Gender Divergence in Premarket Skill Acquisition and Wage Inequality

Sunha Myong*

April 12, 2021

Abstract

Based on the NLSY79 and NLSY97, I document changes in premarket skills and wage distribution across two cohorts, and quantify to what extent changes in premarket skills can explain the recent trend in the wage distribution. I apply the DiNardo, Fortin, and Lemieux decomposition method (DiNardo et al. (1996)) to estimate the relative importance of the change in premarket skill and the change in skill prices in explaining the trend. I find a substantial gender divergence in premarket skill acquisition and wage gain across two cohorts. Women gain substantially more premarket skills than men at all levels of the wage distribution, and the gender difference is greatest in the middle range of the wage distribution. Women's wages increased at all levels of the wage distribution, but men's wages decreased except at the top quartile of the wage distribution. The decomposition result suggests that the change in premarket skills can explain most of the decrease in the gender wage gap in the middle range of the wage distribution, and explain about half of the decrease at both ends of the wage distribution. Also, the change in premarket skills has different implications on wage inequality for men and women; it widened the upper-tail wage inequality for men, but the opposite is true for women. I argue that gender-specific shifts in demand-side factors can contribute to the diverging trend by gender.

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

1 Introduction

Heterogeneous premarket skills explain substantial variations in wage inequality.[1] The extent to which premarket skills explain the wage inequality can vary across time, depending on the time specific technology, institution, and labor demand for different skills.[2] Previous studies extensively investigated factors that explain the rise in U.S. wage inequality between 1980 and mid 2000s.[3] In this paper, I quantify the role of changes in premarket skills in explaining the recent trend in the wage inequality between 1990s and 2010s by comparing the premarket skills and the wage distribution of two particular birth cohorts: 1957-1964 birth cohorts from the National Longitudinal Study of Youth 1979 (NLSY79) and 1980-1984 birth cohorts from the National Longitudinal Study of Youth 1997 (NLSY97).[4]

As discussed in Autor et al. (2005) and Lemieux (2006), both changes in the skill composition of the population and changes in the relative price for skilled workers determine the evolution of the wage distribution over time. There are several reasons to re-examine the role of shifts in premarket skills in explaining the wage distribution of the recent cohorts.

First, I find that the NLSY79 (the 1960s birth cohorts) and the NLSY97 cohorts (the 1980s birth cohorts) exhibit notable differences in their premarket skill acquisition and wage inequality. For instance, the school to work transition presents greater heterogeneity for the recent cohort. Not only the recent cohort have substantially higher educational attainment, but also they take longer time to finish education and work more at low-paying jobs during the college period (Bound et al., 2012; Ashworth et al., 2017). Also, men and women show markedly different trends in skill acquisition and the wage distribution across two cohorts: i) men and women had similar educational attainment

[1] Taber and Vejlin (2020) find that among four components of the Roy model—variation in premarket skills, search frictions, compensating differentials, and human capital accumulation on-the-job, variation in premarket skills is the most important factor that explains the wage inequality, accounting for between 59% and 82% of wage inequality.

[2] Premarket skills in this paper refer to premarket skills formed before the completion of education, including work experience during schooling period.

[3] Literature finds that computerization and automation play a crucial role in explaining the pronounced rise in the college wage premium since 1980, despite the substantial increase in relative supply of college-educated workers during the same period (Goldin and Katz, 2009; Autor et al., 2008).

[4] Throughout this paper, I use "the 1960s (the 1980s) birth cohorts" to refer to the NLSY79 (NLSY97) sample.

in the NLSY79 cohort, but for the NLSY97 cohort, the share of women with a four-year college degree as of age 35 is 0.11 (38%) higher than that of men; ii) both men and women in the recent cohort increased labor supply during their early 20s, but the share of individuals working at low-paying jobs during the college period increased more for men than women; iii) women in the recent cohort have higher wage than the older cohort at all levels of the wage distribution, but the wage gain of men is negative except for those in the top quartile of the wage distribution.

Second, the labor market after mid 2000s exhibits distinctive features. For instance, Beaudry et al. (2016) find that the demand for cognitive tasks decreased substantially after the 2008 financial crisis.[5] Castex and Kogan Dechter (2014) finds that the return to cognitive skill decreased and the return to education increased between the NLSY79 and the NLSY97. Deming and Kahn (2018) finds that the labor market increasingly rewards social skills, and the jobs requiring social interaction grew substantially between 1980 and 2012. Cortes et al. (2020) show that changes in the relative importance of social tasks explains gender-specific changes in occupational sorting between 1980 and 2016. Given the substantial changes in demand side factors in recent years, it is important to account for changes in skill prices to evaluate the role of premarket skills in recent wage inequality.

