Gender Divergence in Skill Components and Changes in Gender Wage Gap

Sunha Myong*

December 20, 2023

Abstract

Using data from NLSY79 and NLSY97, I find that women experience positive wage gains at all levels of the wage distribution, while men mostly experience negative wage gains, except at the top end of the wage distribution. I analyze how changes in various skill components contribute to the recent shifts in the wage distribution by gender and examine the implications for trends in the gender wage gap. The wage gain from changes in skill components is highest in the middle range of the wage distribution for women and at the top for men. Accordingly, changes in skill components can explain three-quarters of the shift in the gender wage gap in the middle range of the wage distribution, but account for only one third of the trend at the top end. Skill prices declined more for men in the middle wage range, accounting for one-quarter of the gender wage gap reduction. Meanwhile, the increase in skill prices for women, combined with stable prices for men in two different cohorts, accounts for two-thirds of the reduction in the gender wage gap at the top end of the distribution.

*School of Economics, Singapore Management University; E-mail: sunhamyong@smu.edu.sg.

1 Introduction

In this paper, I document how changes in various skill components explain the recent evolution of the wage distribution by gender, and examine the implications for the trend in the gender wage gap. I use data from the National Longitudinal Study of Youth 79 (NLSY79) and the National Longitudinal Study of Youth 97 (NLSY97) to construct comparable samples for the 1960s and the 1980s birth cohorts in the US, respectively.[1] Figure 1 provides motivation to investigate gender-specific trends, highlighting notable gender differences in the evolution of wage distribution across two cohorts. Women experience positive wage gains at all levels of the wage distribution (dashed line), while men mostly experience negative wage gains, except at the top end of the wage distribution (solid line).[2] On the other hand, log wage gains peak in the middle for women but monotonically increase across the wage distribution for men (Panel (a)). Investigating factors explaining the aforementioned gender divergence in wage gain across two cohorts can help to understand the evolution of the gender wage gap in recent years.

The NLSY79 and NLSY97 datasets offer a unique opportunity to explore the evolution of wage distribution for a representative sample of the U.S., providing comparable data for cohorts from the 1960s and 1980s in terms of various skill components, including family structure, cognitive skill, degree attainment, and early career experiences, as well as wage rates. I build upon the analysis conducted by Altonji et al. (2012), who investigates to what extent recent changes in the wage inequality can be explained by changes in skill components of youth population. The novel approach employed in Altonji et al. (2012) is to estimate the counterfactual wage distribution for the NLSY97 cohorts based on the observed wage distribution for the NLSY79 cohort and observed changes in skill components for both NLSY samples, without observing the actual wage distribution for the NLSY97 cohorts. This paper presents three key differences. First, I document the actual wage distribution for the 1980s birth cohorts by using recent data from the NLSY97.

[1]The data for the 1957−1964 birth cohorts are sourced from the National Longitudinal Study of Youth 1979 (NLSY79), while the data for the 1980−1984 birth cohorts are obtained from the National Longitudinal Study of Youth 1997 (NLSY97).

[2]Section 3 explains how the wage distribution is estimated by gender and cohort.

With the actual wage distribution data for the 1980s cohorts in hand, I can precisely document the evolution of the wage distribution at each percentile across two cohorts by gender, and their implications for the gender wage gap. Second, I analyze gender-specific changes in skill components across these two cohorts and estimate the respective roles of different types of skill components in women's relative wage gains across the two cohorts. Third, I conduct decomposition exercises to quantify the extent to which the actual changes in the gender wage gap can be attributed to gender-specific changes in skill components and gender-specific changes in skill prices.

Comparing skill components across two cohorts by gender reveals intriguing patterns of gender divergence. First, on average, men and women have similar AFQT scores in the NLSY79 cohort; but in the NLSY97 cohort, women's AFQT scores are 0.1 standard deviation higher than men's. Substantial gains in cognitive skills among women, relative to men, contrast with other characteristics measured before college age, such as parents' education and family structure, which do not exhibit notable gender differences in their trends. Similarly, men and women have similar educational attainment in the NLSY79 cohort, but for the NLSY97 cohort, the share of women with a four-year college degree as of age 30 is 14 percentage points higher than that of men. I also find that while labor market experience decreased for both men and women as they attained more education in their early 20s, the gender gap in labor market experience, measured by employment and average working hours, decreased across the two cohorts.

Motivated by gender-specific changes in the wage distribution and skill components across two cohorts, I propose a decomposition method to clearly demonstrate the role of gender-specific changes in skill components and prices in explaining the trends in the gender wage gap. The key feature of the proposed decomposition method for the gender wage gap is to focus on the gender specific progress in skill components and skill prices over time. This approach complements the previous decomposition method that focuses on how the cross sectional gender differences in skill components and skill prices change over time, which is agnostic about whether the narrowing gender gap over time is associated with positive skill gains for both men and women, gains for women and losses for men, or losses for both genders.

To quantify the degree to which changes in skill components can explain the wage gain by gender across two cohorts, I apply the DiNardo, Fortin, and Lemieux (DFL) decomposition method (DiNardo et al. (1996)) to obtain the gender-specific counterfactual wage distribution of the NLSY79 cohort when the distribution of skill components is replaced with that of the NLSY97 cohort for each gender.[3] Having observed the actual changes in the wage distribution for each gender across two cohorts in the data, I can quantify the relative importance of the following four components—men's wage gains explained by changes in skill components, women's wage gains explained by changes in skill components, men's wage gains explained by changes in skill prices, and women's wage gains explained by changes skill prices—in explaining the observed trends. The gender difference in the first two components—wage gain explained by changes in skill components—measures to what extent women's relative progress in skill acquisition across two cohorts contributed to the decrease in the gender wage gap. On the other hand, the gender difference in the last two components—wage gain explained by changes in skill prices—measures to what extent changes in women's relative progress in skill prices across two cohorts contributed to the decrease in the gender wage gap. I also estimate the marginal effects of specific skill components to quantify the relative importance of pre-college skills, education, and work experience in the above analysis.

I find that wage gains attributed to changes in skill components are positive for both men and women across all wage levels. However, gender-specific patterns emerge. Women exhibit the highest log wage gain attributed to the skill change in the mid-range of the distribution, while men's log wage gains monotonically increase along the distribution. Consequently, women's wage gains relative to men's explained by changes in skill components are greatest in the middle and smallest at the top end of the wage distribution. I also find that men's wage gain associated with changes in skill components are almost entirely attributed to changes in pre-college skill components while changes in educational attainment are also significant contributors to wage gains for women across two cohorts. Changes in work experience conditional on pre-college and college periods skill

[3] Altonji et al. (2012) employs the same approach to estimate the counterfactual wages for the population, including both men and women, in the NLSY97 cohorts, accounting for changes in skill components across two cohorts.

components have negligible marginal effects on the wage gain across two cohorts for men, but it has negative impacts on women.

The contribution to wage gains varies significantly across different skill components. First, parents' education plays a significant role in explaining the positive wage gain across two cohorts for both men and women, as well as in shaping the overall wage gain across various wage distribution levels for both genders. However, changes in family structure and AFQT scores contribute to explaining why women's wage gain across two cohorts is greater than that of men. On the other hand, after controlling for changes in pre-college skill components, the marginal effects of changes in educational attainment on wage gains across two cohorts are substantially greater for women than for men. In particular, for men, the marginal effect of educational attainment, as measured by highest grade completed and degree attainment, is negligible compared to the wage gain explained by pre-college skill components. For women, changes in educational attainment explain additional positive wage gains after accounting for changes in pre-college skill components, and the marginal effects of pre-college skills and educational attainment on wage gains are similar. Additionally controlling for the major field of study has negligible marginal effect on the wage gain across two cohorts for both men and women. Finally, the marginal effect of labor market experience, as measured by full-time or part-time employment periods during age 22-34, is small for both men and women, whereas changes in occupation across two cohorts has relatively greater negative impacts on women's wage gain. This finding aligns with previous studies Blau and Kahn (2017), suggesting that the persistent occupational segregation by gender is an important factor in explaining the gender wage gap.

Consistent with the findings for the gender-specific wage gain explained by skill changes, the decomposition results implies that changes in skill components can explain three-quarters of the shift in the gender wage gap within the middle range of the wage distribution, but account for only one third of the observed trend at the top end. The skill prices decrease in the middle range of the wage distribution for both men and women, with a larger decline for men, contributing to the reduction of the gender wage gap. Conversely, the decline in the gender wage gap at both ends of the wage distribution is primarily attributed to a significant increase in skill prices for women, while men's skill prices remain relatively stable. This indicates that alterations in the gender wage gap at the both ends of the wage distribution are primarily attributed to gender-specific changes in skill prices across the two cohorts. These findings are consistent with the literature (Galor et al., 1996; Welch, 2000; Weinberg, 2000; Black and Spitz-Oener, 2010; Bacolod and Blum, 2010; Cortes et al., 2020), which documents that changes in skill prices play an important role in explaining the decrease in the gender wage gap at the top of the wage distribution.

The decomposition exercise also offers insights into how gender-specific progress in skill acquisition and skill prices across two cohorts contributes to explaining the gender wage gap. For instance, at the 50th percentile of the wage distribution, if men's skill level did not change across two cohorts, women's skill gain could explain 104.2% of the observed decrease in the gender wage gap. However, men's skill also improved across two cohorts, which contributed to a 27.1% increase in the observed change in the gender wage gap. Together, the net effect of changes in skill components in explaining the decrease in the gender wage gap is 77.1% at the median. On the other hand, at the 90th percentile of the wage distribution, women's skill gain would explain 85.2% of the observed decrease in the gender wage gap, whereas men's skill gain counteracts its impacts by 51.2%, leading to the net effect of skill changes in explaining the observed gender wage gap to be 33.9%. Thus, at the higher end of the wage distribution, despite significant advancements in women's skill gains, men also make substantial progress in skill acquisition across two cohorts. This leads to women achieving a relatively smaller catch-up through skill accumulation.

When comparing the roles of pre-college skills, college education, and work experience in explaining the observed gender wage gap, I find that both pre-college skills and college education significantly contribute to reducing the gender wage gap. However, the inclusion of work experience measures has a negative marginal effect on decreasing the gender wage gap after accounting for pre-college and college-period skill components. For instance, when focusing on the 50th percentile of the wage distribution, pre-college skill components account for 59.1% of the observed change in the gender wage gap. This estimate increases to 86.2% when I additionally include educational attainment to quantify the role of skills but decreases to 77.1% if I further include work experience measures.

The impacts of pre-college skills, college education, and work-experience on the gender wage gap also differ substantially across the level of the wage distribution. For instance, parents' education and family structure play a significant role in reducing the gender wage gap in the middle range of the distribution, reducing the log gender wage gap by 0.054 log point at the 50th percentile of the wage distribution. However, its impacts on the gender wage gap at the top end of the wage distribution is negligible. The marginal effect of the changes in the AFQT score on the gender wage gap is about 0.02 log point decrease between 50th and 90th percentile of the wage distribution, but its impact on the gender wage gap is negligible at the top end of the wage distribution. Changes in labor market experience, conditional on skills formed prior to the working period, have a negligible impact on the gender wage gap below the 70th percentile of the wage distribution but widen the gender wage gap by 0.03 log points at the 90th percentile. Conversely, the effect of occupation on the gender wage gap is most pronounced at the lower end of the wage distribution.

This paper is related to the literature on the trend in the gender wage gap and the relative roles of skill components and skill prices in narrowing the gender wage gap (Oaxaca, 1973; Blinder, 1973; Blau and Kahn, 2017, 1997; O'Neill and Polachek, 1993; O'Neill, 2003; Kassenboehmer and Sinning, 2014; Ñopo, 2008). The contributions of this paper are threefold. First, I document changes in the gender wage gap for recent cohorts (born in the 1960s and 1980s) using two widely used nationally representative samples in the US: NLSY79 and NLSY97. Focusing on two recent cohorts to understand the trend in the gender wage gap complements previous studies that have primarily examined wage distribution among the working-age population. For instance, contrary to earlier findings (Blau and Kahn, 2017; Kassenboehmer and Sinning, 2014; O'Neill, 2003; Blau and Kahn, 2006) that showed a relatively slower decrease in the gender wage gap at the top end of the wage distribution among the working-age population, the decrease in the log gender wage gap is similar across different wage levels between the 1960s and 1980s birth cohorts, and the extent of the decrease in the gender wage gap across two time periods is greater among youth population.

Examining the gender-specific wage evolution for individuals born in the 1960s and 1980s reveals that men experienced significant wage losses in the middle range of the distribution, which differed from the wage evolution observed in the working-age population (Kassenboehmer and Sinning (2014)). This phenomenon contributes to explaining half of the reduction in the gender wage gap at the median. At the upper end of the wage distribution, both men and women experienced substantial wage increases, with women's progress outpacing men's, leading to a decrease in the gender wage gap.

Second, utilizing the rich set of skill components available in the NLSY surveys, I quantify the diverse roles of each skill component, including parents' education, family structure, cognitive skills, as well as education and work experience variables, in explaining gender-specific wage gains across two cohorts and their implications for the gender wage gap. Doing so provides a detailed account of which stage of skill formation and which types of skills have evolved differently by gender across two cohorts, and its implications for the gender wage gap. For instance, women's relative increase in educational attainment over time has been considered a key factor in explaining the rapid decrease in the gender wage gap (Wellington, 1993; O'Neill and Polachek, 1993; Blau and Kahn, 1997; Boeri et al., 2005). Utilizing NLSY data, I can assess the degree to which shifts in women's relative educational achievements can be linked to changes in several pre-college factors, including parental education, family structure, and cognitive abilities. Additionally, I can quantify the progress made by women in educational attainment within the NLSY97 cohort when compared to women possessing similar pre-college skill components within the NLSY79 cohort. I can also conduct this analysis for men. Understanding this relationship may provide deeper insights into the role of various skill components in shaping the evolution of the gender wage gap.

