual wages. Moreover, it can be found that the number of children does not negatively affect the annual wage of fathers and that females with 1-2 children are also not penalized in their wages. However, a significant "child penalty" as the number of children continued to increase is shown in the mother’s annual wages.

5.1.2. The influence of Education factors on average annual income

According to the regression results in Table 10, compared with those who graduated from Non-STEM subjects, those who graduated from STEM subjects have higher incomes for both males and females. However, although graduates from STEM subjects have higher incomes, there are also distinct differences in the more detailed classification of STEM subjects. Compared to the wages of those Non-STEM fields graduates, males and females graduate in the STEM fields of "Computer and Information Sciences", "Engineering", "Mathematics and Statistics", and "Medical and Health Sciences and Services" all earned significantly more. However, males who graduated from "Technologies" and "Other STEM subjects" earned less than Non-STEM male graduates. At the same time, in addition to the above two subjects, females who graduated in "Biology and Life Sciences" and "Physical Sciences" also earn less than Non-STEM female graduates.

According to Lim (2016), female-dominated subjects generally correspond to lower wages, which can be considered to be related to the “societal devaluation of feminine work”, which generally refers to the fact that occupations with high female representation are often associated with lower status and rewards. In this study, it seems to be that the female societal devaluation also can be observed in the degree field as well. It is therefore meaningful to analyze the impact of representation or proportion of females in different subjects on male wages and female wages and to measure their contribution to the gender wage gap. According to the regression results of the variable "*Frate*" in Table 10, the higher the proportion of females in various subjects, the more male and female wages will be punished. And females' wages are punished even more.

Table 10: Regression results of income equation by gender (Education-related characteristics), 2018-2020, U.S.A.

Variables	Male (N=516,815)	Female (N=470,221)
Education-related characteristics		
Whether the individual university bachelor's degree belongs to the STEM field		
STEM (subjects)	0.038*** (0.002)	0.004* (0.003)
STEM subjects by sub-category		
Computer and Information Sciences	0.068*** (0.004)	0.028*** (0.006)
Engineering	0.071*** (0.003)	0.093*** (0.005)
Biology and Life Sciences	0.041*** (0.004)	-0.015*** (0.004)
Mathematics and Statistics	0.097*** (0.006)	0.075*** (0.007)
Technologies	-0.056*** (0.006)	-0.061*** (0.012)
Physical Sciences	-0.004 (0.005)	-0.049*** (0.006)
Medical and Health Sciences and Services	0.140*** (0.012)	0.091*** (0.010)
Other STEM subjects	-0.118***(0.007)	-0.071*** (0.007)
Female/Male ratio in different degree fields		
Female ratio	-0.299*** (0.006)	-0.352*** (0.005)
Highest educational attainment		
Master's degree	0.178*** (0.002)	0.178*** (0.002)
Professional degree	0.476*** (0.004)	0.446*** (0.004)
Doctoral degree	0.274*** (0.004)	0.325*** (0.004)
Whether individual has second degree		
Degree2	0.045*** (0.003)	0.032*** (0.003)
whether enrolled in a public or a private school		
Public school	-0.179*** (0.005)	-0.144*** (0.004)
Private school	-0.150*** (0.007)	-0.010*** (0.005)

Note: Significant at: *p , 0.1, * *p , 0.05 and ***p , 0.01; robust standard errors are in parentheses. For the complete regression table, please check the Appendix.

The National Center for Education Statistics (1998) also mentioned that the concept of female social devaluation is a key element linking gender segregation and the gender wage gap. Given the great interest among many researchers in the difference in female representation in different subjects of the STEM fields, this paper will follow up with a further in-depth exploration of gender differences in the different STEM subjects. Table 11 reports the percentage of graduates in different degree fields by gender and major. And the proportion of males and females in each subject of the STEM field is shown in Figure 5.

According to Table 11, it can be seen that compared with males, women have a higher probability of graduating from Non-STEM degree fields and have a smaller probability of graduating from STEM fields. In addition, of all the eight subjects in the STEM field, males have the highest probability of obtaining an engineering degree and have the lowest probability of graduating from "Medical and Health Sciences and Services". Meanwhile, among the eight STEM subjects, females are the most likely to graduate with a degree in "Biology and Life Sciences", and the least likely to graduate with a technical-related degree.

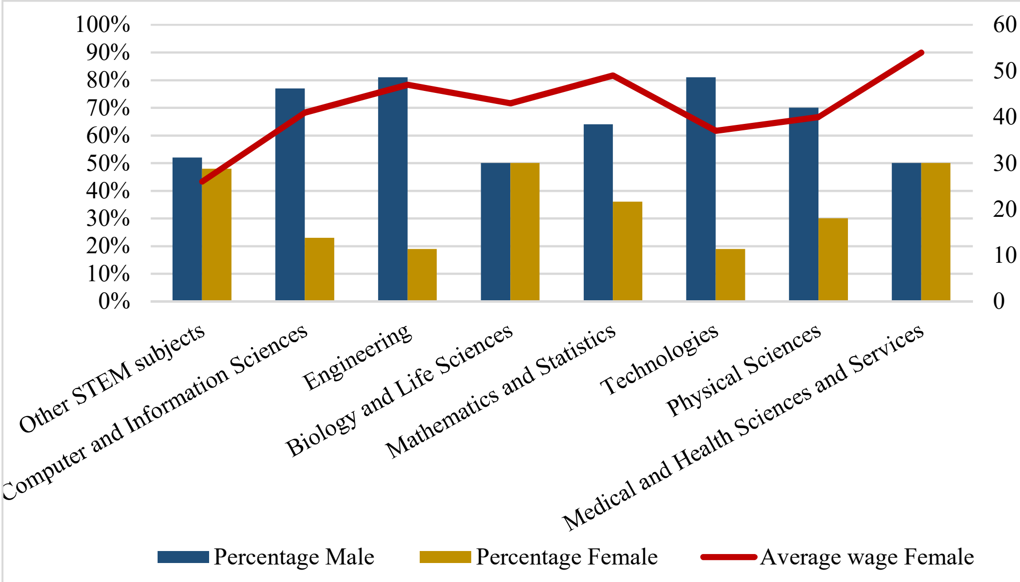
Table 11: Percentage of graduates in different degree fields by gender and major, 2018-2020, U.S.A.

Major	Male (N=516815)	Female (N=470221)	Difference (Male-Female)	Wage ratio (W_f/W_m)
STEM	37.69%	16.88%	20.80%	78.78%
Computer and Information Sciences	6.66%	2.11%	4.56%	83.32%
Engineering	15.66%	3.15%	12.51%	84.83%
Biology and Life Sciences	4.85%	5.37%	-0.52%	73.43%
Mathematics and Statistics	2.46%	1.53%	0.92%	79.38%
Technologies	2.29%	0.48%	1.81%	80.56%
Physical Sciences	3.79%	2.13%	1.66%	75.59%
Medical and Health Sciences and Services	0.54%	0.63%	-0.08%	79.35%
Other STEM subjects	1.43%	1.48%	-0.05%	83.42%
Non-STEM	62.31%	83.12%	-20.80%	74.54%

Note: The data all comes from the IPUMS database.

Figure 5 not only shows the gender differences in the distribution of the eight STEM subjects more intuitively but also reports the average hourly wages of females in each subject. It can be found that although females are equally represented as males in "Biology and Life Sciences" and "Medical and Health Sciences and Services". But females are significantly underrepresented in "Computer and Information Sciences", "Engineering", "Mathematics and Statistics", "Technologies" and "Physical Sciences". What's more, the subjects in which females can earn higher returns tend to be those subjects that are underrepresented by females.

Figure 5: Proportion of males and females in each subject of the STEM field and female wages, 2018-2020, U.S.A.



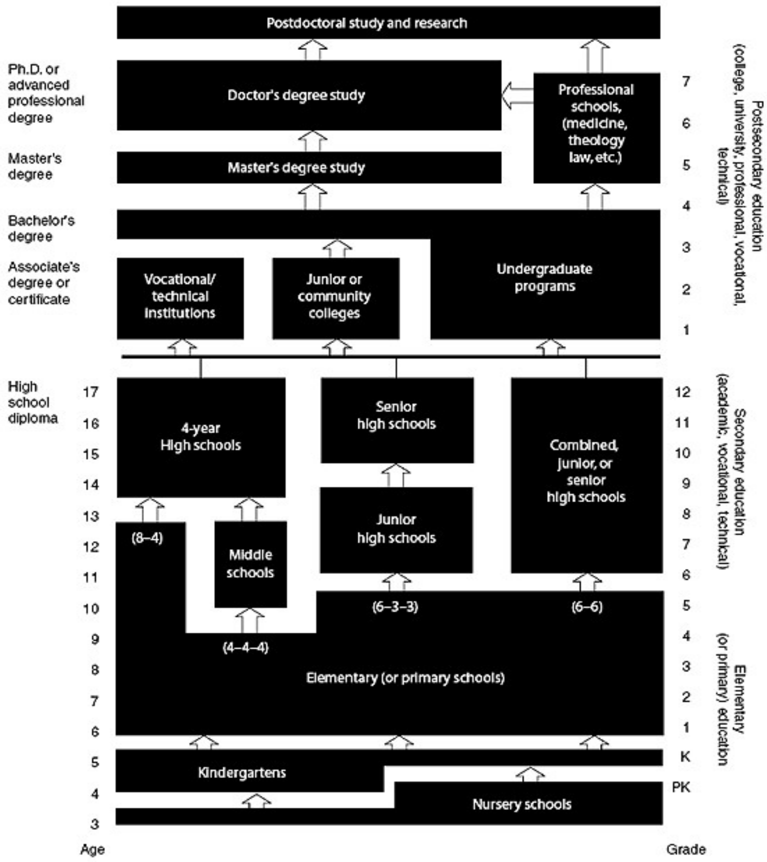
Note: The data all comes from the IPUMS database.