To quantify the extent to which changes in premarket skills can explain the recent change in wage distribution, I apply the DiNardo, Fortin, and Lemieux (DFL) decomposition method (DiNardo et al. (1996)) to obtain the counterfactual wage distribution of the NLSY79 cohort when the distribution of skill components is replaced with that of the NLSY97 cohort. A closely related study is Altonji et al. (2012) who apply the DFL method to quantify the extent to which changes in premarket skills across two cohorts explain the wage gain of the NLSY97 cohort relative to the NLSY79 cohorts.[6] I extend Altonji et al. (2012) in the following ways. First, I improve the

[5] Beaudry et al. (2016) claim that the technological progress can exhibit a boom-and-bust cycle in its demand for cognitive skills: the demand for cognitive-task workers is high and growing which peaks during the key investment periods following a technological revolution; but once the technology reaches a maturity stage and new capital is in place, the demand for cognitive-task workers can decrease.

[6] By matching two cohorts based on various premarket skills—cognitive skill (AFQT score), parental characteristics, educational attainment, and school-to-work transition as of age 22, Altonji et al. (2012) find that changes in premarket skills between the NLSY79 and the NLSY97 cohorts would substantially increase the wage inequality of

measurement of educational attainment and the school-to-work transition by using more recent data for the NLSY97 cohort. I argue that due to the increasing heterogeneity in the school-to-work transition across the two cohorts, it is important to track educational attainment and early work experience until the early 30s to document the skill acquisition of the youth. Because the NLSY97 cohort has a longer schooling period and a more complex school-to-work transition compared to the NLSY79 cohort, using skill measures as of age 22 substantially understates the increase in educational attainment across the two cohorts, the gender difference in premarket skill acquisition, and the heterogeneity in premarket skills among college educated individuals. Second, since the wage data for the NLSY97 cohort during their mid-30s have become available, I can compare the counterfactual wage gain obtained from the DFL method to that in the actual data. This allows me to quantify respective roles of supply side factor—i.e., premarket skills—and demand side factors—i.e., changes in the price for skilled-labor and shifts in employment share across occupations—in explaining the trend in the wage distribution across the two cohorts.

Focusing on population including both men and women, I find that changes in premarket skills explain 5-9% wage gains of the recent cohort relative to the old cohort across different percentiles of the wage distribution. The increased premarket skill across two cohorts plays an essential role for the recent cohort to have non-negative wage gain, as it could counteract the substantial reduction in skill prices in the middle range of the wage distribution. Without the increase in premarket skill across two cohorts, the NLSY97 cohort would have had 7-8% lower wage rates than the NLSY79 cohort in the 30-60 percentiles of the wage distribution. Among various premarket skills, changes in parental education and family structure have the greatest marginal contribution to the wage gain across two cohorts, by 4-7%. However, if I use the the degree attainment as of age 35 as a measure for education, the marginal effect of education on the wage gain of the recent cohort is also significantly positive (2-5%) even after accounting for changes in family background factors.[7]

the NLSY97 cohort at their prime ages.

[7]This finding is different from Altonji et al. (2012) who find that the marginal contribution of education on wage gain across two cohorts is not significant after accounting for changes in parents' education and family structure, when education is measured by the highest grade completed as of age 22.

Finally, changes in early labor market experience during the college period, as measured by working hours during age 18-22 and occupation at age 22 and 25, reduces the wage gain of the recent cohorts by 1-2%.

On the other hand, I find substantial gender differences in to what extent changes in premarket skills can explain the wage gain across two cohorts. First, women's wage gain associated with changes in premarket skill is greater than that of men's at all levels of the wage distribution. For women, changes in premarket skill significantly increases the wage rates at all levels of the wage distribution, accounting for 5-18% increase in the wage rates across two cohorts. On the other hand, men's wage gain associated with changes in premarket skills across two cohorts is not significantly different from zero except at the top end of the wage distribution.[8] Second, for women, the wage gain associated with changes in premarket skill is greatest in the middle range of the wage distribution , whereas it is greatest at the top end of the wage distribution for men. As a result, changes in premarket skills across the two cohorts explain most of decrease in the gender wage gap in the middle range of the wage distribution, but changes in premarket skill can explain about half of the decrease in the gender wage gap at both ends of the wage distribution.

Next, I evaluate to what extent changes in premarket skill can explain the evolution of wage inequality across two cohorts—increasing upper-tail wage inequality and decreasing lower-tail wage inequality. First, focusing on the wage distribution of the population including both men and women, the change in the premarket skills do not play a significant role in explaining the change in the wage inequality between the NLSY79 and NLSY97 cohorts. Instead, the diverging pattern of lower- and upper-tail wage inequality is largely explained by changes in the demand side factors, which is consistent with Autor et al. (2008). [9] I also find substantial gender differences in the role of premarket skills in explaining the wage inequality. For men, changes in educational attainment contribute to widening upper-tail wage inequality. The opposite is true for women. This finding

[8]Once I include early labor market experiences as premarket skills, changes in premarket skill do not singificantly increase men's wage even at the top end of the wage distribution.