Third, this paper is related to the extensive literature that examines various decomposition methods used to understand the wage distribution (Oaxaca, 1973; Blinder, 1973; Juhn et al., 1993; DiNardo et al., 1996; Gosling et al., 2000; Machado and Mata, 2005; Fortin et al., 2011; Kassenboehmer and Sinning, 2014). I propose a decomposition method that can clearly break down the roles of gender-specific progress in skill components and skill prices in understanding the trend in the gender wage gap. A key distinction from previous studies (Blau and Kahn, 2017; Kassenboehmer and Sinning, 2014) is that I decompose the wage gain across two cohorts for men and women separately into two parts: one explained by gender-specific changes in skill components and the other by gender-specific changes in skill prices. Afterward, I examine how these components relate to the change in the gender wage gap over time. This approach does not require estimating the reference skill price in each period to measure the extent to which men's and women's skill prices differ from the reference level. Accordingly, the proposed decomposition method offers two ways to analyze the trend: first, by decomposing to discuss the roles of skill components and skill prices, and second, by examining the progress of women and and the progress of men across two cohorts. While the first approach has been extensively discussed in the literature, the second approach provides new insights. For instance, I find that although changes in women's skill components alone could explain most of the observed decrease in the gender wage gap in the middle, the decrease in women's skill prices in the middle offset about half of the progress driven by skill gains. On the other hand, the decrease in men's skill prices in the middle range of the wage distribution alone could account for three-quarters of the observed decrease in the gender wage gap. However, men's skill gains in the middle offset these effects by roughly half. In summary, women's wage gains, primarily explained by increased skill, and men's wage losses, mainly attributed to price decreases, have equally contributed to the decrease in the gender wage gap in the middle range of the wage distribution.

The paper proceeds as follows. Section 2 presents the data and a sample of the NLSY79 and NLSY97 surveys. Section 3 describes the trends in the gender-specific wage distribution and their implications for the gender wage gap. Section 4 discusses the decomposition methods. Section 5 explains the DFL method for estimation. Section 6 presents the estimation results and the decomposition of the trend in the gender wage gap. Section 7 concludes.

2 Data

The NLSY79 and NLSY97 are longitudinal data that follow population-representative samples of American youth. The NLSY79 consists of 12,686 individuals who were born between 1957 and 1964 and were 14–22 years old in 1979, while the NLSY97 consists of 8,984 individuals who were born between 1980 and 1984 and were 12–18 years old in 1997.

To document the demographic characteristics, educational attainment, and school-to-work transition of each NLSY cohort, I use the 1979–2004 survey data for the NLSY79 cohort and the 1997–2017 survey data for the NLSY97 cohort. To ensure comparability with the literature, I construct the sample closely following Altonji et al. (2012). For instance, to construct a comparable measure of cognitive ability across the two cohorts, I standardize the AFQT score based on the mapping provided by (Segall, 1997). I extend the analysis in Altonji et al. (2012) by incorporating measures of work experience and occupation from the more recent NLSY97 survey, which provides information on labor market outcomes at around age 35. Because the NLSY97 does not include a military sample or supplemental samples on poor whites, I exclude those supplemental samples from the NLSY79, which reduces the number of individuals in the NLSY79 from 12,686 to 9,763. I keep only individuals whose reported race is White, Black, or Hispanic, which reduces the sample size of the NLSY97 from 8,984 to 8,901 but does not affect the NLSY79 sample. I adjust the weight for migrants depending on the age of arrival by putting zero weight on those who arrived in the U.S. after age 16. Dropping those individuals reduces the sample size of the NLSY79 to 9,058 but does not affect the sample size of the NLSY97. I restrict the sample to individuals who have non-missing variables for demographic characteristics and educational attainment at age 22, which reduces the sample size to 8,649 for the NLSY79 and 7,718 for the NLSY97. I further drop individuals with missing AFQT scores, which leaves 8,305 and 6,156 individuals in the final sample for the NLSY79 and NLSY97, respectively.[4]

I use the cross-sectional sample weight for the original sample (in the first survey for each

[4] The slight difference in sample size from Altonji et al. (2012) is driven by an update in the raw data in the survey.

NLSY cohort) and update the weight to account for the attrition rate by age 22 and missing AFQT score. In doing so, I estimate the attrition rate and the probability of having a missing AFQT score based on a rich set of individual characteristics in the first survey.[5] The weight is divided by the predicted probability of attrition or missing AFQT score.

For the wage data, I have utilized the 1988, 1990, 1992, and 1994 surveys for the NLSY79 and the 2011, 2013, 2015, and 2017 surveys for the NLSY97. Wages have been standardized through regression to the year 2002, considering individuals with 13 years of experience.[6] The choice of 13 years of experience is primarily aimed at closely aligning with the average years of work experience after schooling for the NLSY97 cohort, which consists of individuals aged 32 to 38 in the 2017 survey.[7]

Table 1 documents gender-specific demographic characteristics for the NLSY79 and the NLSY97 cohorts. First, as widely documented, the educational attainment of both mothers and fathers increased across the two cohorts for both men and women, while the share of US-born parents decreased for both genders across the two cohorts. Second, the AFQT score exhibits significant gender divergence across two cohorts. The AFQT score decreased from 0.008 to -0.037 for men, whereas it increased from 0.023 to 0.060 for female.[8] For the NLSY79 cohort, there was little gender difference in the average AFQT score, but for the NLSY97 cohort, women have almost 0.1 standard deviation higher mean AFQT score than men. Second, family structures changed substantially across the cohorts, with the share of individuals living with both parents decreasing from 76% and 74% to 55% and 52% for men and women, respectively, across the two cohorts.

[5]The list of variables used to estimate the attrition rate and the probability of having a missing AFQT score are as follows: race, gender, highest grade completed by their mother and father, whether the individual lived with their mother/father/both at age 14, whether the individual lived in an urban/standard metropolitan statistical area (SMSA), the attitude displayed during the interview, whether the 1997 interview was done in 1998 instead. The same weights are used in (Altonji et al., 2012).

[6]I calculate potential experience by subtracting the highest grade completed from age minus 6. Then, I run separate OLS regressions for each educational group (classifying individuals into five groups with different levels of education) with log hourly wage as the dependent variable, and experience, its square, and triple squares as independent variables. Once I obtain the coefficients from these OLS regressions, I use them to compute the wage rate for each individual by substituting experience with 13.

[7]Altonji et al. (2012) standardized the wage of the NLSY79 cohorts to 23 years of experience to study the wage distribution of workers at their prime wage; however, they did not use wage data for the NLSY97 cohorts.

[8]The AFQT score is standardized to have zero mean and standard deviation of 1 for the NLSY79 cohort.

Panel A of Table 2 documents gender-specific changes in degree attainment as of age 30, major field of study, and work experiences for the NLSY79 and the NLSY97 cohorts. First, degree attainment increased substantially for both men and women, with a greater extent of increase observed for women. For instance, the share of individuals who received a Bachelor's degree by age 30 increased from 0.26 to 0.28 for men, whereas it increased from 0.25 to 0.39 for women. Similarly, the share of individuals who received a Master's degree by age 30 increased modestly from 5% to 6% for men, whereas it increased from 4% to 10% for women. While women gained substantially more education than men across two cohorts, the gender differences in major choice remained relatively stable across two cohorts (Panel B of Table 2). In the NLSY79 cohorts, men's share with an applied STEM major (0.123) is significantly higher than women's (0.056). However, in the NLSY97 cohorts, the gap widened, as the share for men remained similar at 0.118, while it decreased by half for women to 0.025. While the overall decrease in the share of applied STEM majors is consistent with the findings in Deming (2017), the above finding shows that the decrease in STEM related career across two cohorts is pronounced among female population.

Panel A of Table 3 documents variables summarizing the labor market experience of individuals during ages 22-34 by gender across two cohorts. The labor market experience is measured by the number of years an individual worked positive hours and full-time (more than 1600 annual hours) between ages 22 and 34. The occupation is measured by dummy variables that takes a value of 1 if the number of years the individual worked in a certain occupation is greater than 5 during ages 22-34. The gender difference in labor market experience narrows on various fronts across the two cohorts. For instance, the number of years employed is 9.57 for men and 8.88 for women in the NLSY79 cohort, whereas it becomes 7.95 for men and 8.01 for women in the NLSY97 cohort. Given that women have a higher level of education than men in the recent cohort, the decreasing gender gap in labor market experience suggests that increasing schooling for women did not necessarily widen the gender gap in labor market experience. Similar trends are found for full-time work experience as well as annual working hours. Panel B of Table 3 presents the occupation choices made by men and women in each cohort. To summarize occupation-specific experience between the ages of 22 and 34, a dummy variable is created. This variable takes a value of 1 if the individual worked in a particular occupation for more than 5 years during the age range of 22 to 34. I find that the gender-specific distribution of occupations remains relatively stable across two cohorts, with the exception of women in clerical occupations, where the share decreased from 0.25 to 0.14, and men in operative occupations, where the share decreased from 0.31 to 0.18. Additionally, there was a slight increase in the share of women with at least five years of experience in education and social occupations.

3 Trends in the Gender Specific Wage Distribution and Its Implications

In this section, I document the unadjusted wage gain for each gender across the two cohorts at various percentiles of the wage distribution. I will also discuss its implications for changes in the gender wage gap and wage inequality within each gender.

3.1 Trends in Gender Specific Wage Gain Across Two Cohorts

First, I estimate the gender-specific wage distribution for the NLSY79 and the NLSY97 cohorts. I use the standardized hourly wage for 13 years of potential work experience, which may not be directly comparable to the wage distribution for the working-age population. However, understanding the trend among a relatively young working population complements previous studies because the skill accumulation of the young working population could be more affected by recent trends in the labor market. Following Altonji et al. (2012), I multiply the reciprocal of the frequency of valid wage observations and the sample weight to account for heterogeneous selection into positive wage earnings. Therefore, the reported wage distribution is the wage distribution of the population, if all individuals participate in the labor force.

Panel (a) of Figure 1 displays the difference in the log wage rate between the NLSY97 and

NLSY79 cohorts at each percentile of the wage distribution for men (solid line) and women (dashed line). The point estimates for the wage gain across two cohorts by gender, as well as the statistical significance of the log wage difference between the two cohorts, are presented in column (1) of Table 5. First, men experienced negative wage gains across the two cohorts, except at the top end of the wage distribution. Men's log wage decreased by 0.015, 0.072, 0.061, and 0.031 at the 10th, 25th, 50th, and 75th percentiles of the wage distribution, while it increased by 0.080 at the 90th percentile of the wage distribution. In contrast, women's log wage increased at all levels of the wage distribution across the two cohorts. It increased by 0.098, 0.040, 0.070, 0.098, 0.206 log points at the 5th, 10th, 25th, 50th, 75th, and 90th percentiles of the wage distribution. Therefore, the findings suggest a significant gender divergence in wage gains across the 1960s and 1980s birth cohorts. Note that although the wage gain is negative for men and positive for women except at the top end, both men and women have positive and large wage gains at the top end of the wage distribution.

The log wage gain quantifies the relative increase in hourly wages compared to the baseline level. Considering that wage rates tend to be lower for women than for men, it is informative to examine the wage gains in absolute dollars. Not surprisingly, the direction of the change at each percentile is same for Panels (a) and (b) in Figure 1. However, despite the substantial gender wage gap in the NLSY79 cohort, women's wage gains in absolute dollars across two cohorts exceed those of men as well.

The significant gender difference in the evolution of the wage distribution across two cohorts has important implications for the trend in the gender wage gap. In the following sections, I will document the trends and discuss the relevance of the gender-specific wage gains across the two cohorts.

3.2 Trend in the Gender Wage Gap

I document the gender wage gap for each NLSY cohort and its change along the wage distribution. The solid line with solid circles and the solid line with plus mark in Panel (a) of Figure 2 depict the gender wage gap for the NLSY79 cohort and that for the NLSY97 cohort, respectively, as measured by the log wage difference between men and women at each percentile of the gender specific wage distribution. The log wage difference between men and women for each cohort, along with the corresponding statistical significance, is presented in columns (1) and (2) of Table 6.

The gender wage gap is estimated to be 0.22–0.35 log point for the NLSY79 cohort and 0.09–0.19 log point for the NLSY97 cohort. The female/male wage ratios are 76.1, 74.0, and 76.5 percent for the NLSY79 cohort at 10th, 50th, and 90th percentile of the wage distribution, and the corresponding numbers for the NLSY97 cohorts are 87.4, 87.1, and 89.2 percent. The decrease in the gender wage gap across two cohorts is greater than the trend reported for the working-age population. For instance, focusing on full-time workers of age 25–64 in the Panel Study of Income and Dynamics (PSID) Blau and Kahn (2017) documents that the female/male wage ratios in year 1998 were 80.3, 79.8, and 73.8 percent at 10th, 50th, and 90th percentile of the wage distribution. The corresponding findings in 2010 are 81.5, 82.4, and 73.9 percent at 10th, 50th, and 90th percentiles of the wage distribution. Therefore, the gender wage gap decreases more rapidly in the NLSY samples than in the PSID sample. Also, unlike relatively slower decreases in the gender wage gap over time at the top end of the distribution found in PSID for working-age poulation, the gender wage gap from the NLSY samples decreases similarly at different percentiles of the wage distribution. Given that the NLSY97 sample consists of younger cohorts (born between 1980-1984) than the sample in Blau and Kahn (2017) (born between 1946-1985), the different patterns between two data sets could be explained by smaller gender wages gap for younger cohorts.

Panel (b) of Figure 2 describes the gender wage gap in absolute dollar terms, $(w_m - w_f)$. First, the gender wage gap in absolute dollar differences increases monotonically along the wage distribution for the NLSY79 cohort (represented by the solid line with filled circles in panel (b)). At the 10th, 25th, 50th, 75th percentiles of the wage distribution, the gap is 1.5, 2.5, 3.3, and 4.1 dollars, respectively. The gender wage gap in absolute dollars increases drastically at the top end of the wage distribution, increasing from 4.1 dollars at the 75th percentile of the distribution to 5.9 dollars at the 90th percentile of the distribution. Second, focusing on the NLSY97 cohort, the gender wage gap in absolute dollar term is greatest at the top end, for instance, 3.1 dollars at the top 10 percentile of the wage distribution. However, unlike the NLSY79 cohort, the gender wage gap in absolute dollar term did not increase between 30th and 75th percentiles of the wage distribution at around 1.7 dollars. This finding is consistent with the results in panel (a), where the gender wage gap in relative term decreases significantly in the middle range of the wage distribution.