5.1.3. The influence of job factors on average annual income

In order to obtain the data of individual's work experience, this paper first estimates the years of education for each individual according to their educational attainment (bachelor, master, professional and doctorate). The basic education in the United States is called "K-12" which means that American citizens receive basic education from kindergarten until the 12th grade. Figure 4 provides a complete picture of the educational structure in the United States. When estimating the years of education, this paper starts from the first grade of primary school but not the kindergarten. In the dataset used in this paper, if an individual obtains a bachelor's degree, I add 4 years to the 12-year basic education; if an individual obtains a master's degree or a professional degree, I add 6 years[11] to the 12-year basic education; for those with a doctorate degree, I add 5.8 years to the 18-year master's degree.

11 The number of years required to earn a professional degree is derived from the document "Structure of the U.S. Education System: First-Professional Degrees" by the U.S. Department of Education International Affairs Office. According to this document, a first-professional degree is an award that requires completion of a program that meets the following three criteria: (1) Completion of the academic requirements to begin the practice of the profession; (2) At least 2 years of university work experience before entering this major; (3) Completion of a total of at least 6 academic years of college work in the first professional degree program, including the previously required college work plus the length of the professional program itself.

Figure 4: The structure of education, U.S.A.



Note: Chart reflects typical patterns of progression rather than all possible variations.
Source: U.S. Department of Education, National Center for Education Statistics.

Therefore, the number of years of education for those with a bachelor's degree is estimated to be 16 years in this paper, while the number of years of education for those with a master's and professional degrees is estimated to be 18 years, and the number of years of education for a Ph.D. is estimated to be 23.5 years. Then the work experience can be estimated according to the following formula:

$$\mathrm{Experience_i} = \mathrm{Age_i} - \mathrm{Education_i} - 5 \qquad (5)$$

where Age_i is the age of individual i and $Education_i$ is the number of years of education of individual i. The value 5 is the preschool age including kindergarten (See Figure 3 for details and it should be noted that the education received in kindergarten is not included in the calculated years of education in this study).

Finally, the regression results of all Job-related variables including "work experience" are shown in Table 12. Compared to the worker who works in Non-STEM fields, both males and females with STEM-related jobs earn higher earnings, and this advantage is more great for females. Besides, although for both males and females, working in the private sector has higher earnings, there are also significant gender differences across different sectors. Between the government sector and the nonprofit sector, females working in government earn less than that working in nonprofit sectors, while the opposite for males. As expected and empirically verified by typical human capital models, the incomes of both males and females increase with work experience and the work experience shows a diminishing marginal effect. In addition, the increase in working hours can also significantly increase the average annual income of both males and females.

Table 12: Regression results of income equation by gender (Job-related characteristics), 2018-2020, U.S.A.

Variables	Male (N=516,815)	Female (N=470,221)
Job-related characteristics		
Whether the occupation of an individual belongs to the STEM field		
STEM (occupation)	0.170*** (0.002)	0.238*** (0.002)
Job sector		
Government	-0.237*** (0.002)	-0.151*** (0.002)
Non-profit	-0.282*** (0.003)	-0.145*** (0.002)
Working experience	0.041*** (0.001)	0.040*** (0.001)
Working experience²	-0.001*** (0.000)	-0.001*** (0.000)
Usual hours worked per week	0.014*** (0.000)	0.014*** (0.000)

Note: Significant at: *p , 0.1, * *p , 0.05 and ***p , 0.01; robust standard errors are in parentheses. For the complete regression table, please check the Appendix.

5.2. Blinder-Oaxaca decomposition results

5.2.1. Blinder-Oaxaca decomposition of college graduates (STEM & Non-STEM graduates)

Table 13 reports the decomposition results of all subsamples (including STEM graduates and Non-STEM graduates) based on pooled regression results with the group indicator (female) in the pooled regression model. This paper conducts two separate wage decomposition using the binary degree field dummy variable as well as the detailed STEM subjects variable, and obtain two similar results in the end. Regardless of which degree field variable is used in Blinder-Oaxaca decomposition, the final decomposition results all show that among the highly-educated labor force in the U.S. for the period 2018-2020, the Demographic characteristics, Family characteristics, Education-related characteristics, and Job-related characteristics combined can explain 55.96% of the gender pay gap.

Table 13: Blinder-Oaxaca decomposition results of college graduates, 2018-2020, U.S.A.

	With binary degree variable	With detailed STEM subjects variable
difference	0.3028928***	0.3028928***
explained	0.1694874***	0.1694906***
(%)	55.96%	55.96%
unexplained	0.1334054***	0.1334022***
(%)	44.04%	44.04%

Note: Significant at: *p , 0.1, * *p , 0.05 and ***p , 0.01; The complete decomposition results can be found in the Appendix.The data all comes from the IPUMS database.

Table 14: Percentages of each Demographic and Family variable can explain the gender pay gap, 2018-2020, U.S.A.

	Variables	With binary degree variable		With detailed STEM subjects variable	
Background/Demographic Characteristics	Asian	0.00%	2.15%	0.00%	2.16%
	Black	1.47%		1.48%	
	Indian	0.12%		0.12%	
	Other	-0.28%		-0.29%	
	More-than-one	0.06%		0.06%	
	Northeast	0.00%		0.00%	
	Midwest	0.21%		0.21%	
	South	0.57%		0.57%	
Family Characteristics	Number of children	2.12%	6.46%	2.11%	6.43%
	Married	5.16%		5.14%	
	Single	-0.82%		-0.82%	

Note: The data all comes from the IPUMS database.

Tables 14 and 15 show us the percentage of each variable explaining the gender pay gap, respectively. In the results, all other variables are significant except the racial variable: "Asian" and the regional variable: "Northeast" which cannot significantly explain the gender wage gap. In addition, the two tables also provide a comparison of the changes of each variable in the two wage decomposition. According to Table 14, the gender wage gap explained by Demographic characteristics combined with Family characteristics is very small, about 8.16% - 8.59%. In particular, Demographic characteristics can only explain 2.15% - 2.16% of the gender pay gap.

Table 15: Percentages of each Educational and Job-related variable can explain the gender pay gap, 2018-2020, U.S.A.

	Variables	With binary degree variable		With detailed STEM subjects variable	
	STEM/Non-STEM	1.35%			
	Computer and Information Sciences			0.78%	
	Engineering			2.40%	
	Biology and Life Sciences			-0.02%	
	Mathematics and Statistics			0.26%	
	Technologies			-0.35%	
	Physical Sciences			-0.15%	
Educational Characteristics	Medical and Health Sciences and Services		20.62%	-0.03%	20.82%
	Other STEM subjects			0.02%	
	Female ratio	19.97%		18.55%	
	Public school	0.64%		0.63%	
	Private school	0.26%		0.26%	
	Master's degree	-3.39%		-3.35%	
	Professional degree	0.92%		0.93%	
	Doctoral degree	0.89%		0.90%	
	STEM/Non-STEM	4.49%		4.35%	
	Non-profit	4.91%		4.88%	
Job-related Characteristics	Government	5.79%	26.72%	5.76%	26.55%
	Working experience	3.36%		3.35%	
	Hours worked per week	8.17%		8.19%	

Note: The data all comes from the IPUMS database.

By contrast with Demographic characteristics and Family characteristics, Education and Job characteristics explain a much more large proportion of the gender pay gap respectively. Nevertheless, in the Education-related variables, the gender pay gap explained by degree fields which are most concerned in this study is very small, and even when I subdivide the STEM subjects, it explains only 2.91% of the gender pay gap, which only explained 1.56% more than using the binary degree variable. Additionally, the representation of females in various degree subjects can explain a large part of the gender pay gap, about 18.55% - 19.97%. Also, it should be noted that, in the Education-related variables, there is a variable shown with a negative percentage explaining the gender wage gap: "*Master's degree*", which means that this variable instead helps to narrow the gender pay gap. Based on what we observed earlier in the mean of each variable, females are more likely to graduate with a master's degree than males. In the opinion of this paper, this should be an important reason why the coefficient of this variable is negative in the decomposition result.

Job characteristics are among all the observable characteristics in this study that can most explain the gender pay gap for US college graduates. According to Table 15, it can be found that whether an individual is employed in a STEM occupation explains 4.35% - 4.49% of the gender pay gap, while work experience can explain 3.35% - 3.36% of the gender pay gap. In addition to this, the 8.17% - 8.19% gender pay gap can be explained by working hours. And of all Job-related variables, the "private sector" combined with the "government sector" together explained 10.64% - 10.7% of the gender pay gap, which illustrates that the job sector is a very important factor affecting the gender pay gap.

5.2.2. Blinder-Oaxaca decomposition of STEM and Non-STEM graduates separately.

As found in the discussion of the gender pay gap in Section 3.2.2, the gender pay gap calculated based on median and mean wages vary considerably. The observed gender pay gap for STEM graduates was greater than for Non-STEM graduates when calculated using the median, whereas the opposite results were obtained with the mean wage. In addition, it would be interesting to discuss the respective gender pay gaps in STEM and Non-STEM fields, thus, I will report the respective wage gaps and wage decomposition results in STEM graduates and Non-STEM graduates as follows, and give a brief discussion. Besides, what should be noted is that the Blinder-Oaxaca decomposition method is a typical mean decomposition method, thus the conclusions obtained are likely to be consistent with the mean-wage-based calculation of the gender pay gap. That is, the observed gender pay gap for STEM graduates is smaller than for non-STEM graduates.