[9]This finding is consistent with Autor et al. (2008) who find the change in skill prices is main driving force explaining the diverging upper- and lower-tail wage inequality between 1963 and 2005.

is mainly driven by the diverging pattern of skill acquisition of men and women in the middle range of the wage distribution. A large increase in educational attainment by middle-women plays a significant role in counteracting the increasing upper-tail wage inequality associated with shifts in skill prices. However, the middle-range men do not accumulate as much human capital as comparable women before they enter the labor market.

Given the pronounced gender divergence in their skill acquisition and wage distribution across the NLSY79 and NLSY97 cohorts, I further investigate the source of gender specific change in demand side factors which could have affected skill acquisition of men and women. First, I find that the skill price for men and women changes differently across the two cohorts; the return to a four-year college degree is higher for women in both cohorts, but the gap doubled across the two cohorts.[10] The gender gap in the return to a baccalaureate degree increased from 0.097 to 0.187, and the gender gap in the return to a master's degree or above increased from 0.223 to 0.275. Gender-specific shifts in skill prices can explain the intensified gender differences in educational attainment in recent years. Second, there has been gender-specific change in occupational sorting, especially in the middle range of the wage distribution. Women move up from middle to upper-middle occupations, whereas men move down from middle to lower-middle occupations. The rapid growth in female-dominant occupations such as health care, education, and occupations that involve community/social service and media/communication (hereafter "health, education, and social occupations") plays an important role in creating gender-specific shifts in the middle range of the occupation. In particular, based on job-posting data from Atalay et al. (2020),[11] I find that the share of jobs that require a college degree increased rapidly for occupations with an increasing share of female. Therefore, shifts in demand across occupations and changes in educational requirement jointly explain gender differences in educational attainment, especially in the middle

[10] To facilitate comparison with previous findings based on the NLSY79 and NLSY97 cohorts, I closely follow Castex and Kogan Dechter (2014) when estimating the wage regression, using more recent data for wage and educational attainment for the NLSY97 cohort.

[11] Atalay et al. (2020) construct a dataset from the text content of 7.8 million job ads from the *Boston Globe*, *New York Times*, and *Wall Street Journal* and provide the skill requirements/task contents of each job across 1945-2000. I use their public data which aggregate skill requirement and other information at 3-digit standard occupational classification (SOC) level.

range of the wage distribution.

This paper contributes to the literature on wage inequality. First, I document more recent trends in wage inequality between the 1960s and 1980s birth cohorts. Because the 1980s birth cohorts enter the labor market after mid 2000s, comparing the skill acquisition and labor market outcome of the 1960s and the 1980s birth cohorts can shed light on how the recent changes in the labor market influence the human capital formation and the labor market outcome of the recent cohort. Second, I show that heterogeneous school to work transition explains substantial wage variation at age 35. Third, I evaluate to what extent shifts in the skill components of the 1960s and the 1980s birth cohorts explain the changes in the observed wage distribution, thereby extending the findings of Altonji et al. (2012).[12]

This paper is also related to the previous studies which find that technological change plays an important role in closing the gender wage gap. Galor et al. (1996) and Welch (2000) claim that gender differences in brains and brawn are important to explain the narrowing gender gap in the labor market. Weinberg (2000) provides empirical evidence that a decline in emphasis on physical skills following computerization has increased the demand for female workers. Based on the task-based Roy model, recent studies find evidence that the shift in demand from manual tasks to analytical tasks, increasing importance of social skills in cognitive and high-wage occupation, and decreasing returns to motor skills play important roles in the narrowing gender wage gap during 1980-2000 (Black and Spitz-Oener, 2010; Bacolod and Blum, 2010; Yamaguchi, 2018; Cortes et al., 2020). Most of these studies take education as exogenously given characteristics and focus on the role of skill prices in the narrowing gender wage gap. However, as discussed in Yamaguchi (2018), the growth of women's general and cognitive skills relative to men plays an important role in explaining the narrowing gender wage gap after 2000. In this paper, I claim that

[12] A large literature investigates the source of increasing earnings inequality over the last few decades and evaluates the role of skill components and skill prices in explaining the trend (Autor et al., 2008; Lemieux, 2006; Autor et al., 2003; Juhn et al., 1993). For instance, Lemieux (2006) finds that due to the greater within-group wage inequality for college-educated workers, the increasing share of college educated-workers can explain most of the growing wage inequality between 1973 and 2003. In contrast, Autor et al. (2008) claim that changes in skill composition cannot explain the steady increase in upper-tail wage inequality (90-50th wage inequality) during 1963-2005, and emphasize the role of changes in relative skill price.

increased demand for social skills has closed the gender wage gap for recent cohort by not only changing gender-specific skill prices but also by affecting the skill acquisition of men and women before they enter the labor market.

The paper proceeds as follows. Section 2 presents facts regarding the skill acquisition and wage inequality of the NLSY79 and NLSY97 cohorts. Section 3 describes the econometric methods, and Section 4 discusses the data. Section 5 explains the estimation, and Section 6 presents the counterfactual analysis to discuss the role of changing skill components on the wage inequality of men and women and the gender wage gap. Section 7 discusses the mechanism behind the gender-specific change in skill acquisition, and Section 8 concludes.