It is worth noting that the decreased gender wage gap across 1960s and 1980s birth cohorts except at the top end of the wage distribution is driven by wage loss for men and wage gain for women. In other words, except at the top end of the wage distribution, the gender wage gap decreased because the wage gain is negative for men and positive for women. On the other hand, at the top end of the wage distribution, the wage gains across two cohorts are positive for both men and women, while the magnitude of the increase is greater for women.

4 Decomposition of the Trends in the Gender Wage Gap

In this section, I investigate further on the trends in the gender wage gap in relation to the gender-specific changes in skill components and skill prices. I first discuss decomposition exercise widely used in previous studies. Equation (1) illustrates the Oaxaca-Blinder decomposition that decompose the wage difference between men and women into (i) the wage gap explained by different characteristics and (ii) the wage gap explained by different skill prices by gender. w_{jt}, X_{jt}, and b_{jt} for $j \in \{m, f\}$ indicate the wage rate, the characteristic, and the skill price, respectively, for each gender $j \in \{m, f\}$ in period t. Let $\overline{X}_{jt}$ is the mean characteristic of gender $j \in \{m, f\}$ in period t. Part (ii) is interpreted as unexplained differential and corresponds to the average female residual from the male wage equation and the exponential of this unexplained differential is used to demonstrate the female-to-male wage ratio, controlling for specific observable characteristics.

$$ln(w_{mt}) - ln(w_{ft}) = \underbrace{b_{mt}(\overline{X}_{mt} - \overline{X}_{ft})}_{\text{(i) gap explained by different characteristics}} - \underbrace{\overline{X}_{ft}(b_{mt} - b_{ft})}_{\text{(ii) gap explained by different skill prices}} \tag{1}$$

Note that previous studies propose various ways to measure the gender wage gap explained by gender differences in skill prices, which is denoted as part (ii) in the above equation.[9] In particular, $(b_{mt} - b_{ft})$ can be further decomposed into $(b_{mt} - b_t^*)$ and $(b_t^* - b_{ft})$ to measure to what extent men's and women's skill prices are different from the reference price denoted as b_t^*. Extending the Oaxaca-Blinder decomposition to the trend in the gender wage gap, Juhn et al. (1993) and Blau and Kahn (1997) provide the following decomposition:

$$\begin{aligned}&(lnw_{1m} - lnw_{1f}) - (lnw_{0m} - lnw_{0f})\\ &= \underbrace{\left[\left(\Delta\overline{X}_1 - \Delta\overline{X}_0\right)b_{1m}\right]}_{\text{(i) Effect of Changing Characteristics}} + \underbrace{\left[\Delta\overline{X}_0(b_{1m} - b_{0m})\right]}_{\text{(ii) Effect of Changing Skill Prices for Men}} + \underbrace{\left[\overline{X}_{1f}(b_{1m} - b_{1f}) - \overline{X}_{0f}(b_{0m} - b_{0f})\right]}_{\text{(iii) Effect of Changes in Unexplained Gaps}}\end{aligned} \tag{2}$$

The Δ_t prefix denotes the male-female difference for a variable in period $t \in 0,1$. In this specification, part (i) of equation (2), the effect of changing means, measures to what extent changes in gender gap in characteristics over time contributed to the change in the gender wage gap when its impacts on the gender wage gap is evaluated with the skill price of men in year 1 (b_{1m}). Part (ii) of equation (2), the effect of changing skill prices, measures to what extent changes in skill price of men between year 0 and year 1 contributed to the change in the gender wage gap when the male-female characteristic differences remains the same as in year 0. Therefore, part (ii) specifically focuses on how shifts in wage structures for men over time impact the gender wage gap. One

[9]Recent studies estimating the reference price include (Fortin, 2008; Jann, 2008; Elder et al., 2010; Kassenboehmer and Sinning, 2014).

thing to note is that if the skill price changed differently for men and women for a particular skill component, part (ii) might understate or overstate the impacts of price changes on the gender wage gap. Part (iii) assesses the impact of changes in the unexplained gap, which is measured by the shift in the gender wage gap attributed to different skill prices for men and women with comparable characteristics over time. Part (iii) is often interpreted as changes in the unexplained gender wage gap in previous studies. It remains unclear whether this change can be attributed primarily to disparities between X_{f0} and X_{f1} or if it is primarily driven by variations between $(b_{m1} - b_{f1})$ and $(b_{m0} - b_{f0})$.

In the following equation (3), I provide an alternative decomposition method that allows to focus on gender specific progress across cohorts in terms of skill acquisition and skill prices. The key difference from the method above is that I initially decompose the gender-specific wage gain across two cohorts into that explained by changes in skill components and that explained by changes in skill prices. Then, I relate this gender-specific change to account for the changes in the gender wage gap. In particular, the change in the gender wage gap can be attributed to the following four components—men's wage gains explained by changes in skill components, women's wage gains explained by changes in skill components, men's wage gains explained by changes in skill prices, and women's wage gains explained by changes skill prices—in explaining the observed trends as follows:

$$
\begin{aligned}
&(lnw_{1m} - lnw_{1f}) - (lnw_{0m} - lnw_{0f}) \\
&= \underbrace{\left[(\overline{X}_{1m} - \overline{X}_{0m})b_{0m}\right]}_{\text{(i) Effect of Changing Characteristics for Men}} - \underbrace{\left[(\overline{X}_{1f} - \overline{X}_{0f})b_{0f}\right]}_{\text{(ii) Effect of Changing Characteristics for Women}} \\
&+ \underbrace{\left[\overline{X}_{1m}(b_{1m} - b_{0m})\right]}_{\text{(iii) Effect of Changing Coefficients for Men}} - \underbrace{\left[\overline{X}_{1f}(b_{1f} - b_{0f})\right]}_{\text{(iv) Effect of Changing Coefficients for Women}}
\end{aligned}
\tag{3}
$$

Part (i) in equation (3) calculates the wage gain for men between period 0 and 1 that can be attributed to changes in men's characteristics over time when using the skill price for men from period 0. Similarly, part (ii) in equation (3) represents the wage gain for women between period 0 and 1 that can be attributed to changes in women's characteristics over time, while using the skill price for women from period 0. Therefore, both part (i) and part (ii) provide counterfactual wage gains for each gender, which can be explained solely by the changes in their characteristics over time, assuming the wage structure for each gender remains the same as in period 0. The difference between part (i) and part (ii) in equation (3) quantifies the extent to which the gender wage difference has changed due to alterations in gender-specific characteristics while keeping the gender-specific skill prices constant at their period 0 levels.

Part (iii) in equation (3) calculates the wage gain for men across two cohorts that is explained by changes in skill prices for men between periods 0 and 1 when men's characteristics are kept at the period 1 levels. Similarly, part (iv) in equation (3) calculates the wage gain for women across two cohorts that is explained by changes in skill prices for women between period 0 and 1 when women's characteristics are kept at the period 1 levels. Therefore, both part (iii) and part (iv) provide counterfactual wage gains for each gender, which can be explained solely by the changes in the gender-specific skill prices over time, assuming the characteristics for each gender remains the same as in period 1. Consequently, the difference between part (iii) and part (iv) quantifies the extent to which changes in gender-specific skill prices across two periods can explain the change in the gender wage gap, while holding the characteristics of each gender at their period 0 levels.

The advantage of the proposed decomposition method is that it enables us to focus on the relative progress of men and women across two periods in terms of skill components and skill prices. As the gender-specific wage gain is decomposed into gender-specific progress in skill components and skill prices, this approach does not require estimating the reference skill price in each period to measure the extent to which men's and women's skill prices differ from the reference level.

Also, the proposed decomposition method offers two ways to analyze the trend: first, by decomposing to discuss the roles of skill components and skill prices, and second, by examining the progress of women and and the progress of men across two cohorts. While the first approach has been extensively discussed in the literature, the second approach has not been employed. Using gender-specific skill prices to evaluate the impact of price changes on gender-specific wage gains makes it straightforward to discuss the respective roles of skill prices and skill components in explaining the gender-specific changes in wages across two cohorts, because this discussion does not depend on the estimation of reference skill prices. The second approach could yield meaningful insights, including whether changes in skill acquisition and skill prices for each gender have a positive or negative impact on the declining gender wage gap, as well as the extent of that impact.

5 Estimation of the Counterfactual Wage Gain across Two Cohorts

In this section, I discuss the estimation method of the gender-specific counterfactual wage gain across two cohorts explained by changes in skill components.

5.1 The Dinardo, Fortin, and Lemieux (DFL) Mothod

To quantify the degree to which changes in premarket skills explain the observed trend in the wage distribution, I construct a counterfactual wage distribution —the wage distribution of the NLSY79 cohort if they had the skill components of the NLSY97 cohort while facing the NLSY79 cohort's wage function that determines the relationship between the skill components and the adult wage. I apply the density reweighting procedure introduced by DiNardo et al. (1996), reweighting the NLSY79 sample to have the same distribution of the skill components as the NLSY97 sample and evaluate how the wage distribution changes in the reweighted NLSY79 sample compared with the sample prior to reweighting. While the econometric method I employ closely follows Altonji et al. (2012), I present the method below for the sake of expositional clarity.

Let z be a vector of observed skill components and let u be a vector of unobservable skills and all other factors that affect wages of cohort $j \in \{79, 97\}$. Let w^j ($j \in \{79, 97\}$) be the adult wage of cohort j, which is determined by $w^j = W^j(z,u)$, where $W^j(z,u)$ is the wage determination function of cohort j. Let $f(w^j|z,j) \equiv f(W^j(z,u)|z,j)$ be the density of adult wages of cohort $j \in \{79,97\}$ conditional on z, where the conditional distribution of u on z follows that of cohort j. The difference in the wage density between the two cohorts can be written as

$$
\begin{aligned}
& f(w^{97}|97) - f(w^{79}|79) \\
& = \underbrace{\int \left[f(W^{97}(z,u)|z_{97},97) - f(W^{79}(z,u)|z_{97},97) \right] f(z_{97},97)dz}_{\text{(i) wage difference explained by changes in the wage determination function}} \\
& + \underbrace{\left[(W^{79}(z,u)|z_{97},97) f(z_{97},97)dz - \int f(W^{79}(z,u)|z_{79},79) f(z_{79},79)dz \right]}_{\text{(ii) wage difference explained by changes in skill components}}
\end{aligned}
\tag{4}
$$

The first term on the right-hand side (Part (i)) captures the difference in the log wage rate between the NLSY79 and the NLSY97 cohorts that is associated with the change in the wage determination function W^j across two cohorts. The second term (part (ii)) captures the wage difference between the two cohorts that is driven by differences in observed skill components (z) and unobservable skills and other factors (u) across two cohorts, when the wage determination function is $W^{79}(z,u)$. Because u is not observable, an estimate for the second term can be obtained under the following assumption.

Assumption 1. Let $g(u|z,79)$ and $g(u|z,97)$ be the conditional densities of u given z for the NLSY79 and NLSY97 cohorts, respectively. Then $g(u|z,79) = g(u|z,97)$ holds.

As noted in Altonji et al. (2012), Assumption 1 is unlikely to hold exactly. Changes in skill prices can alter the distribution of unobservable characteristics conditional on observable characteristics. For instance, as more individuals choose to attend college over time, the unobservable ability of college-educated workers in the more recent cohort could be lower than that of the older cohort (Bowlus and Robinson (2012)). However, as discussed in Altonji et al. (2012), it is not possible to directly test Assumption 1 because u is not observed. Moreover, it is not clear whether the estimation of the counterfactual wage distribution $f(w^{79}|z,97)$ would be overstated or understated under Assumption 1. While acknowledging the limitation, I follow the approach of Altonji et al. (2012) and discuss the counterfactual wage distribution of the NLSY79 cohort that is valid under Assumption 1.

Under Assumption 1, $f(w^{79}|z,79) = f(w^{79}|z,97)$ holds. This equality allows me to calculate part (ii) of equation (4), the counterfactual wage distribution of the NLSY79 cohort when the distribution of characteristic z follows that of the NLSY97 cohort. Let $f(z|j)$ be the density distribution function of z for cohort $j \in \{79, 97\}$. Under Assumption 1, the DFL method implies:

$$\begin{aligned} f(w^{79}|97) &= \int f(w^{79}|z,97) f(z|97) dz \\ &= \int f(w^{79}|z,79) f(z|79) \psi(z) dz, \end{aligned} \tag{5}$$

where

$$\psi(z) = \frac{f(z|97)}{f(z|79)} = \frac{p(97|z)}{p(79|z)} \frac{p(79)}{p(97)}. \tag{6}$$

and $p(97|z)$ and $p(79|z) = 1 - p(97|z)$ are the respective propensity scores to observe z in the NLSY97 sample and the NLSY79 sample from the pooled sample. Basically, I reweight the distribution of z for the NLSY79 cohort so that the reweighted distribution represents the distribution of z for the NLSY97 cohort. Then I use the observed relationship between z and w^{79} to estimate the counterfactual distribution $f(w^{79}|z_{97}, 97)$.

Once I estimate part (ii), the estimate for part (i)—the wage difference between the two cohorts that is explained by the different wage determination function across the two cohorts—can be estimated by subtracting the estimate for part (ii) from the actual wage difference between the two cohorts as written in equation (4). Since I observe wage data for the NLSY97 cohort, I can extend Altonji et al. (2012) and quantify the relative importance of part (i) and part (ii) in explaining the actual wage difference between the two cohorts.