Based on the limited sample of this study, I successively removed STEM graduates and Non-STEM graduates from the subsample. Then decompose the gender wages with the same method, and explain the gender wage gap with similar Demographic characteristics, Family characteristics, Educational characteristics, and Job-related characteristics in the STEM graduates sample and Non-STEM graduates sample.

The decomposition results are shown in Table 16. As expected, when the mean decomposition method is used to study and decompose the gender pay gap. It can be found that the gender pay gap among Non-STEM graduates is larger than that of STEM graduates. In addition, all the four categories of characteristics involved in this study combined can explain 46.55% of the gender pay gap among STEM graduates and can explain 42.34% of the pay gap among Non-STEM graduates. Figure 5 compares the percentage each type of characteristics that explains the gender pay gap for STEM graduates and Non-STEM graduates, respectively. And Figure 6 shows the percentage of education-related variables explaining the gender pay gap for STEM graduates and Non-STEM graduates, respectively.

Table 16: Blinder-Oaxaca decomposition results of STEM graduates and Non-STEM Graduates respectively, 2018-2020, U.S.A.

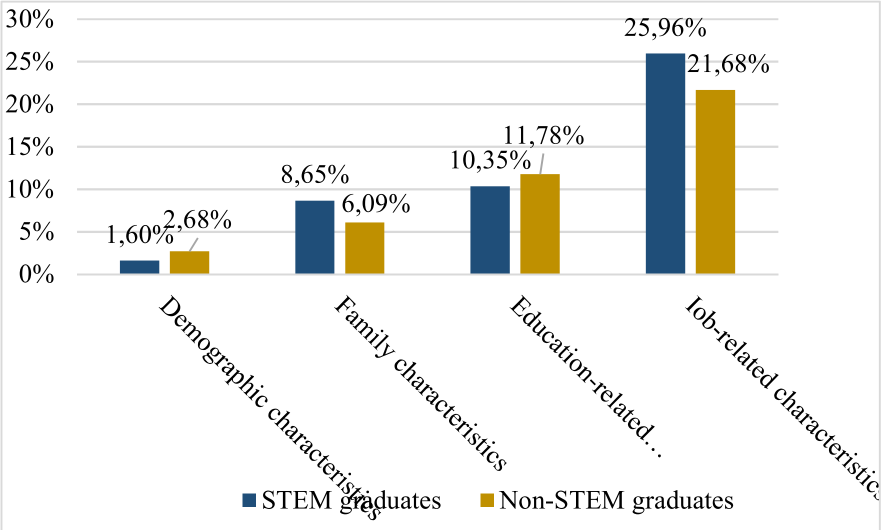
	STEM graduates (N= 274,171)	Non-STEM graduates (N= 712,865)
group 1	11.54083	11.32992
group 2	11.28969	11.10652
difference	0.2511403***	0.2606121***
explained	0.1172169***	0.110347***
(%)	46.55%	42.34%
unexplained	0.1339234***	0.1502651***
(%)	53.45%	57.66%

Note: Significant at: *p , 0.1, * *p , 0.05 and ***p , 0.01; The data all comes from the IPUMS database.

According to Figure 5, it can be found that Job-related variables are the most influential characteristics of the gender pay gap both for STEM graduates and Non-STEM graduates, which explained 25.96% of the gender pay gap for STEM graduates and 21.68% of the pay gap for Non-STEM graduates, respectively. In addition, the Education-related characteristics also have a significant impact on explaining the gender wage gap in the two groups respectively. According to Figure 6, among all education variables, the percentage of females in subjects is the one variable that has the greatest impact on the gender pay gap. Besides, although the degree fields variable contributes very little to explaining the gender pay gap for Non-STEM graduates, as well as explaining the gender pay gap for the entire subsample. However, in the sample of

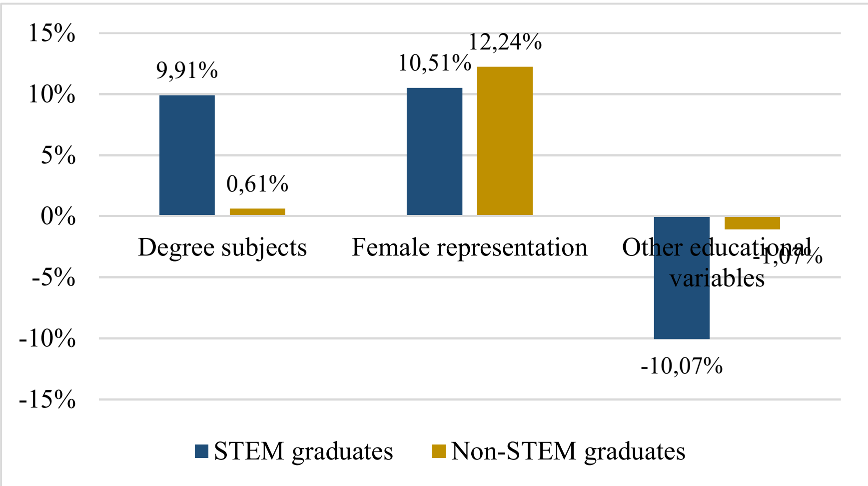
STEM graduates, this variable can explain 9.91% of the gender pay gap of STEM graduates.

Figure 5: Comparison of percentage of four basic characteristics explaining the gender pay gap for STEM graduates and Non-STEM graduates, respectively, 2018-2020, U.S.A.



Note: The data all comes from the IPUMS database.

Figure 6: Comparison of percentage of education-related variables explaining the gender pay gap for STEM graduates and Non-STEM graduates, respectively, 2018-2020, U.S.A.



Note: The data all comes from the IPUMS database.

6. Conclusion

6.1. Summary and Discussion

With the development and progress of society in the field of science and technology and the great success and progress of the United States in the field of science and technology, both the private sector and the public sector all demand more talents in the STEM field urgently. In addition, not only in the United States but more and more other countries and scholars in different countries have begun to pay more attention to the STEM field and advocate increasing the diversity in the STEM field. Thus it is necessary and meaningful to explore gender segregation and gender pay gaps in STEM and non-STEM fields in the labor market.

The main purpose of this paper is to use the 2018 – 2020 U.S. Census and American Community Survey microdata to examine the extent to which gender differences in graduates' choice of degree fields - STEM or Non-STEM contribute to the gender pay gap in the United States. With the Blinder-Oaxaca wage decomposition method, this paper finds that among the highly-educated labor force in the U.S. for the period 2018-2020, the individual background characteristics(race, region), family characteristics (marital status, number of children), education-related characteristics (degree fields, the highest educational attainment, and school type), and job-related characteristics (hours worked, work experience, job sector, and occupation type) combined can explain 55.96% of the gender pay gap. Among them, whether an individual chooses a college major related to a STEM field (ie. whether or not to obtain a degree in a STEM field) contributes 2.4% - 5.2% of the "explanatory part", and can only explain 1.35% - 2.91% of the gender pay gap. Finally, the results of this study showed that, although the content of the field of study (STEM / Non-STEM) seems not to explain too much of the current gender wage gap between the U.S. highly-educated workforce, it is still hard to deny that educational segregation is a barrier to gender equality.

Let us recall the concept mentioned earlier - "societal devaluation of feminine work". According to The National Center for Education Statistics (1998), the concept of female social devaluation is a key element linking gender segregation and the gender wage gap. Thus, in order to make up for the few empirical analyses that measure the contribution of female dominance in professions to the gender pay gap. Similar to Lim (2016), this study also includes a variable to measure the proportion of females in each subject. Although the representation of females in different subjects is not a statistically significant determinant of wages in Lim (2016), this is an important variable in explaining the gender pay gap among all education-related variables in my research, which can explain 18.55% - 19.97% of the gender pay gap of American university graduates. Furthermore, this result is very close to Bobbitt -Zeher (2007). In the research of Bobbitt-Zeher (2007), they found that the "gender dominance of the major" (ie, the percentage of females in the field of study) is one of the most pronounced differences between university majors, explaining about 14% of the overall gender pay gap. In general, gender segregation in education across degree fields or subjects is indeed a very tough issue for gender equality.

In summary, the research of this paper mainly focuses on three aspects:

(1) Research and explain what percentage of the gender pay gap among U.S. highly-educated workers can be explained by the choice of graduating from a STEM field or a non-STEM field; (2) Although the effect of "degree field choice" on the gender pay gap of US college graduates is very small, degree field is not the only factor in educational segregation. By adding a continuous variable that measures the percentage of females in each degree subject, it can be found that, except for Job-related variables, this variable has the most significant impact on explaining the gender pay gap; (3) Based on the interest and concern of many researchers on the gender gap in STEM fields. This paper further divides the subsample into a STEM graduate sample and a Non-STEM graduate sample. The gender wages of the two sample populations are also decomposed using the Blinder-Oaxaca decomposition method. Finally, this study found that the gender pay gap for STEM graduates is 3.77% smaller than the gender pay gap for Non-STEM graduates. Besides, the choice of degree subjects can explain 9.91% of the gender gap for STEM graduates.

6.2. Implications

Based on the above research results, this research mainly proposes two policy implications:

(1) Increase the diversity of STEM fields and encourage more women to study STEM subjects, especially those subjects with very low female representation in STEM fields.