2 Data

The NLSY79 and NLSY97 are longitudinal data that follow population-representative samples of American youth. The NLSY79 originally consists of 12,686 individuals who were born between 1957 and 1964 and were 14-22 years old in 1979. The NLSY97 originally consists of 8,984 individuals who were born between 1980 and 1984 and were 12-18 years old in 1997. The NLSY79 follows the sample from 1979 to 2016, and the NLSY97 follows the sample from 1997 to 2017. The NLSY79 respondents are 52-60 at the time of the 2016 survey, and the NLSY97 respondents are 32-38 at the time of the 2017 survey.

I use 1979-2004 survey data for the NLSY79 cohort, and 1997-2017 survey data for the NLSY97 cohort to document demographic characteristics, educational attainment, and school-to-work transition of each NLSY cohort.[13] Because the NLSY97 does not include a military sample or white poor supplemental samples, I exclude those supplemental samples from the NLSY79 sample, which reduces the number of individuals in the sample 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

[13] To provide results comparable to the literature, I closely follow Altonji et al. (2012) regarding sample construction. The key difference from the sample construction in Altonji et al. (2012) is the use of more recent survey years for the NLSY97 cohorts to construct school-to-work transition variables and wage at around age 35.

NLSY97 from 8,984 to 8,901, whereas it does not affect the NLSY79 sample. I adjust the weight for migrants depending on the age of arrival by putting zero weight on those arrived in the U.S. after age 16. Dropping those individuals reduces the sample size of the NLSY79 to 9,665, while it does not affect the sample size of the NLSY97. I restrict the sample to individuals who have valid information at age 22, which reduces the sample size to 9,230 for the NLSY79 and 7,557 for the NLSY97. I further drop individuals with missing AFQT scores, which leaves 8,848 and 6,028 individuals in the final sample for the NLSY79 and NLSY97 sample, respectively.[14]

I use the cross-sectional sample weight for the original sample (in the first survey for each 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 characteristics of individuals in the first year survey.[15] The weight is divided by the predicted probability of attrition/missing AFQT score. The same weights are used in (Altonji et al., 2012).

For the wage data, I use the 1988, 1990, 1992, and 1994 surveys for the NLSY79 and 2011, 2013, 2015, and 2017 surveys for the NLSY97. I closely follow Altonji et al. (2012) to get the wage distribution of the population. Wages are regression standardized to year 2002 and 13 years of experience. The main reason for choosing 13 years of experience is to closely reflect the average years of work experience after the schooling period for the NLSY97 cohort who is aged 32-38 in the 2017 survey.[16] Details on the standardization are discussed in section 4.

Table A1 shows summary statistics on the demographic characteristics of the sample for the NLSY79 and NLSY97 cohorts. The NLSY97 cohort consists of fewer White (8% fewer) and more Black (1% more) and Hispanic (7% more) compared with the NLSY79 cohort. Parents of

[14]The slight difference in sample size from Altonji et al. (2012) is driven by an update in the raw data for the highest grade completed.

[15]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 mother and father, whether the individual lived with mother/father/both at age 14, whether the individual lived in urban/standard metropolitan statistical area (SMSA), attitude displayed at interview, whether the 1997 interview was done in 1998 instead.

[16]Altonji et al. (2012) standardized the wage of the NLSY79 cohorts to 23 years of experience to discuss the wage distribution of workers at their prime wage, but did not use wage data for the NLSY97 cohorts.

the NLSY97 cohort have higher educational attainment than the NLSY79 cohort; the highest grade completed by the mother and father increased by 1.35 and 1.04 years, respectively. Although parents' education increased significantly over time, more youth live with only one parent at age 14. The proportion of youth who live with both mother and father decreased by 21%, whereas the proportion of individuals who live with only the mother increased by 17%. As discussed in Altonji et al. (2012), changes in the family structure are among the most notable differences between the two cohorts. AFQT scores are slightly higher for the NLSY97 cohort than the NLSY79 cohort.[17] Table A2 summarizes demographic characteristics by gender for the NLSY79 and NLSY97 cohorts. The change in demographic characteristics does not show substantial gender differences, except for the AFQT score.[18]

2.1 Skill Acquisition and Wage Inequality of the NLSY79 and NLSY97 Cohorts

I first document trends in skill acquisition of the NLSY79 and the NLSY97 cohorts. I find that the NLSY79 (the 1960s birth cohorts) and the NLSY97 cohorts (the 1980s birth cohorts) exhibit notable differences in their premarket skill acquisition.

Fact 1: Not only the NLSY97 cohort has substantially higher educational attainment, but also spends longer time to finish education and works more at low-paying jobs during the college period.