5.2 Sequential Marginal Effects of Subsets of Characteristics

To quantify the contribution of subsets of characteristic z between the actual and counterfactual wage distribution for the NLSY79 cohort, I can define the sequential marginal effect (SME). For simplicity, consider the case in which z is divided into two subvector (z_1, z_2). Altonji et al. (2012) show that under Assumption 1, $f(w^{79}|97) - f(w^{97}|79)$ can be decomposed as follows:

$$\begin{aligned} f(w^{79}|97) - f(w^{79}|79) = & \underbrace{\int f(w^{79}|z_1, z_2, 79)\left[f(z_1, z_2, 97) - f(z_2|z_1, 79)\psi(z_1)f(z_1|79)\right]dz}_{\text{(a) SME of } z_1} \\ & + \underbrace{\int f(w^{79}|z_1, z_2, 79)\left[f(z_2|z_1, 79)\psi(z_1)f(z_1|79) - f(z_2|z_1, 79)f(z_1|79)\right]dz,}_{\text{(b) SME of } z_2} \end{aligned} \tag{7}$$

where $\psi(z_1) = f(z_1|97)/f(z_1|79) = \left[p(97|z_1)/p(79|z_1)\right]\left[p(79)/p(97)\right]$. Thus, the difference in $f(w^{79}|97) - f(w^{79}|79)$ can be decomposed into two components: (a) a component explained by the difference in z_1 across the two cohorts when the density of z_2 conditional on z_1 remains the same as in the NLSY79 cohort (SME of z_1), and (b) a component that is explained by the additional change in z_2 across the two cohorts that is not accounted for by part (a) (SME of z_2). The SME of z_1 has two effects: the direct effect from the different distributions of z_1 between the two cohorts, and the indirect effect from the different distributions of z_2 that is induced by the relationship between z_2 and z_1 for the NLSY79 cohort ($f(z_2|z_1, 79)$). The decomposition result crucially depends on the order of (z_1, z_2). In particular, the decomposition is based on the strong assumption that changes in z_1 will translate into changes in z_2 to the extent implied by the conditional density of z_2 on z_1 for the NLSY79 cohort. However, the observed relationship between (z_1, z_2) is not necessarily driven by a causal impact of z_1 on z_2, and changes in z_1 may not have the exact same impact on z_2 for the NLSY97 cohort as it does for the NLSY79 cohort. Therefore, the SME of z_1 could overstate the true impact of changes in z_1 on the wage distribution.

I divide z according to the timing of the variable, to the extent that the timing when the variable is determined is obvious. For instance, I first control for race, family background factors, and the cognitive skills. Then I also control for highest grade completed, degree attainment, major field of study to capture the educational attainment. For work experience, I control for the number of years of working between ages 22-34, the number of full time work, and occupation. Despite that the principle to determine ordering of the variables is not clearly defined, the decomposition does not require a parametric assumption for the wage function $W(z,u)$ and can be applied to the entire wage distribution.

Table 4 summarizes the model specification for the estimation of the SME of different skill components. Let z_j be the skill components included in Model j. Because I construct z_j to satisfy $z_j \subset z_{j+1}$, the sequential marginal effect of $\bar{z}_j = z_{j+1} \setminus z_j$ can be estimated by the difference between the estimated wage gain from Model $j+1$ and that from Model j.

6 Results

6.1 Gender Specific Wage Gain Across Two Cohorts Explained by Changes in Skill Components

In this section, I document the estimation result of the counterfactual wage distribution for each model in Table 4.

6.1.1 Pre-College skill Components

Panels (a) and (b) of Figure 3 present the estimation results for Models 0-3 on the log wage gain across two cohorts for men and women, respectively. Panels (c) and (d) of Figure 3 present the corresponding estimation results on the wage rate in absolute dollars for men and women, respectively.[10] Changes in the racial composition has a small negative impact on the wage gain for men,

[10]Columns (2)-(9) in Table 5 present the estimated log wage gain across two cohorts when I control for different sets of skill components in the estimation as specified as in Models 0-8 in Table 4.

but its impact is negligible for women. Consistent with Altonji et al. (2012), the increase in parental education (Model 1: solid line with plus marks) across two cohorts explains substantial wage gain across two cohorts for both men and women. The log wage gains explained by parents' education in the top half of the wage distribution are similar across different levels of the wage distribution for men but decrease along the percentile of the distribution for women. Consequently, the wage gain in absolute dollars explained by parents' education (Panel (c) of Figure 3) monotonically increases in the top half of the wage distribution for men, while it is similar between the 50th and 80th percentiles of the wage distribution for women.

Note that the marginal effect of changes in family structure that is included in Model 2 but not in Model 1 can be estimated by the differences in the counterfactual wage gain in Models 1 and 2. While changes in parental education contributes positively to the wage gain at all levels of the wage distribution for both men and women, the marginal effects of the changes in family structure—increasing share of single-parent families—conditional on parents' education are negative on the wage gain across two cohorts for both men and women. Interestingly, the negative marginal effect of changes in family structure on the wage gain is significantly greater for men than for women, both in relative and absolute terms. This finding is somewhat puzzling, considering that changes in family structure themselves do not show significant gender differences (Panel B of Table 1). This could be explained by gender differences in how much family structure affects wages for each gender, with a greater negative impact on wages for men when there are more single parents.

The wage gain estimated in Model 3, which includes the AFQT score along with race, parents' education, and family structure, shows a significantly greater increase for women than for men in both relative and absolute terms across two cohorts. In fact, women's relative wage gain compared to men's increases in Model 3 compared to Model 2, in both absolute and relative terms. This finding is consistent with the fact that the AFQT score initially showed no gender differences in the NLSY79 cohorts but diverged significantly over two cohorts, increasing for women and decreasing for men as discussed in Section 2. Therefore, the gender-divergence in the AFQT score across two cohorts conditional on other pre-college skill components seem to play an important role in closing the gender wage gap.

Note that the overall shape of the wage gain across the wage distribution is similar between Models 1, 2, and 3, indicating that changes in parents' education across two cohorts play important roles in determining the relative wage gains at each percentile. On the other hand, the negative sequential marginal effect of the AFQT score on wage gain implies that in comparison to changes in race, parents' education, and family structure, the observed AFQT score for the NLSY97 cohorts is relatively smaller than what would have been anticipated based on the correlation between skill components and wage rates observed in the NLSY79 cohorts. Such a change in the correlation between the AFQT score and other skill components is more pronounced for men across two cohorts, as men's SME of the AFQT score has greater negative values.

Overall, I find that the wage gain across two cohorts, explained by skill components formulated prior to the college period, is positive for both men and women. However, women's wage gain associated with changes in pre-college period skill components is greater than that of men. The sequential marginal effect of different skill components suggests that while parents' education plays a determinant role in shaping the overall wage gain across various wage distribution levels, changes in the family structure and AFQT scores can further explain why women's wage gain across two cohorts is greater than that of men.

6.1.2 Educational Attainments

Panels (a) and (b) of Figure 4 display the estimation results for Models 3, 5, and 6 regarding the log wage gain across two cohorts for men and women, respectively. Model 3 serves as the benchmark for explaining wage gains based on pre-college period skill components. Model 5 incorporates the highest grade completed and degree attainment in higher education, while Model 6 further includes the major field of study. Panels (c) and (d) of Figure 4 present the corresponding estimation results on the wage rate in absolute dollars for men and women, respectively.

First, I find substantial gender difference in the sequential effect of educational attainment controlling for the pre-college skill components. For men, the SME of educational attainment, as measured by highest grade completed and degree attainment, is negligible compared to the wage gain explained by pre-college skill components. For women, changes in the educational attainment explains additional wage gain after controlling for changes in pre-college skill components. In particular, the log wage gain explained by Model 3 for women is 0.07, 0.12, 0.09 (Column (5) of Panel B in Table 5) at 10th, 50th, and 90th percentiles of the wage distribution. On the other hand, the corresponding numbers for Model 5, which additionally includes the highest grade completed and degree attainment compared to Model 3, are 0.09, 0.16, and 0.13, respectively. Thus, the SME of educational attainment measures for women on the wage gain across two cohorts is about 0.02 to 0.04 log points. Overall, the findings imply that while women's wage gain associated with increased educational attainment cannot be predicted by changes in women's pre-college period skill components, men's wage gain associated with changes in educational attainment is almost entirely explained by changes in men's pre-college skill components.

Secondly, after accounting for the highest grade completed and degree attainment, changes in college major do not provide further explanation for the wage gain across two cohorts for both men and women. This observation aligns with Panel B of Table 2, which indicates modest changes in major choices across two cohorts for both genders. However, it's worth noting that the persistence of a gender gap in the Applied STEM major, where wages are generally higher than in other majors, could partially contribute to the enduring gender wage gap.

6.1.3 Work Experience

Figure 5 presents the estimation results for Models 6,7, and 8 by gender. Model 6 serves as the benchmark for the wage gain explained by pre-college and during-college period skill components, while Model 7 incorporates labor market experience, and Model 8 further includes occupation. Note that lhe labor market experience is measured by the number of years an individual worked positive hours and full-time (more than 1600 annual hours) between ages 22 and 34. The occupation is measured by dummy variables that takes a value of 1 if the number of years the individual worked in a certain occupation is greater than 5 during ages 22-34.

First, the SME of labor market experience is small for both men and women. For men, the labor market experience has negligible SME after controlling for skill components measured prior or during the college period at all levels of the wage distribution. For women, its SME is negligible in the bottom half of the distribution, while it is slightly negative in the top-half of the wage distribution. This implies that the change in the wage gain associated with changes in labor market experience across two cohorts is well predicted by changes in pre- and during-college skill components for men. Women exhibit similar results in the bottom half of the wage distribution, but at the top end of the distribution, the labor market experience of women is smaller than what can be explained by skill components prior to and during the college periods.

Second, the SME of occupation is negligible for men, whereas it is negative for women. The negative marginal effect of occupation on women's wages is more pronounced in the lower half of the wage distribution, decreasing by about 0.03 log points, while it is nearly zero at the top end. The negative marginal effect of changes in occupational sorting on women's wage gains in the lower half of the wage distribution can be attributed to the substantial decrease in the share of clerical occupations, which was not fully replaced by other types of middle-level occupations, such as education or social occupations (Table 3). Simultaneously, the nearly zero marginal effect of changes in occupation on wage gains for women at the top end can be attributed to women's stagnant progress in sorting into high-paying occupations such as STEM occupations. These findings are consistent with previous studies Blau and Kahn (2017), suggesting that persistent occupational segregation by gender is an important factor in explaining the gender wage gap.

6.2 Implications on the Gender Wage Gap

In this section, I explore the extent to which gender-specific changes in skill components and skill prices can account for the trend in the gender wage gap across two cohorts.

Panels A, B, and C of Table 7 summarize the decomposition results of the log gender wage gap at the 10th, 50th, and 90th percentiles of the wage distribution, respectively. Columns (1) and (2) of Row (a) present the log wage difference between men and women observed in the data for the NLSY79 and the NLSY97 cohort, respectively. Columns (3), (4), and (5) of Row (a) present the estimated log gender wage gap explained by changes in skill components accounted for in Model 3, Model 6, and Model 8, respectively. Note that Model 3 includes pre-college skill components, Model 6 additionally incorporates educational attainment, while Model 8 further encompasses work experience. Columns (2)-(5) in Row (b) illustrate the reduction in the gender wage gap in the data for the NLSY97 cohort as well as in Models 3, 6, and 8, compared to the observed gender wage gap in the NLSY79 cohort.

The results for the decomposition exercise for the change in the gender wage gap are presented in Rows (c)-(h). Row (c) reports the extent to which changes in skill components for men contribute to the decrease in the gender wage gap, corresponding to Part (i) in equation (3). The corresponding figure for women is presented in Row (d), representing the estimate for Part (ii) in equation (3). On the other hand, Row (e) reports the extent to which changes in skill prices for men contribute to the decrease in the gender wage gap, corresponding to Part (iii) in equation (3). The corresponding figure for women is presented in Row (f), representing the estimate for Part (iv) in equation (3). The numbers in parentheses indicate the percentage of the explained decrease in the gender wage gap by each component of the decomposition compared to the actual decrease in the gender wage gap observed in the data. Row (g) reports the extent to which changes in skill components explain the decrease in the gender wage gap, calculated by adding the percent contributions from Rows (c) and (d). Conversely, Row (h) reports the extent to which changes in skill components explain the decrease in the gender wage gap, calculated by adding the percent contributions from Rows (e) and (f).

First, accounting most comprehensive set of skill components in Model 8, the role of changes in skills in explaining the decrease in the log gender wage gap across two cohorts is largest in the middle and smallest at the top end of the wage distribution. In particular, comparing Column (5) Row (g) across Panels A,B, and C, changes in skill components explain 77.1% of the decrease of the gender wage gap at the 50th percentile of the wage distribution, whereas the corresponding numbers at the 10th and 90th percentiles of the wage distribution are 45.6% and 33.9%, respectively. This implies that the relative contribution of changes in skill prices to the decrease in the gender wage gap, as documented in Row (h), is greatest at the 90th percentile (66.1%), followed by the 10th percentile (54.4%) and the 50th percentile (22.9%).

I can further decompose to what extent women's skill gain and men's skill gain across two cohorts can explain the decreased gender wage gap. For instance, at the 50th percentile of the wage distribution, in Column (5) Rows (c) and (d) of Panel B, if men's skill component did not change across two cohorts, women's skill gain could explain 104.2% of the observed decrease in the gender wage gap. However, if men's skill improved across two cohorts while women's skill remained the same as in the old cohort, it could have explained a 27.1% increase in the observed change in the gender wage gap. Together, the net effect of changes in skill components in explaining the decrease in the gender wage gap is 77.1% in the median. Similarly, Row (e) suggests that if men's skill price did not change across two cohorts, women's wage gain from changes in skill prices could widen the gender wage gap by 50.7% of the observed decrease. On the other hand, the change in the skill price for men alone contributes to 73.5% of the observed decrease in the gender wage gap. Together, changes in skill prices for men and women contributes to 22.9% of the decreased gender wage gap.

The proposed decomposition provides an alternative way to quantify the relative importance of women's wage gain to men's. In particular, women's changes in skill components have made a significant positive contribution to the decrease in the gender wage gap in the middle range of the wage distribution. Specifically, the decrease in the gender wage gap at the 50th percentile can be entirely attributed to women's skill gain, accounting for 104.2% of the reduction. However, changes in skill prices for women in the middle range have had a negative impact on women's wage gain across both cohorts, resulting in a net wage gain of 53.5% of the observed decrease in the gender wage gap. This represents a reduction by half of the wage gain explained by skill components. Conversely, men's wage loss associated with changes in skill prices in the middle range of the wage distribution contributes to the decrease in the gender wage gap by 73.5%, but as men also gained wages from changes in skill components, their net contribution to the decrease in the gender wage gap becomes 46.4%. Overall, women's wage gain, mainly due to increased skill, and men's wage loss, primarily attributed to price decreases, have each contributed about half to the reduction in the gender wage gap in the middle range of the wage distribution.