In fact, there are a lot of actions and policies that have been put in place to work on this. In December 2018, "Strategic Planning for STEM Education, Planning the Road to Success: A Strategy for U.S. STEM Education," which jointly published by the Office of the President and the Office of Science and Technology Policy, explicitly mentioned three major goals for U.S. STEM education, and one of them is - "Increase Diversity, Equity, and Inclusion in STEM. Besides". In addition, the Trump Administration has already taken multiple actions for it. The Trump Administration championed and signed into law the Strengthening Career and Technical Education for the 21st Century Act to increase student access across secondary and post-secondary levels to high-quality technical education and credentialing; In addition to this, the Trump Administration also directed the Department of Education (ED) to expand access to high-quality STEM and computer science education to K-12 students.

These actions reflect the administration's recognition of the importance of STEM education and training as a driver of job creation and economic prosperity in the United States. However, according to the research results obtained in this paper using the latest data, in order to effectively increase the participation rate of females in the STEM field, it is necessary to increase the deepness of policy implementation and the wideness of policy coverage.

(2) Mitigate and reduce "female devaluation" in the labor market.

"Female devaluation" is reflected in the fact that occupations with higher representation of women are often those with lower income returns. According to the regression results of the income equation research in this paper, "female devaluation" exists not only in occupations, but is also deeply rooted in education. The variable of the proportion of females in different degree subjects has a negative impact not only on women's earnings but also on men's earnings. Furthermore, the study according to Levanon et al. (2009) found that once women entered the occupation, the average salary of the occupation decreased. This all indicates that the problem of discrimination against females in the labor market has always existed.

The U.S. Equal Pay Act requires men and women in the same workplace to be paid equally for equal work. If there is wage inequality between men and women, employers may not reduce the wages of any gender to make their wages equal. In addition, the law makes it illegal to discriminate on the basis of sex in pay and benefits. In fact, there are many similar laws, regulations, and policies like the Equal Pay Act to protect females obtain equal benefits and basic rights in the workplace, but why are women still discriminated against? In my opinion, policymakers should focus on supervising employers in addition to only focusing on making policies and laws. No matter how perfect the law is, if it is not effectively implemented, then everything will be in vain.

6.3. Limitations and Directions for Future Research

Although this study can make a meaningful contribution to the knowledge base for the gender wage gap among American graduates, to some extent, there are inevitably many limitations. First, this paper fails to effectively deal with the potential problem of sample selection bias. For instance, in order to define "full-time workers", this study excludes workers in the sample who work less than 50 weeks per year and who work less than 40 hours per week. According to descriptive statistical analysis, women are more likely than men to work part-time rather than full-time. This paper, therefore, excludes a disproportionately large sample of part-time women. In conclusion, the potential existence of sample selection bias is likely to affect the results of the study. Second, some variables in this study, such as the "work weeks" in 2018 and "work experience" variables, are calculated by myself based on some literature and official documents, and these variables may not be accurate. Finally, the variables included in this study and the characteristics that can be observed are limited, and this is why the variables included in this study can only explain about half of the gender pay gap. Compare to the study of Lim (2016) which is very similar to mine, he explained up to 82% of the gender pay gap for U.S. STEM graduates by using a different database and incorporating more variables, while in my study only 46.55% of the gender pay gap for STEM graduates was explained.

In future research, I will pay more effort to work on improving the econometric techniques, and try to use different databases and more variables to continue to explore the gender wage gap in the labor market of various countries. In addition to this, I will study more about to the issue of "female devaluation" and the issue of discrimination against women in future research, which is almost the blank part of this paper.

References

Alegria, S. N., & Branch, E. H. (2015). Causes and Consequences of Inequality in the STEM: Diversity and its Discontents. International Journal of Gender, Science and Technology, 7(3), 321-342.

Barón, J. D., & Cobb-Clark, D. A. (2010). Occupational segregation and the gender wage gap in private-and public-sector employment: a distributional analysis. Economic Record, 86(273), 227-246.

Bayard, K., Hellerstein, J., Neumark, D., & Troske, K. (2003). New evidence on sex segregation and sex differences in wages from matched employee-employer data. Journal of labor Economics, 21(4), 887-922.

Benavent, X., de Ves, E., Forte, A., Botella-Mascarell, C., López-Iñesta, E., Rueda, S., ... & Marzal, P. (2020). Girls4STEM: Gender diversity in STEM for a sustainable future. Sustainability, 12(15), 6051.

Benbow, C. P., Lubinski, D., Shea, D. L., & Eftekhari-Sanjani, H. (2000). Sex Differences in Mathematical Reasoning Ability at Age 13: Their Status 20 Years Later. Psychological Science, 11(6), 474–480.

Bender, K. A. (1998). The central government - private sector wage differential. Journal of economic surveys, 12(2), 177-220.

Bergmann, B. R. (1974). Occupational segregation, wages and profits when employers discriminate by race or sex. Eastern economic journal, 1(2), 103-110.

Bilimoria, D., & Lord, L. (2014). Women in STEM careers international perspectives on increasing workforce participation, advancement, and leadership. Northampton, MA: Edward

Elgar Publishing Inc.

Blau, F. D., & Kahn, L. M. (2008). Women's work and wages. The new Palgrave dictionary of economics, 8, 762-72.

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

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. The Journal of Human Resources, 8(4), 436–55.

Bobbitt-Zeher, D. (2007). The Gender Income Gap and the Role of Education. Sociology of Education, 80(1), 1–22.

Bol, T., & Heisig, J. P. (2021). Explaining wage differentials by field of study among higher education graduates: Evidence from a large-scale survey of adult skills. Social Science Research, 99, 102594.

Bottia, M. C., Stearns, E., Mickelson, R. A., Moller, S., & Valentino, L. (2015). Growing the roots of STEM majors: Female math and science high school faculty and the participation of students in STEM. Economics of Education Review, 45, 14-27.

Brainard, G. S. & Carlin, L. (1998). A Six-Year Longitudinal Study of Undergraduate Women in Engineering and Science. Journal of engineering education, 87(4), 369–375.

Brown, R.， M. Moon， and B. Zoloth， "Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials"， Journal of Human Resources， 1980， 15(1)， 3—28.

Brown, T.C. (1993). Descriptions of the Characteristic Sequence of an Irrational. Canadian Mathematical Bulletin, 36, 15 - 21.

Broyles, P. (2009). The gender pay gap of STEM professions in the United States. International journal of sociology and social policy.

Cavaglia, C., Machin, S., McNally, S., & Ruiz-Valenzuela, J. (2020). Gender, achievement, and subject choice in English education. Oxford Review of Economic Policy, 36(4), 816 – 835.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. Psychological Bulletin Journal, 135(2), 218-61.

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? Psychological Bulletin, 143(1), 1–35.

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others?. Psychological bulletin, 143(1), 1.

Choudhury, S. (1994). New evidence on public sector wage differentials. Applied Economics, 26(3), 259-266.

Dauda, R. O. S. (2013). Does female education promote economic performance? Evidence from Nigeria. International Journal of Economics and Finance, 5(1), 201-209.

De Meulemeester, J. L., & Rochat, D. (1995). A causality analysis of the link between higher education and economic development. Economics of education review, 14(4), 351-361.

Elder, T., Goddeeris, J., & Haider, S. (2010). Unexplained gaps and Oaxaca – Blinder decompositions. Labour Economics, 17, 284-290.

Fayer, S., Lacey, A., & Watson, A. (2017). STEM occupations: Past, present, and future. Spotlight on Statistics, 1, 1-35.

Gharehgozli, O., & Atal, V. (2020). Revisiting the gender wage gap in the United States. Economic Analysis and Policy, 66, 207-216.

Good, C., Rattan, A., Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. Journal of Personality and Social Psychology, 102(4), 700-17.

Gradstein, M. & Justman, M. (2002). Education, Social Cohesion, and Economic Growth. The American Economic Review, 92(4), 1192－1204.

Griffin, P., Care, E., McGaw, B. (2012). The Changing Role of Education and Schools. In: Griffin, P., McGaw, B., Care, E. (eds) Assessment and Teaching of 21st Century Skills. Springer, Dordrecht.

Guo Jiqiang, Jiang Li, & Lu Lili. (2011). A Review of Wage Difference Decomposition Methods (Doctoral dissertation).

Gylfason, T. (2001). Natural resources, education, and economic development. European economic review, 45(4-6), 847-859.

Hamidi, F., Meshkat, M., Rezaee, M., & Jafari, M. (2011). Information technology in education. Procedia Computer Science, 3, 369-373.

Heath, R., & Jayachandran, S. (2016). The causes and consequences of increased female education and labor force participation in developing countries (No. w22766). National Bureau of Economic Research.

Heckman, J., "Sample Selection Bias as a Specification Error", Econometrica, 1979, 47(1), 153—161.

Hill, C., Corbett, C., & Rose, A. S. (2010). Why so few women? Women in science, technology, engineering, and mathematics. Washington, DC: American Association of University Women.

Huffman, M. L., & Cohen, P. N. (2004). Racial wage inequality: Job segregation and devaluation across US labor markets. American Journal of Sociology, 109(4), 902-936.

Hutchison, L. F. (2012). Addressing the STEM teacher shortage in American schools: Ways to recruit and retain effective STEM teachers. Action in Teacher Education, 34(5-6), 541-550.

Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. The Stata Journal, 8(4), 453-479.

Juhn, C., K. Murphy, and B. Pierce, "Accounting for the Slowdown in Black-White Wage Convergence", in Kosters, M. (ed.), Workers and Their Wages : Changing Patterns in the United States. Washington: American Enterprise Institute Press, 1991.

Kahn, L. M. (2015). Wage compression and the gender pay gap. IZA World of Labor.