Table 1 tabulates educational attainment and school-to-work transition variables of the NLSY79 and NLSY97 cohorts. The share of the population who received a baccalaureate degree (BA) as of

[17] To construct a comparable measure for the cognitive ability across the two cohorts, I follow Altonji et al. (2012) and standardize the AFQT score based on the mapping provided by (Segall, 1997).

[18] The mean of women's AFQT score is 0.016 standard deviation higher than the mean of men's score for the NLSY79 cohort. The gap widens substantially over time so that for the NLSY97 cohort, the mean of women's AFQT score is 0.098 standard deviation higher than the mean of men's AFQT score. The increase in the AFQT score for the whole sample of the NLSY97 cohort is well documented in Altonji et al. (2012). It is worth noting that there is a substantial divergence between men and women in the trend of cognitive skills.

age 35 increased from 0.22 to 0.35 across the two cohorts.[19] The corresponding share for an associated degree increased from 0.09 to 0.14, and that for master's degree or above (MA) increased from 0.05 and 0.12, respectively.[20] It is also worth noting that the trend would be documented differently if I use other measures for the educational attainment. There is no significant cohort difference in the share of BA degree holders as of 22 (Panel B), while the cohort gap widens substantially at age 27 (0.09) and at age 35 (0.13). Similarly, the recent cohort has 0.39 more schooling years as of age 22 and has 0.89 more schooling years as of age 35 (Panel C).

Panels E-G of Table 1 summarize the school-to-work transition variables of the two cohorts. Students work more during college in the recent cohort. The average weekly working hours while enrolled in college before age 22 increased by 3.35. The increased working hours during college might be related to the delayed school-to-work transition for the recent cohort. As shown in Panel C, the NLSY97 cohort takes longer time to achieve the same academic degree. The mean age to obtain a four-year college degree is 22.90 for the NLSY79 cohort and 24.42 for the NLSY97 cohort, an increase of 1.31 years. Increasing working hours while in college and delayed graduation are consistent with previous studies (Bound et al., 2012; Ashworth et al., 2017). I also find that the type of jobs individuals have during the schooling period dramatically changed as well.[21] The recent cohort works more at low-skilled jobs during early ages.[22] In Panel F of Table 1, I find that the proportion of individuals in manual (sales/clerical) occupations at age 18 increased by 3.9% (6.9%) across the two cohorts. Manual and sales/clerical occupations have the lowest wage rates

[19] This finding implies that unlike the stagnation in the college completion rate for the 1970 birth cohort (Goldin and Katz (2009)), the supply of skilled labor increased substantially for the 1980s birth cohort. Goldin and Katz (2009) document historical changes in the educational attainment of the U.S. population for 1930-1975 birth cohorts. The stagnated supply of college-educated workers for the 1970s birth cohort is considered to contribute to the recent increase in wage inequality.

[20] The finding is comparable Altonji et al. (2012) who find the share of population whose highest grade completed is at least 14 years at age 22 increased from 31.78% to 43.18% across the NLSY79 and NLSY97 cohorts. In my sample, I find that the share of individuals who receive either AA, BA, or MA or above as of 35 increased from 31% to 49%.

[21] The increasing time to get BA degree in American youth has been documented in Bound et al. (2012) between cohort high school graduation in 1972 (NLS72) and in 1992 (NELS:88), and two NLSY cohorts in Ashworth et al. (2017). However, they do not document the occupational sorting during the schooling period.

[22] To define low-skilled jobs, I use the wage data on all workers aged 18-37 in each NLSY sample and tabulate the mean wage by occupation in Table A4. Details on sample construction for wage distribution are described in section 2.

for both cohorts.[23] Individuals move from low-paying occupation to higher-paying occupations as they age, but the share of individuals working at low-paying occupations at age 22 is still 6.5 percentage points higher for the recent cohort.

Fact 2: Skill acquisition during the schooling period exhibits substantial gender divergence across two cohorts.

Figure 1 documents the age-specific share of individuals received an AA, BA, MA, or above by gender. I find that the share of individuals with at least a BA degree as of age 35 increased by 8 percentage points, from 22% to 30%, for men, and by 19 percentage points, from 22% to 41%, for women.[24] Men and women had similar college graduation rates in the NLSY79 cohort. In contrast, for the NLSY97 cohort, the share of women with a four-year college degree is 0.11 (38%) higher than that of men.[25] The gender gap widens after age 22, which suggests the importance of measuring the completed history of education to document gender-specific trends in educational attainment.[26] The gender reversal in educational attainment was documented by Goldin and Katz (2009).[27] The above finding suggests that the increase in educational attainment of women relative to men is substantially greater for the 1980 birth cohorts than earlier cohorts.

The school-to-work transition also presents substantial gender differences. Figure 2 shows

[23] Manual occupation includes jobs related to food preparations and serving, cleaning and building service, construction/trades/extraction, and protective service. Social occupation includes social scientists and related workers, counselors, community/social workers, religious workers, lawyers, judges, and legal support workers, and media and communication workers.