Second, as shown in Panel C, at the 90th percentile of the wage distribution, women's skill gain alone would explain 85.2% of the observed decrease in the gender wage gap, whereas men's skill gain counteracts its impacts by 51.2%, leading to the net effect of skill changes in explaining the observed gender wage gap to be 33.9%. Therefore, at the top end of the wage distribution, despite substantial progress in women's skill gains, men also make substantial progress across two cohorts, resulting in relatively smaller catch-up by women through skill accumulation. On the other hand, women's wage gain associated with changes in skill price alone can explain 78.5% of the decrease in the gender wage gap, whereas men's wage gain associated with changes in skill prices widens the gender wage gap by 12.4% of the observed decrease. Therefore, at the top end of the wage distribution, women's relative progress in skill prices compared to men plays a more significant role in closing the gender wage gap than their relative progress in skill acquisition compared to men.

It is worth noting the different roles of skill price changes on the decrease in the gender wage gap between 50th and 90th percentile of the wage distribution. Skill price changes have a large negative impact on men's wage gain at the 50th percentile (-0.0959 log points), whereas they have a relatively small positive impact at the 90th percentile (0.0156 log points). Likewise, changes in skill prices for women negatively impact women's wage gain at the 50th percentile (-0.0661 log points) while significantly boosting wage gain at the 90th percentile (0.0989 log points). Consistent with the polarization literature (Autor et al. (2006)), both men and women experience wage loss with the same skill sets in the middle range of the wage distribution, while both genders experience wage gain associated with increasing skill prices at the top end of the distribution. Consequently, the impact of changes in skill prices on the gender wage gap depends on women's relative wage loss in the middle and their relative wage gain at the top. The findings from comparing two NLSY cohorts suggest that men experienced greater wage loss in the middle due to decreasing skill prices, while women had greater wage gains at the top end. Consequently, changes in skill prices had positive impacts on reducing the gender wage gap at both the 50th and 90th percentiles of the wage distribution, with the effects being more pronounced at the 90th percentile.

Third, the wage gain explained by changes in skill components is smallest at the 10th percentile of the wage distribution for both men and women. Men's wages at the bottom are rarely affected by changes in their skill components, while women's wages increase slightly by 0.0428 log points. Consequently, changes in skill components contribute positively to the decreasing gender wage gap, accounting for 45.6% of the decrease. Similarly, the impact of skill price changes on men's wages at the bottom end is negligible, whereas women's wages increase by 0.05 log points due to increased skill prices across both cohorts. Overall, at the lower end of the wage distribution, nearly all of the reduction in the gender wage gap can be attributed to women's wage increases resulting from enhanced skill components and higher skill prices, while men's wages remain relatively stable across the two cohorts.

When comparing the role of skill components during the pre-college, during-college, and post-college periods in explaining the observed gender wage gap (represented in Models 3, 6, and 8), I find that changes in pre-college skills and educational attainment positively contribute to the decrease in the gender wage gap, while changes in work experience have a negative impact. For instance, at the 50th percentile of the wage distribution (Panel B), skill components account for 59.1%, 92.8%, and 77.1% of the explained gender wage gap in Models 3, 6, and 8, respectively. The marginal effect of pre-college skills on the decrease in the gender wage gap is greatest in the middle (59.1%) and smallest at the top end (12.6%) of the wage distribution, while the marginal effect of education is highest in the middle (33.7%) and smallest at the bottom end (11.3%). On the other hand, the negative marginal effect of work experience on the gender wage gap is most pronounced in the middle (15.7%) and least significant at the top (6.9%).

The trend in the gender wage gap and its decomposition along the entire wage distribution can be found in Figure 6. Panel (a) of Figure 6 displays the log gender wage gap observed in the NLSY79 and NLSY97 cohorts, represented by solid lines with solid circles and hollow circles, respectively. I then plot the estimated counterfactual gender wage gap using the wage distribution for each gender, attributing it to changes in skill components included in Model 2,3,5,6,7, and 8 while maintaining the wage determination function the same as for the NLSY79 cohorts. Panel (b) presents the corresponding estimation of the gender wage gap measured in absolute dollars.

By applying the decomposition method outlined in Section 4, the difference between the observed gender wage gap for the NLSY79 cohorts (solid line with solid circles) and the counterfactual gender wage gap in each model captures the change in the gender wage gap explained by changes in the skill components considered in each model. This part corresponds to $\left[(\overline{X}_{1m} - \overline{X}_{0m})b_{0m}\right] - \left[(\overline{X}_{1f} - \overline{X}_{0f})b_{0f}\right]$ in equation (3), representing the difference between men's and women's wage gain explained by a certain set of skill components across two cohorts (Part (i) for men and Part (ii) for women). Conversely, the difference between the counterfactual gender wage gap estimated in each model and the observed gender wage gap for the NLSY97 cohort (solid line with hollow circles) captures the gender wage gap that can be explained by changes in gender-specific skill prices across two cohorts. This part corresponds to $\left[\overline{X}_{1m}(b_{1m} - b_{0m})\right] - \left[\overline{X}_{1f}(b_{1f} - b_{0f})\right]$, representing the difference between men's and women's wage gain explained by gender-specific changes in skill prices (Part (iii) for men and Part (iv) for women). Table 6 presents the point estimates for the observed gender wage gap in the data (columns (1) and (2)) and counterfactual gender wage gap explained by changes in skill components (columns (3)-(8)).

First, the solid line presents the counterfactual gender wage gap explained by changes in parents' education and and family structure for each gender (Model 2), which results in a 0.054 log-point decrease at the 50th percentile of the wage distribution. However, its impact on the gender wage gap is negligible at the top end of the wage distribution. Second, when I additionally account for changes the AFQT score, represented by the dotted line, the gender wage gap further decreases by 0.02 log points at the 50th and 90th percentiles of the wage distribution. However, the gender wage gap does not decrease further at the top 10th percentile of the wage distribution. This finding demonstrates the heterogeneous impacts of changes in skill components in explaining the trend in the gender wage gap across different levels of the wage distribution. Third, the dashed line represents the counterfactual gender wage gap explained by Model 5, which additionally incorporates the highest grade completed and degree attainment of the individuals. Not surprisingly, the gender wage gap decreased substantially, explaining almost all the observed decrease between the 30th and 50th percentiles of the wage distribution. The inclusion of the major field of study (represented by the solid line with triangles) results in a small, negative marginal effect on the gender wage gap between the 50th and 70th percentiles of the wage distribution.

Third, the dashed line shows the counterfactual gender wage gap when considering changes in labor market experience (full-time and part-time employment) between ages 22-34. Changes in the labor market experience have minimal impact on the gender wage gap below the 70th percentile after adjusting for pre- and during-college skill components. However, at the 90th percentile, they increase the gender wage gap by 0.03 log points. Finally, the solid line with a plus mark represents the counterfactual gender wage gap when I further consider changes in occupation to estimate the counterfactual wage distribution. Consistent with the discussion in Section 6, gender-specific changes in occupational choices have adverse impacts on the change in the gender wage gap. The marginal effect on the gender wage gap is most pronounced at the lower end of the wage distribution, with, for instance, a 0.05 log point increase in the gender wage gap at the 20th percentile of the wage distribution.

In summary, changes in skill components can explain three-quarters of the shift in the gender wage gap within the middle range of the wage distribution, but account for only one third of the observed trend at the top end. The skill prices decrease in the middle range of the wage distribution for both men and women, with a larger decline for men, contributing to the reduction of the gender wage gap. Conversely, the decline in the gender wage gap at both ends of the wage distribution is primarily attributed to a significant increase in skill prices for women, while men's skill prices remain relatively stable. This indicates that alterations in the gender wage gap at the high end of the wage distribution are primarily attributed to gender-specific changes in skill prices across the two cohorts. Overall, the findings are consistent with the literature (Galor et al., 1996; Welch, 2000; Weinberg, 2000; Black and Spitz-Oener, 2010; Bacolod and Blum, 2010; Cortes et al., 2020), which documents that changes in skill prices play an important role in explaining the decrease in the gender wage gap at the top of the wage distribution.

7 Conclusion

Based on the NLSY79 and NLSY97, I find substantial changes in skill components across the 1957–1964 and 1980–1984 birth cohorts in the U.S. I also find notable gender divergence in cognitive skill and educational attainment, favoring women across two cohorts. I analyze how changes in various skill components contribute to the recent shifts in the wage distribution by gender and examine the implications for trends in the gender wage gap. I find that changes in pre-college skills, including family structure and cognitive skills, and changes in educational attainment have significantly positive impact on women's relative wage gain, while changes in work experience have a negative impact. Changes in skill components can explain three-quarters of the shift in the gender wage gap within the middle range of the wage distribution, but account for only one third of the observed trend at the top end. The decline in the gender wage gap at both ends of the wage distribution is primarily attributed to a significant increase in skill prices for women, while men's skill prices remain relatively stable.

References

Altonji, J. G., Bharadwaj, P., and Lange, F. (2012). Changes in the characteristics of american youth: Implications for adult outcomes. *Journal of Labor Economics*, 30(4):783–828.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2006). The polarization of the us labor market. *The American economic review*, 96(2):189–194.

Bacolod, M. P. and Blum, B. S. (2010). Two sides of the same coin us âresidualâ inequality and the gender gap. *Journal of Human resources*, 45(1):197–242.

Black, S. E. and Spitz-Oener, A. (2010). Explaining women's success: technological change and the skill content of women's work. *The Review of Economics and Statistics*, 92(1):187–194.

Blau, F. D. and Kahn, L. M. (1997). Swimming upstream: Trends in the gender wage differential in the 1980s. *Journal of labor Economics*, 15(1, Part 1):1–42.

Blau, F. D. and Kahn, L. M. (2006). The us gender pay gap in the 1990s: Slowing convergence. *Ilr Review*, 60(1):45–66.

Blau, F. D. and 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. *Journal of Human resources*, pages 436–455.

Boeri, T., Del Boca, D., and Pissarides, C. A. (2005). Women at work: An economic perspective.

Bowlus, A. J. and Robinson, C. (2012). Human capital prices, productivity, and growth. *American Economic Review*, 102(7):3483–3515.

Cortes, G. M., Jaimovich, N., and Siu, H. E. (2020). The growing importance of social tasks in high-paying occupations: Implications for sorting. Technical report.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4):1593–1640.

DiNardo, J., Fortin, N., and Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64(5):1001–1044.

Elder, T. E., Goddeeris, J. H., and Haider, S. J. (2010). Unexplained gaps and oaxaca–blinder decompositions. *Labour Economics*, 17(1):284–290.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition methods in economics. In *Handbook of labor economics*, volume 4, pages 1–102. Elsevier.

Fortin, N. M. (2008). The gender wage gap among young adults in the united states: The importance of money versus people. *Journal of Human Resources*, 43(4):884–918.

Galor, O., Weil, D. N., et al. (1996). The gender gap, fertility, and growth. *American Economic Review*, 86(3):374–387.

Gosling, A., Machin, S., and Meghir, C. (2000). The changing distribution of male wages in the uk. *The Review of Economic Studies*, 67(4):635–666.

Jann, B. (2008). The blinder–oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4):453–479.

Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of political Economy*, 101(3):410–442.

Kassenboehmer, S. C. and Sinning, M. G. (2014). Distributional changes in the gender wage gap. *ILR Review*, 67(2):335–361.

Machado, J. A. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, 20(4):445–465.

Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *The review of economics and statistics*, 90(2):290–299.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, pages 693–709.

O'Neill, J. (2003). The gender gap in wages, circa 2000. *American Economic Review*, 93(2):309–314.

O'Neill, J. and Polachek, S. (1993). Why the gender gap in wages narrowed in the 1980s. *Journal of Labor economics*, 11(1, Part 1):205–228.

Segall, D. O. (1997). Equating the cat-asvab.

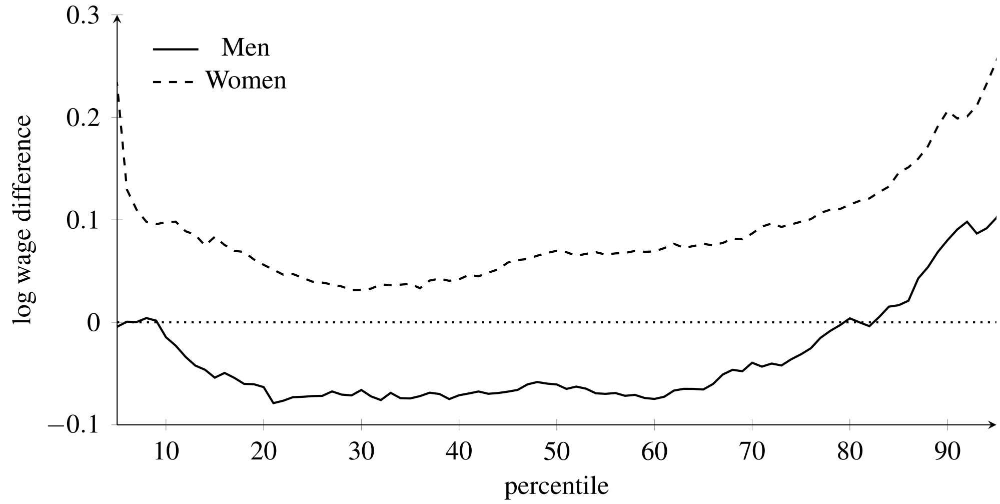
Weinberg, B. A. (2000). Computer use and the demand for female workers. *ILR Review*, 53(2):290–308.

Welch, F. (2000). Growth in women's relative wages and in inequality among men: One phenomenon or two? *American Economic Review*, 90(2):444–449.