Kanematsu, H., Barry, D.M. (2016). The Importance of STEM for Modern Education. In: STEM and ICT Education in Intelligent Environments. Intelligent Systems Reference Library, vol 91. Springer, Cham.

Lamo, A., Pérez, J. J., & Schuknecht, L. (2012). Public or private sector wage leadership? An international perspective. The Scandinavian Journal of Economics, 114(1), 228-244.

Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). STEM: Good Jobs Now and for the Future. ESA Issue Brief# 03-11. US Department of Commerce.

Levanon, A., England, P., & Allison, P. (2009). Occupational feminization and pay: Assessing causal dynamics using 1950-2000 US census data. Social Forces, 88(2), 865-891.

Lim, K. M. (2016). Major matters: Exploration of the gender wage gap among STEM graduates. University of California, Los Angeles.

Livanos, I. & Pouliakas, K. (2012). Educational Segregation and the Gender Wage Gap in Greece. Journal of Economic Studies, 39(5), 554–575.

McDonald, C. V. (2016). STEM Education: A review of the contribution of the disciplines of science, technology, engineering and mathematics. Science Education International, 27(4), 530-569.

Michelmore, K., & Sassler, S. (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), 194-215.

Miriti, M. N. (2020). The elephant in the room: race and STEM diversity. BioScience, 70(3), 237-242.

Morris, V. R., & Washington, T. M. (2017). The role of professional societies in STEM diversity. Not Am Math Soc, 65.

National Center for Education Statistics. (1998). Gender Differences in Earnings Among Young Adults Entering the Labor Market (NCES 98-086). By Suzanne B. Clery, John B. Lee, and Laura G. Knapp. Washington, DC: U.S. Government Printing Office.

Noonan, R. (2017). STEM Jobs: 2017 Update. ESA Issue Brief# 02-17. US Department of Commerce.

Noonan, R. (2017). Women in STEM: 2017 Update. ESA Issue Brief# 06-17. US Department of Commerce.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. International Economic Review, 14(3), 693–709.

Oblova, I. S., Gerasimova, I. G., & Sishchuk, J. M. (2020). Gender segregation in STEM education and careers in Russia. Global Journal of Engineering Education, 22(2), 130-135.

O'Bryant, K. (2003). Fraenkel's partition and Brown's decomposition. ArXiv preprint math/0305133.

Olson, J. E. (2013). Human capital models and the gender pay gap. Sex Roles, 68(3), 186-197.

Ramaley, J. A. (2007). Facilitating change: Experiences with the reform of STEM education.

Rothwell, J. (2013). The hidden STEM economy. Washington, DC: Metropolitan Policy Program at Brookings.

Rutledge, M. S. (2018). What Explains the Widening Gap in Retirement Ages by Education? Center for Retirement Research of Boston College. Retrieved July 25th, 2022, from https://crr.bc.edu/wp-content/uploads/2018/05/IB_18-10.pdf

Sahoo, S., & Klasen, S. (2018). Gender segregation in education and its implications for labour market outcomes: Evidence from India. IZA Institute of Labor Economics, Discussion Paper Series, No. 11660.

Sattari, N. & Sandefur, R. L. (2019). Gender in Academic STEM: A Focus on Men Faculty. Gender, work, and organization, 26(2), 158–179.

Scott, K. A., & Elliott, S. (2019). STEM diversity and inclusion efforts for women of color: A critique of the new labor system. International Journal of Gender, Science and Technology, 11(3), 374-382.

Tabvuma, V., Georgellis, Y., & Lange, T. (2015). Orientation training and job satisfaction: A sector and gender analysis. Human Resource Management, 54(2), 303-321.

Tienxhi, J. Y. (2017). The Gender Gap in Malaysian Public Universities: Examining the'Lost Boys. Journal of International and Comparative Education (JICE), 1-16.

Van den Hurk, A., Meelissen, M., & van Langen, A. (2019). Interventions in education to prevent STEM pipeline leakage. International Journal of Science Education, 41(2), 150-164.

van Langen, A. (2015). Girls and STEM choice in Dutch education: The strong gender segregation and the good practice of the stimulation policy. Gender Studies & Research, 13, 26-39.

Von Keyserlingk, L., Becker, M., Jansen, M., & Maaz, K. (2020). Leaving the pond—Choosing an ocean: Effects of student composition on STEM major choices at university. Journal of Educational Psychology, 112(4), 751–764.

Wang, M. T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. Educational psychology review, 29(1), 119-140.

Wang, X., Xu, W., & Guo, L. (2018). The Status Quo and Ways of STEAM Education Promoting China's Future Social Sustainable Development. Sustainability, 10(12), 4417.

Winters. (2014). STEM graduates, human capital externalities, and wages in the U.S. Regional Science and Urban Economics, 48(September), 190–198.

Appendix

Table A-1: List of degree fields classified as STEM

Code	Label
	Agriculture
1103	Animal Sciences
1104	Food Science
1105	Plant Science and Agronomy
1106	Soil Science
	Environment and Natural Resources
1301	Environmental Science
1302	Forestry
	Communication Technologies
2001	Communication Technologies
	Computer and Information Sciences
2100	Computer and Information Systems
2101	Computer Programming and Data Processing
2102	Computer Science
2105	Information Sciences
2106	Computer Information Management and Security
2107	Computer Networking and Telecommunications
	Engineering
2400	General Engineering
2401	Aerospace Engineering
2402	Biological Engineering
2403	Architectural Engineering
2404	Biomedical Engineering
2405	Chemical Engineering
2406	Civil Engineering
2407	Computer Engineering
2408	Electrical Engineering
2409	Engineering Mechanics, Physics, and Science
2410	Environmental Engineering
2411	Geological and Geophysical Engineering
2412	Industrial and Manufacturing Engineering
2413	Materials Engineering and Materials Science
2414	Mechanical Engineering
2415	Metallurgical Engineering
2416	Mining and Mineral Engineering
2417	Naval Architecture and Marine Engineering

2418	Nuclear Engineering
2419	Petroleum Engineering
2499	Miscellaneous Engineering
	Engineering Technologies
2500	Engineering Technologies
2501	Engineering and Industrial Management
2502	Electrical Engineering Technology
2503	Industrial Production Technologies
2504	Mechanical Engineering Related Technologies
2599	Miscellaneous Engineering Technologies
	Biology and Life Sciences
3600	Biology
3601	Biochemical Sciences
3602	Botany
3603	Molecular Biology
3604	Ecology
3605	Genetics
3606	Microbiology
3607	Pharmacology
3608	Physiology
3609	Zoology
3611	Neuroscience
3699	Miscellaneous Biology
	Mathematics and Statistics
3700	Mathematics
3701	Applied Mathematics
3702	Statistics and Decision Science
	Military Technologies
3801	Military Technologies
	Interdisciplinary and Multi-Disciplinary Studies (General)
4002	Nutrition Sciences
4003	Neuroscience
4005	Mathematics and Computer Science
4006	Cognitive Science and Biopsychology
	Physical Sciences
5000	Physical Sciences
5001	Astronomy and Astrophysics
5002	Atmospheric Sciences and Meteorology
5003	Chemistry
5004	Geology and Earth Science
5005	Geosciences
5006	Oceanography

5007	Physics
5008	Materials Science
5098	Multi-disciplinary or General Science
	Nuclear, Industrial Radiology, and Biological Technologies
5102	Nuclear, Industrial Radiology, and Biological Technologies
	Psychology
5205	Industrial and Organizational Psychology
5206	Social Psychology
	Transportation Sciences and Technologies
5901	Transportation Sciences and Technologies
	Medical and Health Sciences and Services
6106	Health and Medical Preparatory Programs
6108	Pharmacy, Pharmaceutical Sciences, and Administration
	Business
6202	Actuarial Science
6212	Management Information Systems and Statistics

Table A-2: List of occupations classified as STEM

Code	Label
Management, Business, and Financial Occupations:	
110	Computer and information systems managers
300	Architectural and engineering managers
350	Medical and health services managers
360	Natural sciences managers
Computer, Engineering, and Science Occupations:	
1005	Computer and information research scientists
1006	Computer systems analysts
1007	Information security analysts
1010	Computer programmers
1021	Software developers
1022	Software quality assurance analysts and testers
1031	Web developers
1032	Web and digital interface designers
1050	Computer support specialists
1065	Database administrators and architects
1105	Network and computer systems administrators
1106	Computer network architects
1108	Computer occupations, all other
1200	Actuaries
1220	Operations research analysts
1240	Other mathematical science occupations
1305	Architects, except landscape and naval
1306	Landscape architects
1310	Surveyors, cartographers, and photogrammetrists
1320	Aerospace engineers
1340	Biomedical and agricultural engineers
1350	Chemical engineers
1360	Civil engineers
1400	Computer hardware engineers
1410	Electrical and electronics engineers
1420	Environmental engineers
1430	Industrial engineers, including health and safety
1440	Marine engineers and naval architects
1450	Materials engineers
1460	Mechanical engineers