[24] Focusing on only male population from the NLSY79 and NLSY97 cohorts, Ashworth et al. (2017) finds less dramatic increase (3-6% increase) in the share of men who complete a college (including 2- and 4-year colleges) as of age 29. The difference could be partially explained by differences in sample selection, the use of sample weight, and nontrivial proportion individuals who receive 2- or 4-year college degree after age 29 (Figure 1).

[25] The gender gap not only increased for the four-year college degree. The share of individuals with a MA or above increased from 6% to 9% for men and from 5% to 14% for women. The share of individuals with an AA increased from 7% to 12% for men and from 10% to 17% for women.

[26] Table A3 tabulates more variables on the educational attainment and school-to-work transition variables of the NLSY79 and NLSY97 cohorts by gender. It shows that the college dropout rate, as measured by the share of individuals who ever attended a college but did not obtain any degree, decreased more for women (by 16%) than men (by 5%). Also, the age at obtaining a BA increased more for men (by 1.42 years) than women (by 1.24 years).

[27] Men had higher educational attainment, as measured by the four-year completion rate, than women before the 1960 birth cohort, and the trend was reversed after the 1960 birth cohort: For the 1970 birth cohort, the college completion rate was 5% higher for women than men.

gender-specific changes in occupational sorting across two cohorts. First, the share of individuals working at low-paying jobs increased more for men than women across the two cohorts. At age 18, the share of men working in a low-paying occupation (manual of sales/clerical) increased by 13 percentage points across the two cohorts, whereas the corresponding number for women is 8 percentage points. The occupational shift from low-paying to middle- and high-paying jobs between age 18 and 22 is greater for women than men. Focusing on the occupational share at age 22, women's share in a low-paying occupation is similar across two cohorts (increased by 1%), but the men's share at low-paying jobs increased by 11%.

2.2 Trends in Wage Distribution

In this section, I document the wage gain across two cohorts at each percentile of the wage distribution and changes in wage inequality. Figure 3 shows log wage gain of the NLSY97 cohort relative to the NLSY79 cohort at each percentile of the wage distribution for the whole population including men and women.[28] Table 3 presents the distribution of wage for each NLSY cohort and the statistical significance of the log wage difference between the two cohorts. First, the wage gain across two cohorts is non-negative at all levels of the wage distribution. Second, the wage gain across two cohorts is U-shaped; the wage between the 20th and 65th percentiles of the distribution did not change significantly across the two cohorts, whereas the wage of the NLSY97 cohort is 12% (17%) higher than that of the NLSY97 cohort at the top 10th (5th) percentile of the distribution and higher by 6% (5%) at the bottom 10th (5th) percentile. Third, the U-shaped wage gain implies an increasing upper-tail wage inequality (90th-50th wage gap) and a decreasing lower-tail wage inequality (50th-10th wage gap) across two cohorts. This finding is consistent with Autor et al. (2008) who document diverging paths of upper-tail and lower-tail wage inequality between 1963 and 2005.[29]

[28] Note that I multiply the inverse of the frequency of valid wage observations in the weight following Altonji et al. (2012).

[29] Altonji et al. (2012) focus on full-time, full-year workers aged 16 to 64 from the March Current Population Survey (CPS). The U-shaped wage gain between the 1960s and 1980s birth cohorts is somewhat different from Autor et al.

However, once I look at the wage distribution separately for men and women, the evolution of the wage distribution across two cohorts shows substantial gender difference.

Fact 3: Women's wage increased at all levels of the wage distribution, whereas men's wage did not increase except for the top quartile of the wage distribution.

Panel (b) of Figure 3 plots the difference in the log wage rate between the NLSY97 and NLSY79 cohorts at each percentile of the wage distribution for men (solid line with plus marks) and women (solid line). The wage distribution of men and women in each cohort, and the statistical significance of the log wage difference between the two cohorts for men and women are presented in Table 4.

Women's wage increased at all levels of the wage distribution across the two cohorts. The wage gain of women is greater at both ends of the wage distribution; It increased by 0.100 log point at the 10th percentile of the wage distribution and by 0.182 log point at the 90th percentile at 1% significance level. Although the wage gain is smallest at around the 25th percentile of the wage distribution, the wage significantly increased by 0.047 log point across the two cohorts. In contrast, men's wage did not significantly increase except for those at the top quartile of the wage distribution. The decrease in men's wage across the two cohorts is largest in the middle range of the wage distribution, decreasing by 0.069-0.058 log point in the 25th-50th percentiles.

The above gender specific wage gain across two cohorts implies substantial decrease in the gender wage over time. In Panel (c) of Figure 3, I document the gender wage gap at each percentile of the wage distribution for the NLSY79 (solid line) and NLSY97 (dashed line) cohorts. For the NLSY79 cohort, the gender wage gap is greater than 0.2 log point (20%) at all levels of the wage distribution. The gender wage gap decreased by about 0.1 log point at all levels of the wage distribution between the NSLY79 and NLSY97 cohorts.[30]

(2008), who find a monotonically increasing wage gain along the percentile of the wage distribution between 1963 and 2005. However, the diverging pattern between upper- and lower-tail wage inequality is consistent.