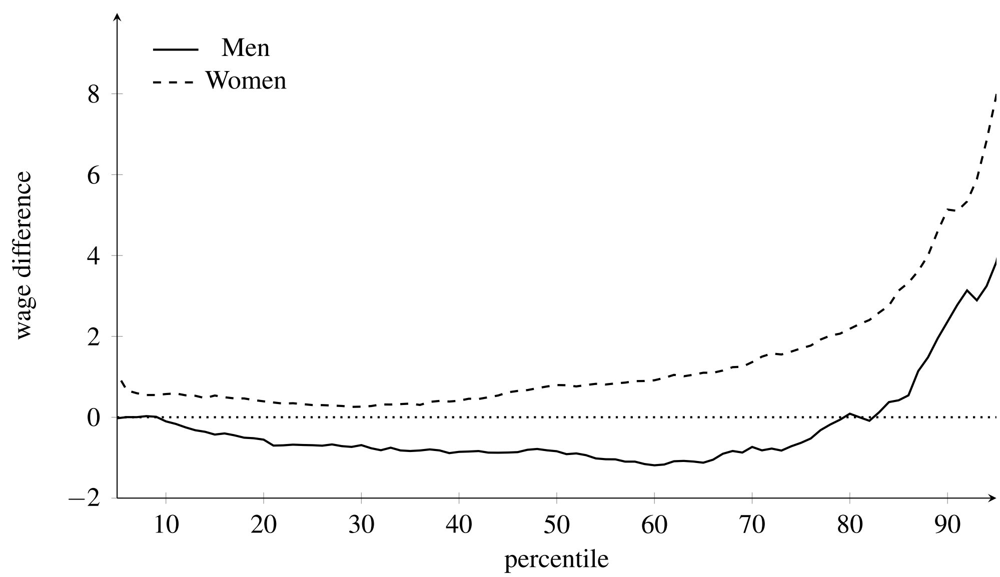
Wellington, A. J. (1993). Changes in the male/female wage gap, 1976-85. *Journal of Human resources*, pages 383–411.

Figure 1: Wage Gain between the NLSY97 and NLSY79 cohorts

(a) Changes in Log Wage between the NLSY79 and NLSY97 Cohorts by Gender



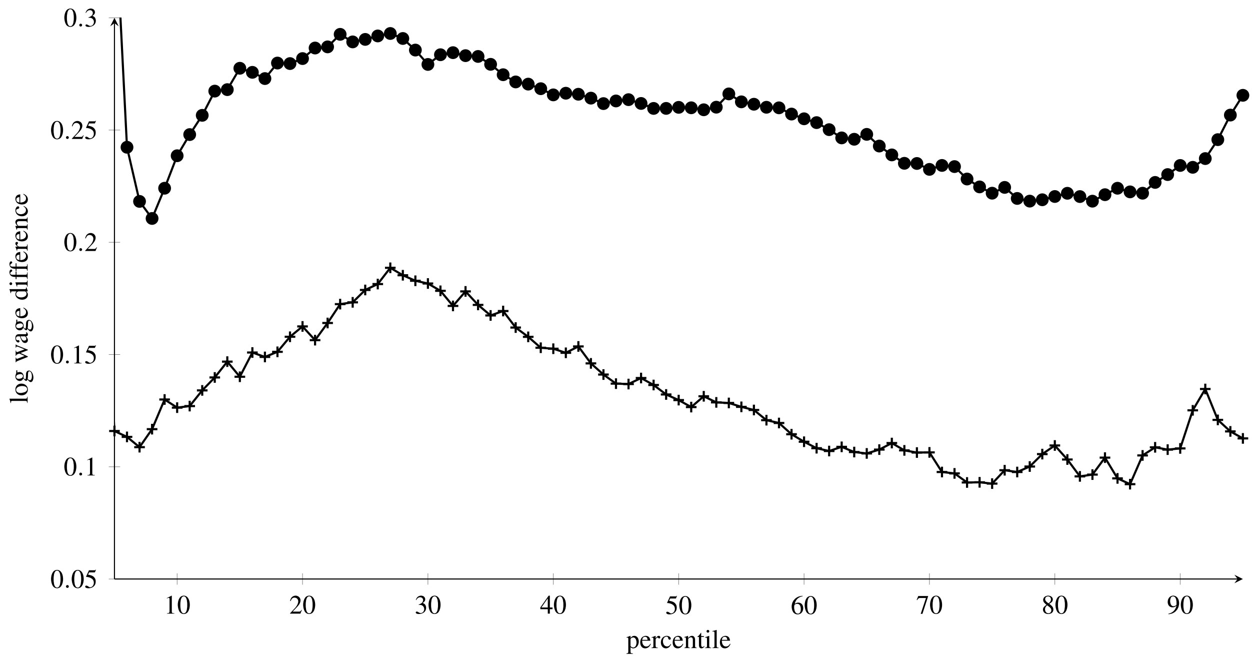
(b) Changes in Wage between the NLSY79 and NLSY97 Cohorts by Gender in Absolute Dollars



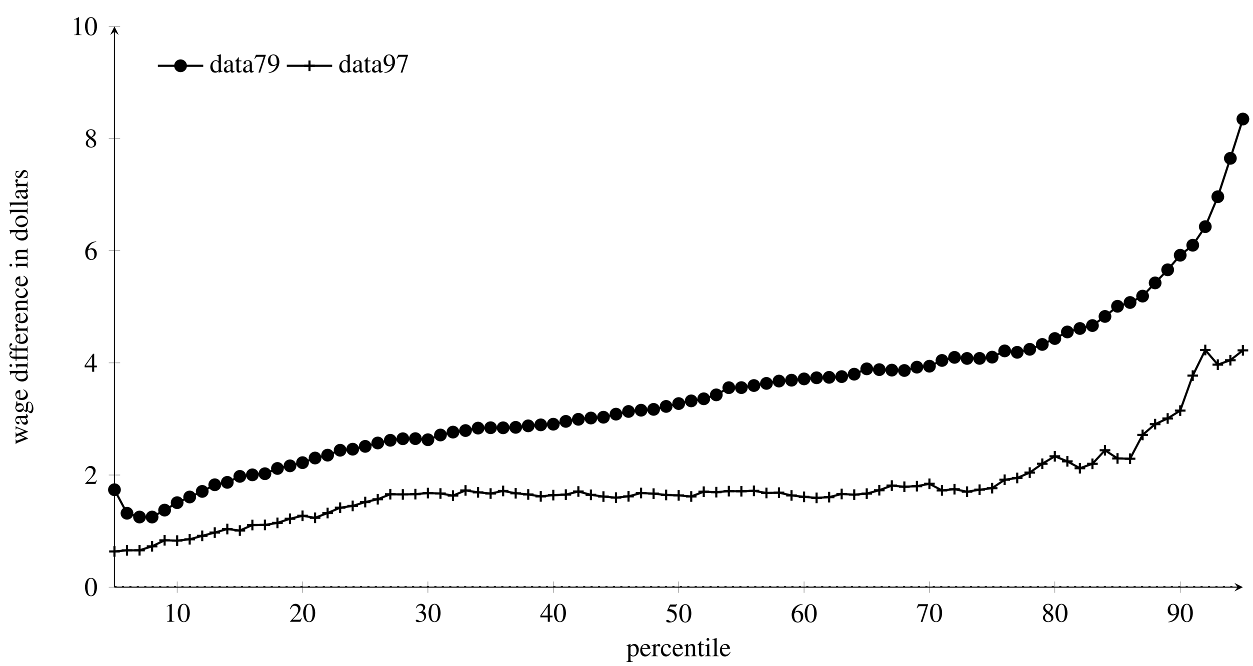
Note: Panel (a) displays the log wage gain of the NLSY97 cohort compared to the NLSY79 cohort at each percentile of the wage distribution. The hourly wage is standardized through regression with 13 years of potential work experience. Panel (b) presents the corresponding wage gain across the two cohorts, measured in absolute dollars.

Figure 2: Gender Wage Gap explained by Changes in Precollege Skills across Two Cohorts

(a) Gender Wage gap (log wage difference)

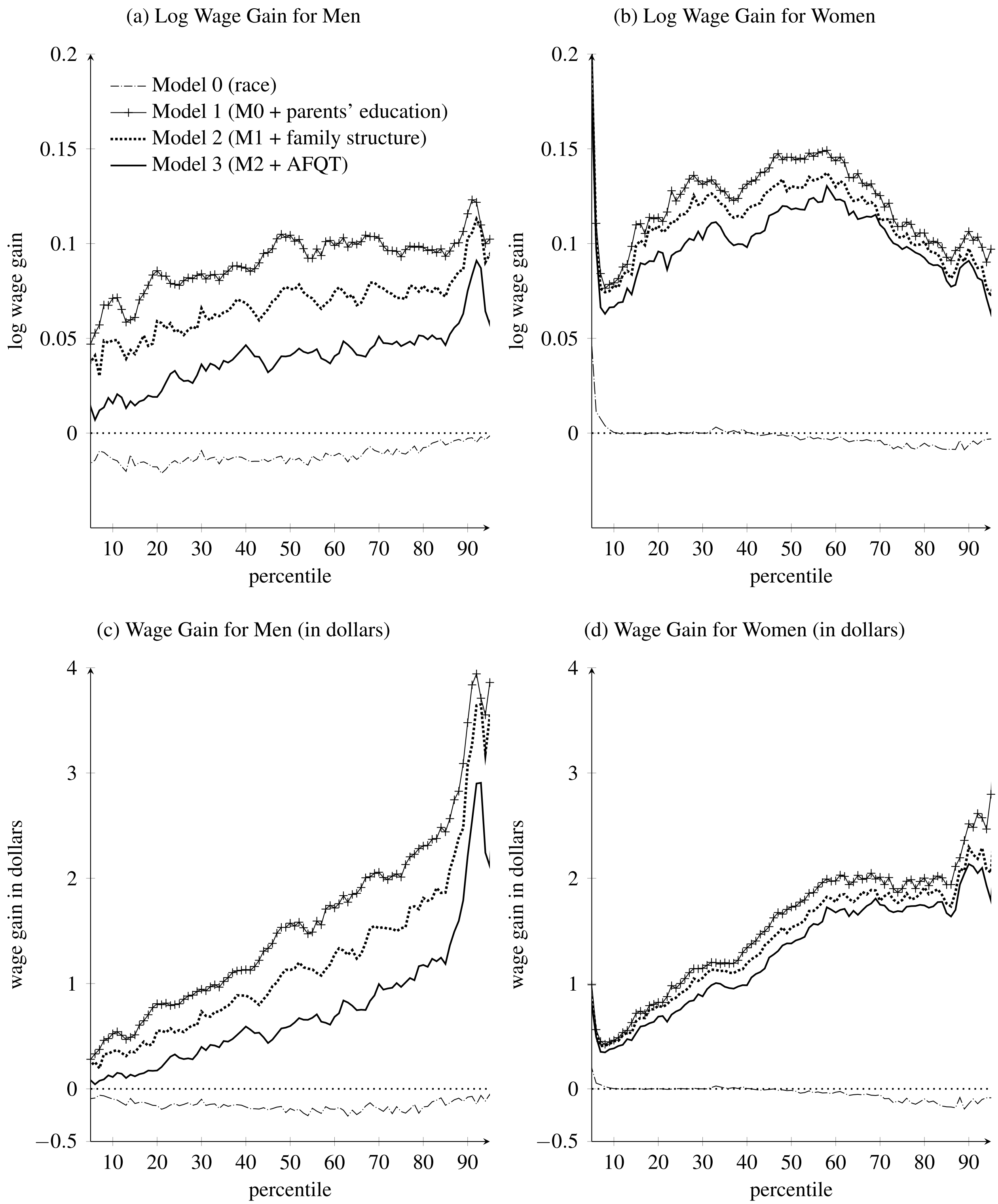


(b) Gender Wage gap (in dollars)



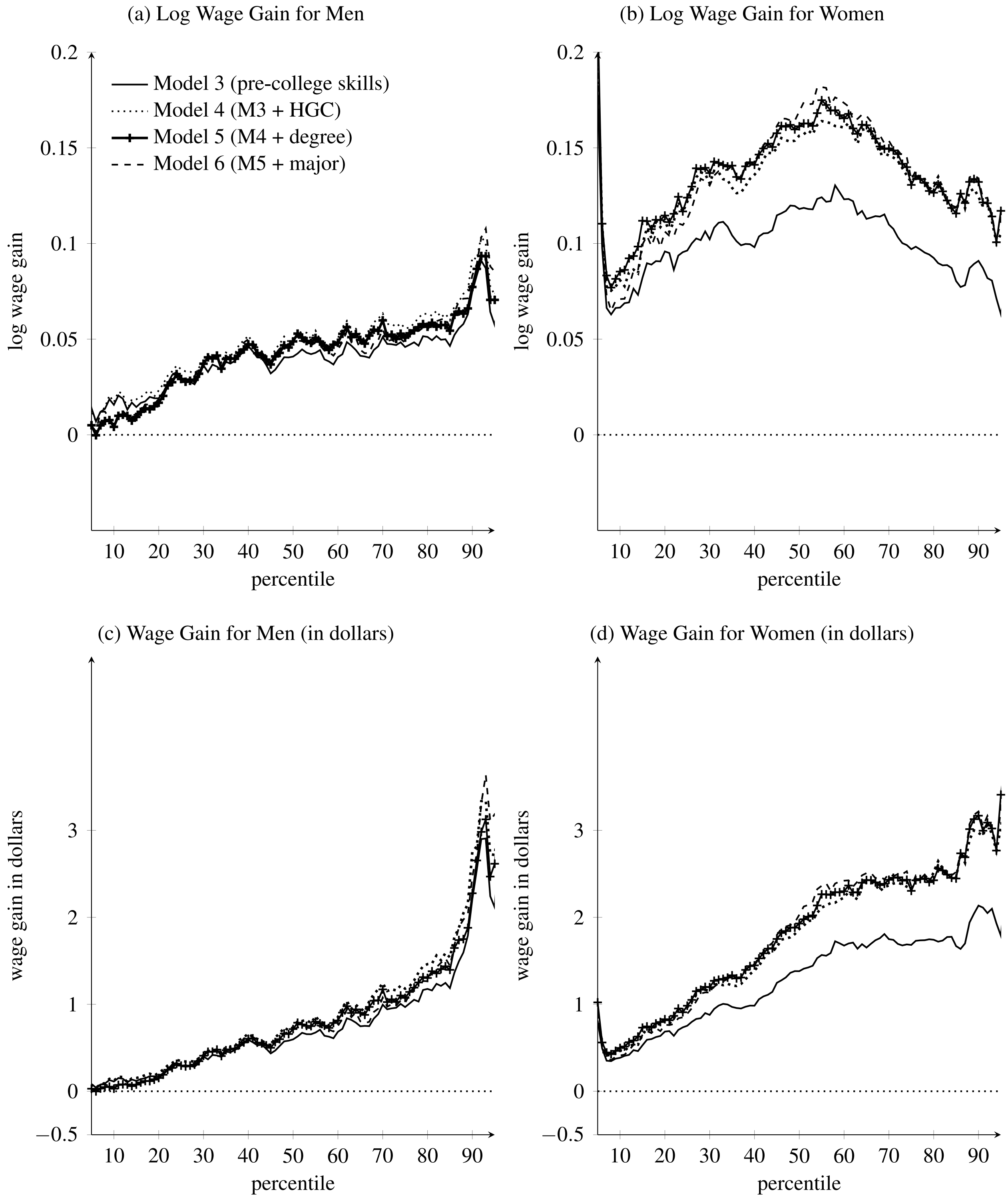
Note: Panel (a) illustrates the log wage difference between men and women at each percentile of the wage distribution for the NLSY79 cohort (solid line with solid circles) and the NLSY97 cohort (solid line with plus marks). Panel (b) presents the corresponding gender wage gap in absolute dollars.