1520	Petroleum, mining and geological engineers, including mining safety engineers
1530	Other engineers
1541	Architectural and civil drafters
1545	Other drafters
1551	Electrical and electronic engineering technologists and technicians
1555	Other engineering technologists and technicians, except drafters
1560	Surveying and mapping technicians
1600	Agricultural and food scientists
1610	Biological scientists
1640	Conservation scientists and foresters
1650	Other life scientists
1700	Astronomers and physicists
1710	Atmospheric and space scientists
1720	Chemists and materials scientists
1745	Environmental scientists and specialists, including health
1750	Geoscientists and hydrologists, except geographers
1760	Physical scientists, all other
1800	Economists
1821	Clinical and counseling psychologists
1822	School psychologists
1825	Other psychologists
1840	Urban and regional planners
1860	Other social scientists
1900	Agricultural and food science technicians
1910	Biological technicians
1920	Chemical technicians
1935	Environmental science and geoscience technicians, and nuclear technicians
1970	Other life, physical, and social science technicians
Healthcare Practitioners and Technical Occupations:	
3000	Chiropractors
3010	Dentists
3030	Dietitians and nutritionists
3040	Optometrists
3050	Pharmacists
3090	Physicians
3100	Surgeons
3110	Physician assistants
3120	Podiatrists

3140	Audiologists
3150	Occupational therapists
3160	Physical therapists
3200	Radiation therapists
3210	Recreational therapists
3220	Respiratory therapists
3230	Speech-language pathologists
3245	Other therapists
3250	Veterinarians
3255	Registered nurses
3256	Nurse anesthetists
3258	Nurse practitioners, and nurse midwives
3261	Acupuncturists
3270	Healthcare diagnosing or treating practitioners, all other
3300	Clinical laboratory technologists and technicians
3310	Dental hygienists
3321	Cardiovascular technologists and technicians
3322	Diagnostic medical sonographers
3323	Radiologic technologists and technicians
3324	Magnetic resonance imaging technologists
3330	Nuclear medicine technologists and medical dosimetrists
3401	Emergency medical technicians
3402	Paramedics
3421	Pharmacy technicians
3422	Psychiatric technicians
3423	Surgical technologists
3424	Veterinary technologists and technicians
3430	Dietetic technicians and ophthalmic medical technicians
3500	Licensed practical and licensed vocational nurses
3515	Medical records specialists
3520	Opticians, dispensing
3545	Miscellaneous health technologists and technicians
3550	Other healthcare practitioners and technical occupations
Sales and Related Occupations:	
4930	Sales engineers

Table A-3: Classification of subjects in STEM fields

	Code	Label
Other STEM	1103	Animal Sciences
	1104	Food Science
	1105	Plant Science and Agronomy
	1106	Soil Science
	1301	Environmental Science
	1302	Forestry
	4002	Nutrition Sciences
	4005	Mathematics and Computer Science
	4006	Cognitive Science and Biopsychology
	5205	Industrial and Organizational Psychology
	5206	Social Psychology
Computer and Information Sciences	2100	Computer and Information Systems
	2101	Computer Programming and Data Processing
	2102	Computer Science
	2105	Information Sciences
	2106	Computer Information Management and Security
	2107	Computer Networking and Telecommunications
Engineering	2400	General Engineering
	2401	Aerospace Engineering
	2402	Biological Engineering
	2403	Architectural Engineering
	2404	Biomedical Engineering
	2405	Chemical Engineering
	2406	Civil Engineering
	2407	Computer Engineering
	2408	Electrical Engineering
	2409	Engineering Mechanics, Physics, and Science
	2410	Environmental Engineering
	2411	Geological and Geophysical Engineering
	2412	Industrial and Manufacturing Engineering
	2413	Materials Engineering and Materials Science
	2414	Mechanical Engineering
	2415	Metallurgical Engineering
	2416	Mining and Mineral Engineering
	2417	Naval Architecture and Marine Engineering
	2418	Nuclear Engineering

	2419	Petroleum Engineering
	2499	Miscellaneous Engineering
Biology and Life Sciences	3600	Biology
	3601	Biochemical Sciences
	3602	Botany
	3603	Molecular Biology
	3604	Ecology
	3605	Genetics
	3606	Microbiology
	3607	Pharmacology
	3608	Physiology
	3609	Zoology
	3611	Neuroscience
	3699	Miscellaneous Biology
Mathematics and Statistics	3700	Mathematics
	3701	Applied Mathematics
	3702	Statistics and Decision Science
	6202	Actuarial Science
	6212	Management Information Systems and Statistics
Technology	2001	Communication Technologies
	2500	Engineering Technologies
	2501	Engineering and Industrial Management
	2502	Electrical Engineering Technology
	2503	Industrial Production Technologies
	2504	Mechanical Engineering Related Technologies
	2599	Miscellaneous Engineering Technologies
	3801	Military Technologies
	5102	Nuclear, Industrial Radiology, and Biological Technologies
	5901	Transportation Sciences and Technologies
Physical Sciences	5000	Physical Sciences
	5001	Astronomy and Astrophysics
	5002	Atmospheric Sciences and Meteorology
	5003	Chemistry
	5004	Geology and Earth Science
	5005	Geosciences
	5006	Oceanography
	5007	Physics
	5008	Materials Science
	5098	Multi-disciplinary or General Science

Medical and Health Sciences and Services	6106	Health and Medical Preparatory Programs
	6108	Pharmacy, Pharmaceutical Sciences, and Administration

Table A-4： Classification of subjects in Non-STEM fields

	Code	Label
Agriculture	1100	General Agriculture
	1101	Agriculture Production and Management
	1102	Agricultural Economics
	1199	Miscellaneous Agriculture
Environment and Natural Resources	1300	Environment and Natural Resources
	1303	Natural Resources Management
Architecture	1401	Architecture
Area, Ethnic, and Civilization Studies	1501	Area, Ethnic, and Civilization Studies
Communications	1900	Communications
	1901	Communications
	1902	Journalism
	1903	Mass Media
	1904	Advertising and Public Relations
Cosmetology Services and Culinary Arts	2201	Cosmetology Services and Culinary Arts
Education Administration and Teaching	2300	General Education
	2301	Educational Administration and Supervision
	2303	School Student Counseling
	2304	Elementary Education
	2305	Mathematics Teacher Education
	2306	Physical and Health Education Teaching
	2307	Early Childhood Education
	2308	Science and Computer Teacher Education
	2309	Secondary Teacher Education
	2310	Special Needs Education
	2311	Social Science or History Teacher Education
	2312	Teacher Education: Multiple Levels
	2313	Language and Drama Education
	2314	Art and Music Education
	2399	Miscellaneous Education
Linguistics and Foreign Languages	2600	Linguistics and Foreign Languages
	2601	Linguistics and Comparative Language and Literature
	2602	French, German, Latin and Other Common Foreign Language Studies
	2603	Other Foreign Languages
Family and Consumer Sciences	2901	Family and Consumer Sciences
Law	3200	Law
	3201	Court Reporting

	3202	Pre-Law and Legal Studies
English Language, Literature, and Composition	3300	English Language, Literature, and Composition
	3301	English Language and Literature
	3302	Composition and Speech
Liberal Arts and Humanities	3400	Liberal Arts and Humanities
	3401	Liberal Arts
	3402	Humanities
Library Science	3501	Library Science
Interdisciplinary and Multi-Disciplinary Studies (General)	4000	Interdisciplinary and Multi-Disciplinary Studies (General)
	4001	Intercultural and International Studies
	4007	Interdisciplinary Social Sciences
	4008	Multi-disciplinary or General Science
Physical Fitness, Parks, Recreation, and Leisure	4101	Physical Fitness, Parks, Recreation, and Leisure
Philosophy and Religious Studies	4801	Philosophy and Religious Studies
Theology and Religious Vocations	4901	Theology and Religious Vocations
Psychology	5200	Psychology
	5201	Educational Psychology
	5202	Clinical Psychology
	5203	Counseling Psychology
	5299	Miscellaneous Psychology
Criminal Justice and Fire Protection	5301	Criminal Justice and Fire Protection
Public Affairs, Policy, and Social Work	5400	Public Affairs, Policy, and Social Work
	5401	Public Administration
	5402	Public Policy
	5403	Human Services and Community Organization
	5404	Social Work
Social Sciences	5500	General Social Sciences
	5501	Economics
	5502	Anthropology and Archeology
	5503	Criminology
	5504	Geography
	5505	International Relations
	5506	Political Science and Government
	5507	Sociology
	5599	Miscellaneous Social Sciences
Construction Services	5601	Construction Services
Electrical and Mechanic Repairs and Technologies	5701	Electrical and Mechanic Repairs and Technologies

Precision Production and Industrial Arts	5801	Precision Production and Industrial Arts
Fine Arts	6000	Fine Arts
	6001	Drama and Theater Arts
	6002	Music
	6003	Visual and Performing Arts
	6004	Commercial Art and Graphic Design
	6005	Film, Video and Photographic Arts
	6006	Art History and Criticism
	6007	Studio Arts
	6099	Miscellaneous Fine Arts
Medical and Health Sciences and Services	6100	General Medical and Health Services
	6102	Communication Disorders Sciences and Services
	6103	Health and Medical Administrative Services
	6104	Medical Assisting Services
	6105	Medical Technologies Technicians
	6107	Nursing
	6109	Treatment Therapy Professions
	6110	Community and Public Health
	6199	Miscellaneous Health Medical Professions
Business	6200	General Business
	6201	Accounting
	6203	Business Management and Administration
	6204	Operations, Logistics and E-Commerce
	6205	Business Economics
	6206	Marketing and Marketing Research
	6207	Finance
	6209	Human Resources and Personnel Management
	6210	International Business
	6211	Hospitality Management
	6299	Miscellaneous Business and Medical Administration
History	6402	History
	6403	United States History