[30]Focusing on workers aged 25-64 who were full-time, non-farm, wage and salary workers and worked at least 26 weeks during the preceding year in the Panel Study of Income Dynamics (PSID), Blau and Kahn (2017) documents that the gender wage gap decreases much more slowly at the top end of the wage distribution between between 1980-2010 period. Thus, changes in the age-specific gender wage gap across cohorts might have different trends from the gender wage gap shown in the entire working population over time.

2.3 Discussion

To what extent can changes in premarket skill components of labor force explain the evolution of the wage distribution between the 1960s and 1980s birth cohorts? As discussed in Autor et al. (2005) and Lemieux (2006), both changes in the skill composition of the population and changes in the relative price for skilled workers determine the evolution of the wage distribution over time. The literature focuses more on the role of skill prices in the recent wage distribution. However, given the substantial changes in skill acquisition of youth between the 1960s and 1980s birth cohorts, it is worth re-examining the role of the changing skill components of the population in determining the wage distribution.

I also find that skill acquisition and the wage gain across two cohorts exhibits substantial gender divergence across the NLSY79 and NLSY97 cohorts. Because education and the school-to-work are endogenously determined, understanding the source of gender-specific trend in premarket skill acquisition can shed light on how current educational policies would affect the wage inequality for men and women and the gender wage gap.

3 Econometric Methods

To quantify the extent 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 premarket skills of the NLSY97 cohort, while facing the same wage function as the NLSY79 that determines the relationship between premarket skills and the adult wage. I apply the density reweighting procedure introduced by DiNardo et al. (1996). I reweight the NLSY79 sample to have the same distribution of premarket skills as the NLSY97 sample and evaluate how the wage distribution changes in the reweighted NLSY79 sample compared with the sample prior to reweighting.[31]

[31] I closely follow Altonji et al. (2012) in describing the econometric method.

3.1 Constructing the Counterfactual Wage Distribution Based on the DFL Method

Let z be a vector of observed premarket skills 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 cohorts can be written as

$$
\begin{aligned}
f(w^{97}|z_{97}, 97) - f(w^{79}|z_{79}, 79) = \quad & \underbrace{\left[f(W^{97}(z,u)|z_{97}, 97) - f(W^{79}(z,u)|z_{97}, 97)\right]}_{\text{(i) wage difference explained by changes in the wage determination function}} \\
+ \quad & \underbrace{\left[f(W^{79}(z,u)|z_{97}, 97) - f(W^{79}(z,u)|z_{79}, 79)\right]}_{\text{(ii) wage difference explained by changes in the premarket skill components}}
\end{aligned} \tag{1}
$$

The first term on the right-hand side (part (i)) captures the difference in the log wage rate between the 1979 and 1997 cohorts associated with the change in the wage determination function W^j over time. As widely discussed in literature, skill-biased technological change—i.e., changes in the task requirement across and within occupation—can fundamentally change how the real wage is determined, given the (observed/unobserved) characteristics of the individual.

The second term (part (ii)) captures the wage difference of the two cohorts driven by differences in (z, u) over time, 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 1979 and 1997 cohorts, respectively. Then $g(u|z, 79) = g(u|z, 97)$ holds.

As noted in Altonji et al. (2012), the above assumption is difficult to hold exactly. Changes in skill price 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)). On the other hand, the assumption might fail due to changes in college cost and financial aid policies (Lochner and Monge-Naranjo (2011)) or discrimination in the labor market (Hendricks and Schoellman (2014)). However, as discussed in Altonji et al. (2012), it is not possible to directly test assumption 1 because u is not observed. Thus, I follow the approach by Altonji et al. (2012) and discuss the counterfactual wage distribution of the NLSY79 cohort that is valid under the 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 (1), 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:

$$\begin{aligned} f(w^{79}|z,97) &= \int f(w^{79}|z,97) f(z|97) dz \\ &= \int f(w^{79}|z,79) f(z|79) \psi(z) dz, \end{aligned} \tag{2}$$

where

$$\psi(z) = \frac{f(z|97)}{f(z|79)} = \frac{p(97|z)}{p(79|z)} \frac{p(79)}{p(97)}. \tag{3}$$

and $p(97|z)$ and $p(79|z) = 1 - p(97|z)$ are the 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 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 (1). 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.