Figure 3: Counterfactual Wage Gain explained by Changes in Precollege Skills



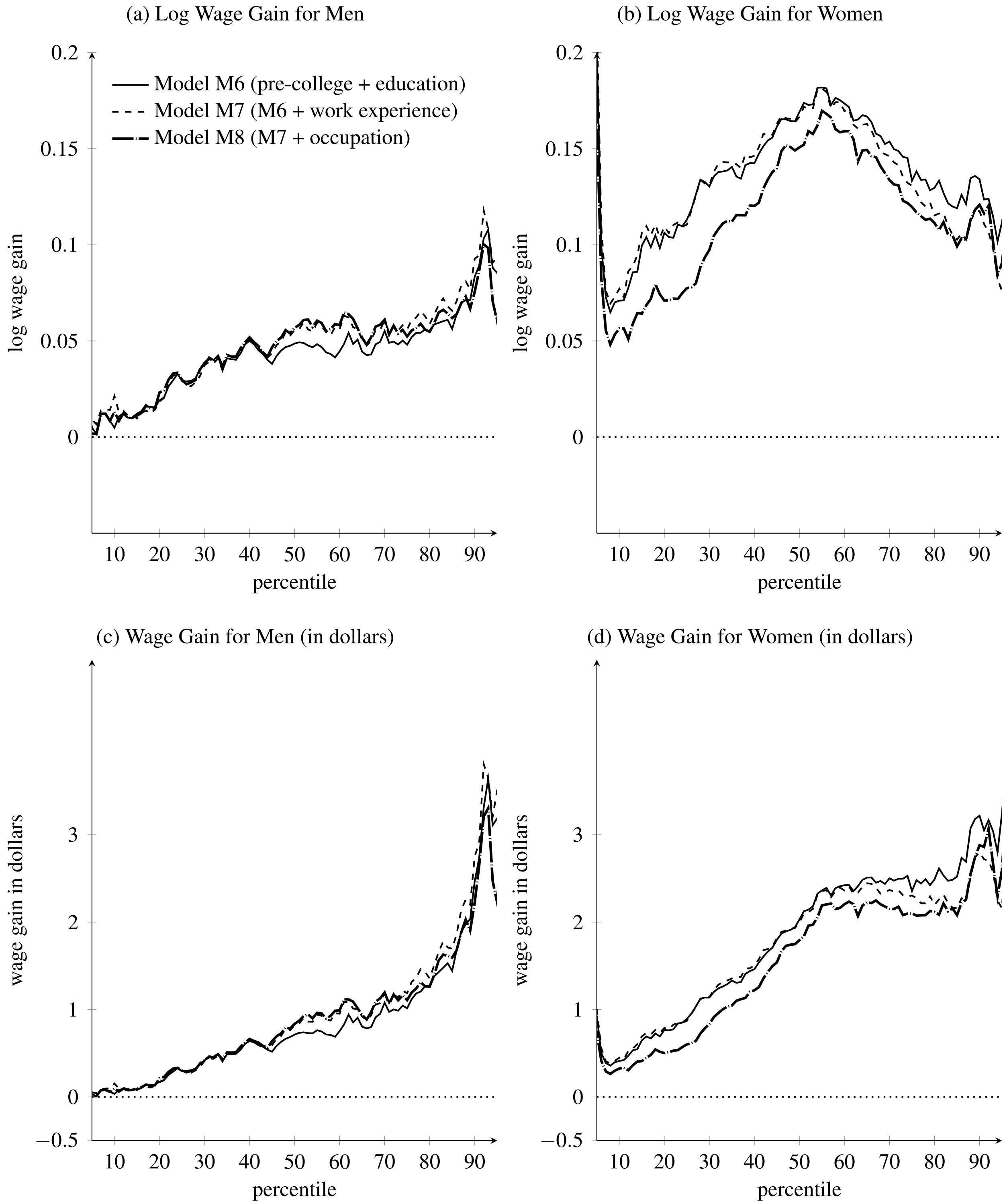
Note: Panels (a) and (b) document the estimated counterfactual log wage gain that is explained by changes changes in the pre-college skill components included in Models 0-3 for men and women, respectively. Each subsequent model (Model n+1) builds upon the previous one (Model n) by incorporating an additional component. Model 0 considers only racial composition, Model 1 adds parents' education, Model 2 adds family structure as measured by which parents the youth lived with, and Model 3 adds the AFQT score. Panels (c) and (d) depict the corresponding wage gain measured in absolute dollars for men and women, respectively.

Figure 4: Counterfactual Wage Gain explained by Changes in Educational Attainment



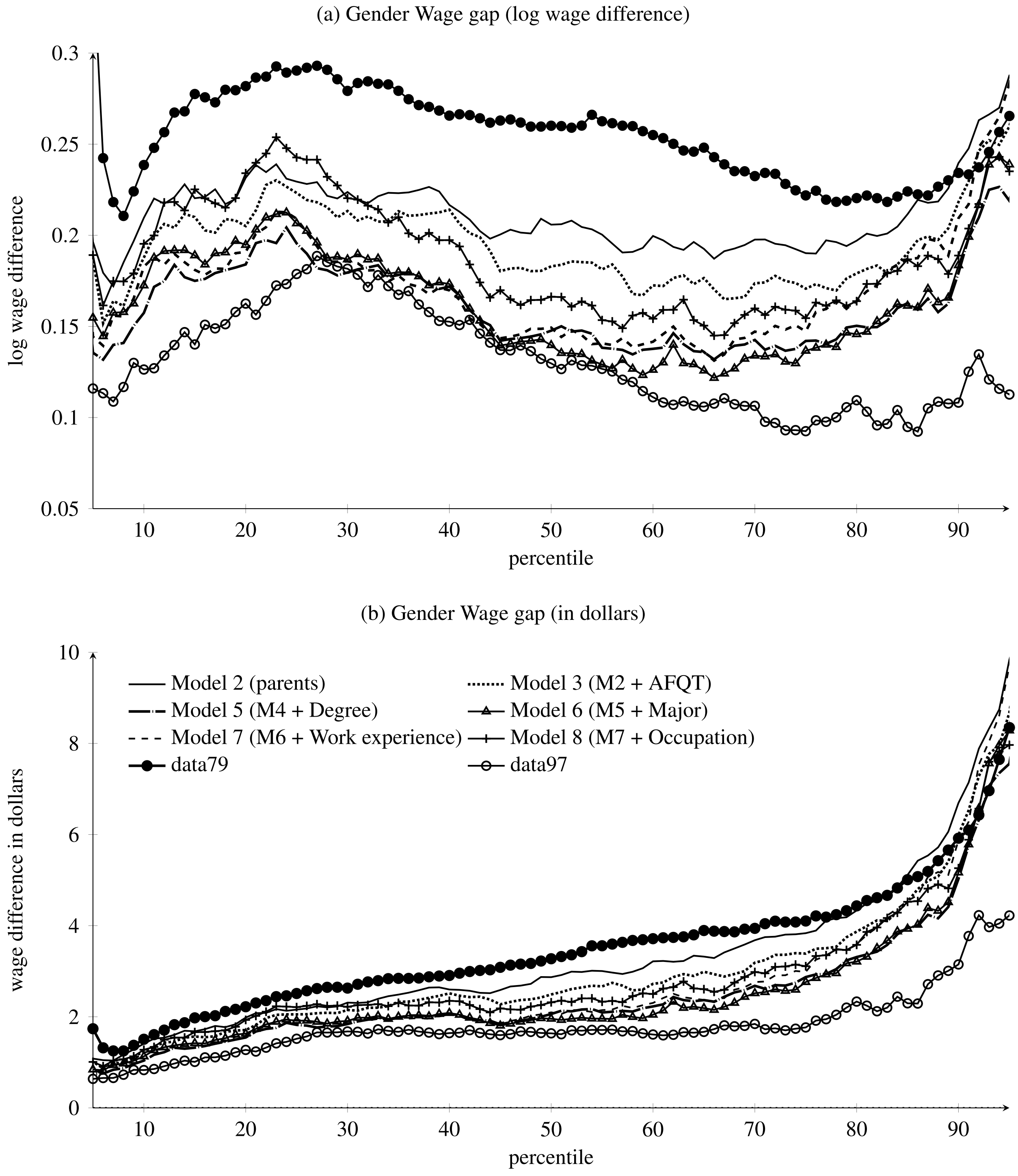
Note: Panels (a) and (b) display the estimated counterfactual log wage gains explained by changes in skill components in Models 3, 4, 5, and 6 for men and women, respectively. Model 3 includes variables such as race, parents' education, family structure, and the AFQT score, serving as a benchmark that accounts solely for pre-college skill components. Each subsequent model (Model n+1) builds upon the previous one (Model n) by incorporating an additional component. Model 4 incorporates the highest grade completed. Model 5 adds degree attainment, and Model 6 further includes major field of study. Panels (c) and (d) illustrate the corresponding wage gains measured in absolute dollars for men and women, respectively.

Figure 5: Counterfactual Wage Gain explained by Changes in Work Experience



Note: Panels (a) and (b) document the estimated counterfactual log wage gain that is explained by changes in the skill components in Models 6-8 for men and women, respectively. Each subsequent model (Model n+1) builds upon the previous one (Model n) by incorporating an additional component. Model 6 includes race, parents' education, family structure, the AFQT score, highest grade completed, highest degree attainment, and major field of study, which serves as a benchmark that accounts for pre-college skills and educational attainment. Model 7 adds work experience, measured by the number of years employed and the number of years employed full-time during ages 22-34. Model 8 adds occupation. Panels (c) and (d) depict the corresponding wage gain measured in absolute dollars for men and women, respectively.

Figure 6: Gender Wage Gap explained by Changes in Skill Components across Two Cohorts



Note: Panels (a) and (b) present the estimated gender wage gaps in Models 2, 3, 5, 6, 7, and 8—the counterfactual gender wage gap when a subset of skill measures of the NLSY79 cohort is replaced with that of the NLSY97 cohort, while the wage determination function and the distribution of other skill correlates conditional on the replaced skill correlates remain the same as for the NLSY79 cohort. The solid line with solid circles and the solid line with hollow circles depict the observed gender wage gap for the NLSY79 and NLSY97 cohorts, respectively. Panel (b) presents the corresponding gender wage gap measured in absolute dollars.

Table 1: Demographic Characteristics by Gender

Sample	Men		Women	
	NLSY79	NLSY97	NLSY79	NLSY97
A. Youth characteristics				
White	0.81	0.71	0.80	0.71
	[0.39]	[0.45]	[0.40]	[0.45]
Black	0.14	0.15	0.15	0.16
	[0.35]	[0.36]	[0.36]	[0.37]
Hispanic	0.05	0.14	0.05	0.13
	[0.22]	[0.34]	[0.22]	[0.33]
Urban	0.79	0.78	0.79	0.78
	[0.41]	[0.51]	[0.41]	[0.51]
AFQT score	0.008	-0.037	0.023	0.06
	[1.04]	[1.07]	[0.944]	[0.96]
B. People living in the household				
Only mother presents (%)	0.18	0.34	0.20	0.38
	[0.39]	[0.47]	[0.40]	[0.48]
Only father presents (%)	0.03	0.06	0.02	0.06
	[0.18]	[0.23]	[0.16]	[0.23]
Both mother and father present (%)	0.76	0.55	0.74	0.52
	[0.43]	[0.50]	[0.44]	[0.50]
C. Parents' characteristics				
Mother's highest grade completed	11.57	13.15	11.48	13.11
	[2.70]	[3.21]	[2.67]	[2.82]
Father's highest grade completed	11.59	13.23	11.40	13.11
	[3.58]	[4.24]	[3.55]	[3.17]
Log family income	10.66	10.68	10.67	10.53
	[0.71]	[1.14]	[0.69]	[1.19]
Mother born in the U.S.	0.96	0.87	0.96	0.87
	[0.20]	[0.33]	[0.19]	[0.33]
Father born in the U.S.	0.96	0.88	0.87	0.90
	[0.19]	[0.32]	[0.33]	[0.30]
Number of observations	4,082	3,093	4,223	3,063

Note: The table documents demographic characteristics of the NLSY79 and the NLSY97 samples by gender, with the standard deviation provided in square brackets.

Table 2: Educational Attainment by Gender

Sample	Men		Women	
	NLSY79	NLSY97	NLSY79	NLSY97
A. Highest Grade Completed as of 22				
HGC	12.65	12.83	12.78	13.34
	[2.03]	[2.22]	[2.06]	[2.32]
B. Degree Received as of 30				
AA	0.07	0.10	0.10	0.14
	[0.25]	[0.30]	[0.30]	[0.34]
BA	0.26	0.28	0.25	0.39
	[0.44]	[0.45]	[0.43]	[0.49]
MA	0.05	0.06	0.04	0.10
	[0.23]	[0.24]	[0.20]	[0.30]
C. Major Field of Study				
Pure STEM	0.024	0.031	0.021	0.035
	[0.155]	[0.173]	[0.142]	[0.184]
Applied STEM	0.123	0.118	0.056	0.025
	[0.329]	[0.322]	[0.231]	[0.155]
Business	0.146	0.128	0.173	0.133
	[0.353]	[0.334]	[0.378]	[0.340]
Social Science	0.024	0.060	0.015	0.060
	[0.154]	[0.238]	[0.120]	[0.238]
Communication	0.014	0.021	0.012	0.021
	[0.117]	[0.142]	[0.110]	[0.142]
Psychology	0.011	0.021	0.021	0.046
	[0.106]	[0.142]	[0.142]	[0.210]
Health	0.027	0.035	0.119	0.156
	[0.161]	[0.183]	[0.324]	[0.363]
Education	0.033	0.021	0.097	0.080
	[0.179]	[0.143]	[0.296]	[0.271]
Art	0.019	0.030	0.019	0.034
	[0.136]	[0.171]	[0.137]	[0.180]
Language/Humanity	0.010	0.031	0.017	0.037
	[0.100]	[0.172]	[0.127]	[0.189]
Number of observations	4,082	3,093	4,223	3,063

Note: The table documents the educational attainment of the NLSY79 and NLSY97 samples by gender, with the standard deviation provided in square brackets. Panel A presents the highest grade completed as of age 22, Panel B presents degree attainment as of age 30, and Panel C presents the major field of study.