Table B-1: Observations and frequency of degree fields and occupations by gender

Year	STEM	Number			Frequency		
		Total	Male	Female	Total	Male	Female
Educational segregation							
2018	Non-STEM	248269	112,739	135,530	72.58%	62.71%	83.51%
	STEM	93801	67,034	26,767	27.42%	37.29%	16.49%
2019	Non-STEM	258416	116,956	141,460	72.30%	62.49%	83.07%
	STEM	99019	70,190	28,829	27.70%	37.51%	16.93%
2020	Non-STEM	206180	92,341	113,839	71.71%	61.60%	82.71%
	STEM	81351	57,555	23,796	28.29%	38.40%	17.29%
2018-2020	Non-STEM	712865	322036	390829	72.22%	62.31%	83.12%
	STEM	274171	194779	79392	27.78%	37.69%	16.88%
Occupational segregation							
2018	Non-STEM	253309	128,136	125,173	74.05%	71.28%	77.13%
	STEM	88761	51,637	37,124	25.95%	28.72%	22.87%
2019	Non-STEM	260156	129,898	130,258	72.78%	69.41%	76.49%
	STEM	97279	57,248	40,031	27.22%	30.59%	23.51%
2020	Non-STEM	207887	102,730	105,157	72.30%	68.53%	76.40%
	STEM	79644	47,166	32,478	27.70%	31.47%	23.60%
2018-2020	Non-STEM	712865	360,764	360,588	72.22%	69.81%	76.68%
	STEM	274171	156,051	109,633	27.78%	30.19%	23.32%

Table B-2: Calculation process and results of Duncan segregation index

Year	STEM	$\lvert F_s - M_s \rvert$		$D = 1/2(\lvert F_s - M_s \rvert)$
		$F_s - M_s$	$\lvert F_s - M_s \rvert$	
Educational segregation				
2018	Non-STEM	20.80%	20.80%	0.2080
	STEM	-20.80%	20.80%	
2019	Non-STEM	20.58%	20.58%	0.2058
	STEM	-20.58%	20.58%	
2020	Non-STEM	21.11%	21.11%	0.2111
	STEM	-21.11%	21.11%	
2018-2020	Non-STEM	20.80%	20.80%	0.2080
	STEM	-20.80%	20.80%	
Occupational segregation				
2018	Non-STEM	5.85%	5.85%	0.0585
	STEM	-5.85%	5.85%	
2019	Non-STEM	7.08%	7.08%	0.0708
	STEM	-7.08%	7.08%	

2020	Non-STEM	7.87%	7.87%	0.0787
	STEM	-7.87%	7.87%	
2018-2020	Non-STEM	6.88%	6.88%	0.0688
	STEM	-6.88%	6.88%	

Table C: Gender wage gap (average hourly wage) in STEM/non-STEM fields

	Fields	Male ($)	Female ($)	Gap
Degree field	STEM	55.89	44.03	0.21
	Non-STEM	46.16	34.41	0.25
Occupation	STEM	56.06	43.35	0.23
	Non-STEM	47.13	33.81	0.28

Table D-1-1: Observations and frequency of detailed STEM classifications by gender

Occupations	**Number**			**Frequency**		
	Total	**Male**	**Female**	**Total**	**Male**	**Female**
*Agriculture	5689	3094	2595	2.07%	1.59%	3.27%
*Environment and Natural Resources	5390	3355	2035	1.97%	1.72%	2.56%
*Communication Technologies	1586	1062	524	0.58%	0.55%	0.66%
*Computer and Information Sciences	44332	34430	9902	16.17%	17.68%	12.47%
*Engineering and Engineering Technologies	104722	88681	16041	38.20%	45.53%	20.20%
*Biology and Life Sciences	50339	25077	25262	18.36%	12.87%	31.82%
*Mathematics and Statistics	14154	8689	5465	5.16%	4.46%	6.88%
*Military Technologies	86	80	6	0.03%	0.04%	0.01%
*Interdisciplinary and Multi-Disciplinary Studies (General)	2594	709	1885	0.95%	0.36%	2.37%
*Physical Sciences	29637	19598	10039	10.81%	10.06%	12.64%
*Nuclear, Industrial Radiology, and Biological Technologies	305	143	162	0.11%	0.07%	0.20%
*Psychology	691	232	459	0.25%	0.12%	0.58%
*Transportation Sciences and Technologies	3121	2807	314	1.14%	1.44%	0.40%
*Medical and Health Sciences and Services	5764	2810	2954	2.10%	1.44%	3.72%
*Business	5761	4012	1749	2.10%	2.06%	2.20%

Table D-1-2: Observations and frequency of detailed non-STEM classifications by gender

Occupations	Number			Frequency		
	Total	Male	Female	Total	Male	Female
*Agriculture	5582	4085	1497	0.78%	1.27%	0.38%
*Environment and Natural Resources	2525	1686	839	0.35%	0.52%	0.21%
*Architecture	7333	4911	2422	1.03%	1.52%	0.62%
*Area, Ethnic, and Civilization Studies	3201	1170	2031	0.45%	0.36%	0.52%
*Communications	41750	18051	23699	5.86%	5.61%	6.06%
*Cosmetology Services and Culinary Arts	782	492	290	0.11%	0.15%	0.07%
*Education Administration and Teaching	86907	22970	63937	12.19%	7.13%	16.36%
*Linguistics and Foreign Languages	8465	2927	5538	1.19%	0.91%	1.42%
*Family and Consumer Sciences	6543	749	5794	0.92%	0.23%	1.48%
*Law	1570	536	1034	0.22%	0.17%	0.26%
*English Language, Literature, and Composition	25925	9736	16189	3.64%	3.02%	4.14%
*Liberal Arts and Humanities	11102	4893	6209	1.56%	1.52%	1.59%
*Library Science	347	52	295	0.05%	0.02%	0.08%
*Interdisciplinary and Multi-Disciplinary Studies (General)	5730	2031	3699	0.80%	0.63%	0.95%
*Physical Fitness, Parks, Recreation, and Leisure	11547	6776	4771	1.62%	2.10%	1.22%
*Philosophy and Religious Studies	6406	4660	1746	0.90%	1.45%	0.45%
*Theology and Religious Vocations	6125	4988	1137	0.86%	1.55%	0.29%
*Psychology	44201	13512	30689	6.20%	4.20%	7.85%
*Criminal Justice and Fire Protection	21565	13528	8037	3.03%	4.20%	2.06%
*Public Affairs, Policy, and Social Work	13123	2880	10243	1.84%	0.89%	2.62%
*Social Sciences	72604	41092	31512	10.18%	12.76%	8.06%
*Construction Services	2449	2250	199	0.34%	0.70%	0.05%

*Electrical and Mechanic Repairs and Technologies	351	328	23	0.05%	0.10%	0.01%
*Precision Production and Industrial Arts						
*Fine Arts	34367	15313	19054	4.82%	4.76%	4.88%
*Medical and Health Sciences and Services	60384	11209	49175	8.47%	3.48%	12.58%
*Business	212575	118460	94115	29.82%	36.78%	24.08%
*History	19406	12751	6655	2.72%	3.96%	1.70%

Note: There are no observations in Precision Production and Industrial Arts.

Table D-2-1:Calculation process and results of Duncan segregation index (STEM)

Occupations	**Frequency**			**|Fj − Mj|**	
	Total	**Male**	**Female**	**Fj − Mj**	**|Fj − Mj|**
*Agriculture	2.07%	1.59%	3.27%	1.68%	1.68%
*Environment and Natural Resources	1.97%	1.72%	2.56%	0.84%	0.84%
*Communication Technologies	0.58%	0.55%	0.66%	0.11%	0.11%
*Computer and Information Sciences	16.17%	17.68%	12.47%	-5.20%	5.20%
*Engineering and Engineering Technologies	38.20%	45.53%	20.20%	-25.32%	25.32%
*Biology and Life Sciences	18.36%	12.87%	31.82%	18.94%	18.94%
*Mathematics and Statistics	5.16%	4.46%	6.88%	2.42%	2.42%
*Military Technologies	0.03%	0.04%	0.01%	-0.03%	0.03%
*Interdisciplinary and Multi-Disciplinary Studies (General)	0.95%	0.36%	2.37%	2.01%	2.01%
*Physical Sciences	10.81%	10.06%	12.64%	2.58%	2.58%
*Nuclear, Industrial Radiology, and Biological Technologies	0.11%	0.07%	0.20%	0.13%	0.13%
*Psychology	0.25%	0.12%	0.58%	0.46%	0.46%
*Transportation Sciences and Technologies	1.14%	1.44%	0.40%	-1.05%	1.05%
*Medical and Health Sciences and Services	2.10%	1.44%	3.72%	2.28%	2.28%
*Business	2.10%	2.06%	2.20%	0.14%	0.14%
$\mathbf{D = 1/2(\|Fj - Mj\|) = 0.6322/2 = 0.3161}$					

Table D-2-2: Calculation process and results of Duncan segregation index (non-STEM)