Table 3: Work Experience and Occupation by Gender

Sample	Men		Women	
	NLSY79	NLSY97	NLSY79	NLSY97
A. Work Experience during ages 22-34				
Employment (#years)	9.57	7.95	8.88	8.01
	[3.02]	[3.31]	[3.26]	[3.16]
Full time work (#years)	7.73	5.73	5.67	4.97
	[3.54]	[3.61]	[3.80]	[3.47]
Average working hours while employed	2053	1875	1615	1608
	[549]	[643]	[549]	[579]
B. Occupation				
Health	0.21	0.21	0.10	0.10
	[0.14]	[0.14]	[0.30]	[0.30]
Education	0.02	0.03	0.05	0.08
	[0.13]	[0.16]	[0.22]	[0.27]
STEM	0.08	0.06	0.02	0.02
	[0.27]	[0.23]	[0.15]	[0.13]
Business	0.11	0.10	0.08	0.08
	[0.31]	[0.30]	[0.27]	[0.27]
Sales	0.08	0.07	0.08	0.08
	[0.27]	[0.25]	[0.27]	[0.27]
Clerical	0.05	0.06	0.25	0.14
	[0.21]	[0.24]	[0.43]	[0.34]
Manual	0.20	0.20	0.04	0.06
	[0.40]	[0.40]	[0.20]	[0.23]
Social	0.03	0.03	0.03	0.06
	[0.17]	[0.18]	[0.17]	[0.23]
Personal Care	0.00	0.00	0.04	0.05
	[0.07]	[0.08]	[0.20]	[0.21]
Operative	0.31	0.18	0.08	0.03
	[0.46]	[0.39]	[0.28]	[0.16]
Number of observations	4,082	3,093	4,223	3,063

Note: The table documents labor market experience of the NLSY79 and the NLSY97 samples by gender, with the standard deviation provided in square brackets. The work experience is measured by the number of years an individual worked positive hours and full-time (more than 1600 annual hours) between ages 22 and 34. The occupation is measured by dummy variables that takes a value of 1 if the number of years the individual worked in a certain occupation is greater than 5 during ages 22-34.

Table 4: Weights Used to Produced Counterfactual Gender-Specific Wage Distributions

Model	Premarket skills included in z
ψ_0	$\psi_{NLSY79} \times \psi_{ATTR-AFQT79}$
Model 0	ψ(race)
Model 1	ψ(M0 + parents' education)
Model 2	ψ(M1 + family structure)
Model 3	ψ(M2 + AFQT score)
Model 4	ψ(M3 + highest grade completed)
Model 5	ψ(M4 + degree attainment)
Model 6	ψ(M5 + major)
Model 7	ψ(M6 + work experience)
Model 8	ψ(M7 + occupation)

Note: The table presents the weights used in estimating the counterfactual wage distributions based on the DFL method. ψ_0 is the weight I obtained by multiplying the cross-sectional sample weight for the NLSY79 cohort (ψ_{NLSY79}) in the data and the weight to adjust for attrition at age 22 and AFQT nonresponse $(\psi_{ATTR-AFQT})$. The weight used to estimate the counterfactual wage distribution for each specification is calculated by multiplying ψ_0 with the weight provided in each cell depending on the model used.

Table 5: Wage Gains across the NLSY79 and NLSY97 Cohorts by Gender: Data vs. Explained by Changes in Skill Correlates

Panel A	Observed	Counterfactual Log Wage Gain for Men								
Men	Wage Gain	M0	M1	M2	M3	M4	M5	M6	M7	M8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
5%	-0.0043	-0.0160	0.0469	0.0375	0.0139	0.0128	0.0050	0.0058	0.0107	-0.0160
	(0.0007)	(0.0008)	(0.0009)	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
10%	-0.0146	-0.0137	0.0711	0.0490	0.0157	0.0171	0.0042	0.0064	0.0220	-0.0084
	(0.0006)	(0.0008)	(0.0011)	(0.0008)	(0.0008)	(0.0009)	(0.0010)	(0.0009)	(0.0009)	(0.0009)
25%	-0.0720	-0.0162	0.0780	0.0544	0.0292	0.0343	0.0292	0.0292	0.0292	0.0131
	(0.0006)	(0.0007)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0009)
50%	-0.0606	-0.0132	0.1044	0.0760	0.0409	0.0502	0.0488	0.0455	0.0506	0.0353
	(0.0005)	(0.0007)	(0.0009)	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
75%	-0.0312	-0.0105	0.0931	0.0714	0.0459	0.0556	0.0508	0.0442	0.0508	0.0398
	(0.0006)	(0.0007)	(0.0008)	(0.0007)	(0.0007)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
90%	0.0801	-0.0028	0.1157	0.1028	0.0750	0.0927	0.0773	0.0762	0.0861	0.0645
	(0.0007)	(0.0009)	(0.0013)	(0.0009)	(0.0009)	(0.0010)	(0.0012)	(0.0013)	(0.0013)	(0.0012)
95%	0.1019	-0.0014	0.1024	0.0944	0.0574	0.0731	0.0705	0.0731	0.0815	0.0819
	(0.0011)	(0.0013)	(0.0018)	(0.0014)	(0.0013)	(0.0014)	(0.0016)	(0.0018)	(0.0019)	(0.0017)

Panel B	Observed	Counterfactual Log Wage Gain for Women								
Women	Wage Gain	M0	M1	M2	M3	M4	M5	M6	M7	M8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
5%	0.2325	0.0469	0.2157	0.1939	0.1793	0.1942	0.2222	0.1948	0.2157	0.1505
	(0.0026)	(0.0035)	(0.0031)	(0.0034)	(0.0034)	(0.0041)	(0.0030)	(0.0033)	(0.0031)	(0.0031)
10%	0.0977	0.0003	0.0793	0.0779	0.0665	0.0808	0.0853	0.0698	0.0751	0.0428
	(0.0005)	(0.0006)	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
25%	0.0396	0.0004	0.1260	0.1154	0.0963	0.1218	0.1240	0.1093	0.1086	0.0679
	(0.0005)	(0.0007)	(0.0010)	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
50%	0.0698	-0.0013	0.1455	0.1303	0.1180	0.1564	0.1613	0.1665	0.1644	0.1359
	(0.0006)	(0.0007)	(0.0009)	(0.0007)	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0008)	(0.0009)
75%	0.0982	-0.0081	0.1090	0.1031	0.0970	0.1321	0.1306	0.1306	0.1283	0.1084
	(0.0006)	(0.0008)	(0.0010)	(0.0008)	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
90%	0.2062	-0.0063	0.1063	0.0974	0.0909	0.1244	0.1322	0.1276	0.1145	0.1073
	(0.0008)	(0.0009)	(0.0012)	(0.0010)	(0.0010)	(0.0010)	(0.0011)	(0.0011)	(0.0011)	(0.0012)
95%	0.2547	-0.0031	0.0970	0.0721	0.0631	0.1048	0.1171	0.1032	0.0655	0.0721
	(0.0010)	(0.0011)	(0.0014)	(0.0011)	(0.0011)	(0.0011)	(0.0017)	(0.0016)	(0.0016)	(0.0017)

Note: The table documents log wage (1/100 dollar). The wage distribution is conditional on reporting positive wages. Wages are regression standardized to year = 2002 and 13 years of potential work experience. Monetary value is adjusted to 2002 USD by using CPI-U. All statistics are weighted by cross-sectional weights, accounting for attrition by age 22 and AFQT nonresponse. Standard errors (in parentheses) are bootstrapped with 500 repetitions, stratified on NLSY cohort, race, and gender. Units are sampled at the individual level. The sample includes only respondents with observed AFQT scores.

Table 6: Gender Wage Gap in Log Wage Differences

percentile	Data 79 (1)	Data 97 (2)	Model 1 (3)	Model 2 (4)	Model 3 (5)	Model 4 (6)	Model 5 (7)	Model 6 (8)	Model 7 (9)	Model 8 (10)
5%	0.3528	0.1159	0.1840	0.1964	0.1874	0.1714	0.1356	0.1638	0.1478	0.1863
	(0.0026)	(0.0006)	(0.0020)	(0.0023)	(0.0023)	(0.0033)	(0.0017)	(0.0022)	(0.0019)	(0.0019)
10%	0.2386	0.1263	0.2304	0.2096	0.1877	0.1749	0.1574	0.1752	0.1855	0.1874
	(0.0007)	(0.0003)	(0.0011)	(0.0007)	(0.0007)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
25%	0.2904	0.1788	0.2423	0.2294	0.2232	0.2028	0.1956	0.2102	0.2110	0.2356
	(0.0007)	(0.0003)	(0.0011)	(0.0008)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0010)
50%	0.2601	0.1297	0.2190	0.2058	0.1830	0.1540	0.1477	0.1391	0.1462	0.1595
	(0.0007)	(0.0003)	(0.0011)	(0.0007)	(0.0007)	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
75%	0.2219	0.0925	0.2059	0.1902	0.1707	0.1454	0.1420	0.1354	0.1444	0.1533
	(0.0008)	(0.0003)	(0.0010)	(0.0008)	(0.0007)	(0.0008)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
90%	0.2342	0.1082	0.2436	0.2396	0.2183	0.2025	0.1793	0.1828	0.2058	0.1915
	(0.0009)	(0.0005)	(0.0015)	(0.0010)	(0.0009)	(0.0010)	(0.0013)	(0.0014)	(0.0015)	(0.0014)
95%	0.2655	0.1126	0.2708	0.2877	0.2597	0.2338	0.2189	0.2354	0.2815	0.2752
	(0.0013)	(0.0007)	(0.0019)	(0.0012)	(0.0012)	(0.0012)	(0.0019)	(0.0021)	(0.0022)	(0.0020)

Note: The table documents the gender wage gap, measured by the log wage difference, at each percentile of the wage distribution. Columns (1) and (2) display the observed gender wage gap in the data for the NLSY79 and the NLSY97 cohorts, respectively. Columns (3)-(8) depict the counterfactual gender wage gap explained by changes in skill components for men and women in models 3-8, respectively.

Table 7: Decomposition of the Changes in the Gender Wage Gap

Log gender wage gap	Data 79 (1)	Data 97 (2)	Model 3 (3)	Model 6 (4)	Model 8 (5)
Panel A. 10th percentile					
a. Log gender wage gap	0.2386	0.1263	0.1877	0.1574	0.1874
b. Decrease in the gap explained by skill changes		0.1123	0.0509	0.0634	0.0512
c. Men's wage gain			0.0157	0.0064	-0.0084
(skill components)			(-14.0%)	(-5.7%)	(7.5%)
d. Women's wage gain			0.0665	0.070	0.0428
(skill components)			(59.2%)	(62.2%)	(38.1%)
e. Men's wage gain			-0.03	-0.02	-0.006
(skill prices)			(26.9%)	(18.7%)	(5.5%)
f. Women's wage gain			0.03	0.03	0.05
(skill prices)			(27.8%)	(24.8%)	(48.9%)
Decomposition					
g. Explained by skills			45.3%	56.5%	45.6%
h. Explained by prices			54.7%	43.5%	54.4%
Panel B. 50th percentile					
a. Log gender wage gap	0.2601	0.1297	0.1830	0.1391	0.1595
b. Decrease in the gap explained by skill changes		0.1304	0.0771	0.1211	0.1006
c. Men's wage gain			0.0409	0.0455	0.0353
(skill components)			(-31.4%)	(-34.9%)	(-27.1%)
d. Women's wage gain			0.1180	0.1665	0.1359
(skill components)			(90.5%)	(127.7%)	(104.2%)
e. Men's wage gain			-0.1015	-0.1061	-0.0959
(skill prices)			(77.8%)	(81.3%)	(73.5%)
f. Women's wage gain			-0.0483	-0.0967	-0.0661
(skill prices)			(-37.0%)	(-74.2%)	(-50.7%)
Decomposition					
g. Explained by skills			59.1%	92.8%	77.1%
h. Explained by prices			40.8%	7.2%	22.9%
Panel C. 90th percentile					
a. Log gender wage gap	0.2342	0.1082	0.2183	0.1828	0.1915
b. Decrease in the gap explained by skill changes		0.1260	0.0159	0.0514	0.0427
c. Men's wage gain			0.0750	0.0773	0.0645
(skill components)			(-59.5%)	(-60.4%)	(-51.2%)
d. Women's wage gain			0.0909	0.1276	0.1073
(skill components)			(72.1%)	(101.2%)	(85.2%)
e. Men's wage gain			0.0051	0.0039	0.0156
(skill prices)			(-4.1%)	(-3.1%)	(-12.4%)
f. Women's wage gain			0.1152	0.0786	0.0989
(skill prices)			(91.4%)	(62.4%)	(78.5%)
Decomposition					
g. Explained by skills			12.6%	40.8%	33.9%
h. Explained by prices			87.4%	59.2%	66.1%

Note: Panel A, B, and C summarize the decomposition results for the change in the log gender wage gap across two cohorts at the 10th, 50th, and 90th percentiles of the wage distribution, respectively. Row (a) presents the log wage difference between men and women, row (b) presents the change in the log gender wage gap across two cohorts. In Panels B and C, rows (c) and (d) present the log wage gain for men and women, respectively. Row (e) presents the extent to which changes in skill components in Models 3,6, and 8 can account for the actual change in the gender wage gap across two cohorts, measured by the changes in the log gender wage gap and its percentage out of the total change across two cohorts. Finally, row (f) presents the extent to which changes in skill prices across two cohorts can account for the actual change in the gender wage gap across two cohorts.