Occupations	Frequency			$\lvert Fj - Mj \rvert$		
	Total	Male	Female	$Fj - Mj$		$\lvert Fj - Mj \rvert$
*Agriculture	0.78%	1.27%	0.38%	-0.89%	0.89%	0.78%
*Environment and Natural Resources	0.35%	0.52%	0.21%	-0.31%	0.31%	0.35%
*Architecture	1.03%	1.52%	0.62%	-0.91%	0.91%	1.03%
*Area, Ethnic, and Civilization Studies	0.45%	0.36%	0.52%	0.16%	0.16%	0.45%
*Communications	5.86%	5.61%	6.06%	0.46%	0.46%	5.86%
*Cosmetology Services and Culinary Arts	0.11%	0.15%	0.07%	-0.08%	0.08%	0.11%
*Education Administration and Teaching	12.19%	7.13%	16.36%	9.23%	9.23%	12.19%
*Linguistics and Foreign Languages	1.19%	0.91%	1.42%	0.51%	0.51%	1.19%
*Family and Consumer Sciences	0.92%	0.23%	1.48%	1.25%	1.25%	0.92%
*Law	0.22%	0.17%	0.26%	0.10%	0.10%	0.22%
*English Language, Literature, and Composition	3.64%	3.02%	4.14%	1.12%	1.12%	3.64%
*Liberal Arts and Humanities	1.56%	1.52%	1.59%	0.07%	0.07%	1.56%
*Library Science	0.05%	0.02%	0.08%	0.06%	0.06%	0.05%
*Interdisciplinary and Multi-Disciplinary Studies (General)	0.80%	0.63%	0.95%	0.32%	0.32%	0.80%
*Physical Fitness, Parks, Recreation, and Leisure	1.62%	2.10%	1.22%	-0.88%	0.88%	1.62%
*Philosophy and Religious Studies	0.90%	1.45%	0.45%	-1.00%	1.00%	0.90%
*Theology and Religious Vocations	0.86%	1.55%	0.29%	-1.26%	1.26%	0.86%
*Psychology	6.20%	4.20%	7.85%	3.66%	3.66%	6.20%
*Criminal Justice and Fire Protection	3.03%	4.20%	2.06%	-2.14%	2.14%	3.03%
*Public Affairs, Policy, and Social Work	1.84%	0.89%	2.62%	1.73%	1.73%	1.84%
*Social Sciences	10.18%	12.76%	8.06%	-4.70%	4.70%	10.18%
*Construction Services	0.34%	0.70%	0.05%	-0.65%	0.65%	0.34%
*Electrical and Mechanic Repairs and Technologies	0.05%	0.10%	0.01%	-0.10%	0.10%	0.05%
*Precision Production and Industrial Arts						
*Fine Arts	4.82%	4.76%	4.88%	0.12%	0.12%	4.82%
*Medical and Health Sciences and Services	8.47%	3.48%	12.58%	9.10%	9.10%	8.47%
*Business	29.82%	36.78%	24.08%	-12.70%	12.70%	29.82%
*History	2.72%	3.96%	1.70%	-2.26%	2.26%	2.72%

$$\mathbf{D = 1/2(\lvert Fj - Mj \rvert) = 0.5573/2 = 0.2787}$$

Table E-1: Income equation regression results (STEM / Non-STEM)

Variables		Male (N=516,815)	Female (N=470,221)
Background and demographic characteristics			
Age Range			
	age36-45	0.019*** (0.004)	0.000 (0.004)
	age46-55	0..030*** (0.007)	‘-0.042*** (0.006)
	age56-65	0.050*** (0.009)	‘-0.022** (0.009)
Race			
	Black	‘-0.231*** (0.004)	‘-0.095*** (0.003)
	Indian	‘-0.254*** (0.015)	‘-0.179*** (0.011)
	Asian	‘-0.020*** (0.005)	0.052*** (0.005)
	Other	‘-0.120*** (0.003)	‘-0.077*** (0.003)
	More-than-one	‘-0.066*** (0.005)	‘-0.028***(0.004)
Region			
	Northeast	0.006** (0.003)	0.009*** (0.003)
	Midwest	‘-0.155*** (0.003)	‘-0.146***(0.003)
	South	‘-0.106*** (0.002)	‘-0.141***(0.002)
Family formation characteristics			
Number of children			
	1-2 Children	0.064*** (0.002)	0.018*** (0.002)
	3-4 Children	0.092*** (0.003)	‘-0.009** (0.003)
	5-6 Children	’-0.004 (0.011)	‘-0.078***(0.016)
	7-8 Children	’-0.036 (0.031)	‘-0.226***(0.058)
	9+ Children	0.009 (0.066)	‘-0.205* (0.124)
Marital status			
	Married	0.197*** (0.003)	0.082*** (0.002)
	Single	0.034*** (0.004)	’-0.011*** (0.003)
Education-related characteristics			
Female/Male ratio in different degree fields			
	Famale ratio	’-0.299*** (0.006)	’-0.352*** (0.005)
Highest educational attainment			
	Master's degree	0.178*** (0.002)	0.178*** (0.002)
	Professional degree	0.476*** (0.004)	0.446*** (0.004)
	Doctoral degree	0.274*** (0.004)	0.325*** (0.004)
Whether the individual has a second degree			
	Degree2	0.045*** (0.003)	0.032*** (0.003)
Whether the individual university bachelor's degree belongs to the STEM field			
	STEM (subjects)	0.038*** (0.002)	0.004* (0.003)

whether respondents attending school were enrolled in a public or a private school (any time in the past 3 months)		
Public school	'-0.179*** (0.005)	'-0.144*** (0.004)
Private school	'-0.150*** (0.007)	'-0.010*** (0.005)
Job-related characteristics		
Whether the occupation of an individual belongs to the STEM field		
STEM (occupation)	0.170*** (0.002)	0.238*** (0.002)
Job sector		
Government	'-0.237*** (0.002)	'-0.151*** (0.002)
Non-profit	'-0.282*** (0.003)	'-0.145*** (0.002)
Working experience	0.041*** (0.001)	0.040*** (0.001)
Working experience²	'-0.001*** (0.000)	'-0.001*** (0.000)
Usual hours worked per week	0.014*** (0.000)	0.014*** (0.000)
Constant	10.172*** (0.008)	10.188*** (0.008)
Obsrvations	516815	470221
Adjusted R^2	0.254	0.222
Wald χ^2 (32)	5502***	4204***

Note: Significant at: *p , 0.1, * *p , 0.05 and ***p , 0.01; robust standard errors are in parentheses.

Table E-2: Income equation regression results (Detailed STEM classification)

Variables		Male (N=516,815)	Female (N=470,221)
	Background and demographic characteristics		
Age Range			
	age36-45	0.019*** (0.004)	‘-0.000 (0.004)
	age46-55	0..029*** (0.007)	‘-0.043*** (0.006)
	age56-65	0.050*** (0.009)	‘-0.023** (0.009)
Race			
	Black	‘-0.232*** (0.004)	‘-0.096*** (0.003)
	Indian	‘-0.254*** (0.015)	‘-0.180*** (0.011)
	Asian	‘-0.024*** (0.005)	0.048*** (0.005)
	Other	‘-0.114*** (0.003)	‘-0.081*** (0.003)
	More-than-one	‘-0.068*** (0.005)	‘-0.028***(0.004)
Region			
	Northeast	0.005* (0.003)	0.008*** (0.003)
	Midwest	‘-0.154*** (0.003)	‘-0.146***(0.003)
	South	‘-0.106*** (0.002)	‘-0.142***(0.002)
	Family formation characteristics		
Number of children			
	1-2 Children	0.063*** (0.002)	0.017*** (0.002)
	3-4 Children	0.091*** (0.003)	‘-0.009** (0.004)
	5-6 Children	’-0.005 (0.011)	‘-0.079***(0.016)
	7-8 Children	’-0.037 (0.031)	‘-0.228***(0.058)
	9+ Children	0.007 (0.066)	‘-0.206* (0.124)
Marital status			
	Married	0.197*** (0.003)	0.081*** (0.002)
	Single	0.034*** (0.004)	’-0.012*** (0.003)
	Education-related characteristics		
STEM subjects by sub-category (including 1st/2nd degree)			
	Computer and Information Sciences	0.068*** (0.004)	0.028*** (0.006)
	Engineering	0.071*** (0.003)	0.093*** (0.005)
	Biology and Life Sciences	0.041*** (0.004)	‘-0.015*** (0.004)
	Mathematics and Statistics	0.097*** (0.006)	0.075*** (0.007)
	Technologies	‘-0.056*** (0.006)	‘-0.061***(0.012)
	Physical Sciences	‘-0.004 (0.005)	‘-0.049***(0.006)
	Medical and Health Sciences and Services	0.140*** (0.012)	0.091*** (0.010)
	Other STEM subjects	‘-0.118*** (0.007)	‘-0.071***(0.007)
Female/Male ratio in different degree fields			
	Famale ratio	’-0.276*** (0.007)	’-0.324*** (0.005)

Highest educational attainment		
Master's degree	0.175*** (0.002)	0.176*** (0.002)
Professional degree	0.476*** (0.004)	0.450*** (0.004)
Doctoral degree	0.275*** (0.004)	0.328*** (0.004)
Whether the individual has a second degree		
Degree2	0.047*** (0.003)	0.035*** (0.003)
whether respondents attending school were enrolled in a public or a private school (any time in the past 3 months)		
Public school	‘-0.178*** (0.005)	‘-0.143*** (0.004)
Private school	‘-0.151*** (0.007)	‘-0.010*** (0.005)
Education-related characteristics		
Whether the occupation of an individual belongs to the STEM field		
STEM (occupation)	0.162*** (0.002)	0.234*** (0.002)
Job sector		
Government	‘-0.235*** (0.002)	‘-0.151*** (0.002)
Non-profit	‘-0.281*** (0.003)	‘-0.144*** (0.002)
Working experience	0.042*** (0.001)	0.041*** (0.001)
Working experience²	‘-0.001*** (0.000)	‘-0.001*** (0.000)
Usual hours worked per week	0.014*** (0.000)	0.014*** (0.000)
Constant	10.160*** (0.008)	10.172*** (0.008)
Obsrvations	516815	470221
Adjusted R^2	0.256	0.224
Wald χ^2 (39)	4557***	3476***

Note: Significant at: *p , 0.1, * *p , 0.05 and ***p , 0.01; robust standard errors are in parentheses.