3.2 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:

$$\begin{aligned} f(w^{79}|97) - f(w^{79}|79) = & \underbrace{\int f(w^{79}|z_1, z_2, 79)\left[f(z_1, z_2, 97) - f(z_2|z_1, 79)\psi(z_1)f(z_1|79)\right]dz}_{\text{(a) SME of } z_1} \\ & + \underbrace{\int f(w^{79}|z_1, z_2, 79)\left[f(z_2|z_1, 79)\psi(z_1)f(z_1|79) - f(z_2|z_1, 79)f(z_1|79)\right]dz,}_{\text{(b) SME of } z_2} \end{aligned} \tag{4}$$

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 1979 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 by part (a) (SME of z_2). The SME of z_1 has two effects: the direct effect from different distributions of z_1 between the two cohorts, and the indirect effect from different distributions of z_2 that is induced by the relationship between z_2 and z_1 for the 1979 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 that the conditional density of z_2 on z_1 for the 1979 cohort implies. 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 1997 cohort as when it affects z_2 for the 1979 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 variables which are difficult to clearly identify the timing of the event, I first account for variables widely used in the literature (i.e., degree attainments), then evaluate the additional effect of controlling for the variable that is not well discussed in the literature before (i.e., working hours during college). Despite the above issue, the decomposition does not require a parametric assumption for the wage function $W(z,u)$ and can be applied to the entire wage distribution. Also, different from the Oaxaca-Blinder decomposition, which focuses on the decomposition at the mean value of the variables, the SME based on the DFL method allows me to investigate the heterogeneity in decomposition results across the wage distribution.

4 Estimation

By applying the DFL method, Altonji et al. (2012) find that changes in the characteristics of the population as of age 22 would increase the wage of the NLSY97 cohort by 3-12 log points relative to the NLSY79 cohort at each percentile of the wage distribution at their prime ages (with 23 years of potential experience). They find that the wage gain associated with shifts in characteristics is greater at the top end of the wage distribution; the wage gain is less than 3% below the 20th percentile, 5% between the 25th and 85th percentiles, and 7%-12% at the top 5th-10th percentiles of the distribution.

In this section, I conduct an exercise similar to Altonji et al. (2012), but focusing on wage with 13 years of potential experience. The main reason for choosing an earlier ages is to compare the counterfactual wage distribution obtained from the DFL method to the actual wage distribution observed in the data. Because the NLSY97 cohort is aged 32-38 in the 2017 survey, I can quantify to what extent the predicted wage distribution based on the DFL method can explain the actual wage distribution of the NLSY97 cohort at around age 35. The residual then includes the impact of changes in skill prices associated with shifts in demand side factors.

The counterfactual wage distribution $f(w^{79}|z_{97},97)$ from the DFL method depends on which variables are used to estimate the propensity matching score. Table 2 documents lists of variables to be used to estimate propensity matching scores for Models 1-6. The benchmark model is Model 6, which includes i) demographic characteristics: race, gender, parents' education (highest grade completed by the mother and father), family structure (whether the individual lived with mother/father/both at age 14), and standardized AFQT score; (ii) education: degree attainments as of age 35, field of study for college education, and age when completed schooling;[32] and (iii) early work experience: average weekly working hours between ages 18 and 22 and occupation held at ages 22 and 25, when estimating the propensity matching score.[33] Models 1-2 include only demographic characteristics, Models 3-4 also include education, and Models 5-6 also include early work experience.

5 Results

5.1 Counterfactual Wage Gains Based on the DFL Method

Benchmark Counterfactual Wage Gain associated with Changes in Premarket Skill

The solid line in Panel (a) of Figure 4 is the wage gain explained by changes in premarket skills across two cohorts in the benchmark model (Model 6). The statistical significance of the wage gain is tabulated in Table 3.

[32] The field of study is for the highest grade completed by the individual.

[33] The selection of variables used in the matching is similar to that of Altonji et al. (2012), who include the same set of demographic characteristics (race, gender, highest grade completed by the mother and father, whether the individual lived with mother/father/both at age 14, and standardized AFQT score), and variables on schooling (highest grade completed as of age 22 and enrollment status at age 22) and school-to-work transition (whether the individual graduates on time/earlier/later for the degree obtained as of age 22). Because I find that the schooling period covers far beyond age 22 and the gap between the two cohorts in educational attainment widens after age 22, I include more variables on educational attainment and the school-to-work transition.

Focusing on population including both men and women, I find that changes in premarket skills explain 5-9% wage gains of the recent cohort relative to the old cohort. The wage gain associated with skill changes across two cohorts is smallest at the bottom end of the wage distribution—4.6% (5.0%) wage increase at the 10th (5th) percentile of the distribution—and largest at the top end of the distribution—6.5% (8.7%) wage increase at the 90th (95th) percentile of the wage distribution. However, the differences in the wage gain associated with changes in premarket skills are not large across different percentiles of the wage distribution. By using the same data sets and econometric methods, Altonji et al. (2012) find similar patterns: changes in the skill explains about 5% increase in the wage between the 25th and 85th percentiles of the wage distribution is about 5%, while the gains for the top decile are in the 7-12%. The wage gain at the top is estimated higher in Altonji et al. (2012) than the finding in this paper, which can be explained by i) different choices for the years of experience when standardizing the wage rate (23 years in Altonji et al. (2012) and 13 years in this paper) and ii) the fact that wage inequality widens along the life-cycle.

Gender Differences in Wage Gain Associated with Changes in Premarket Skills

The solid lines in Panels (b) and (c) of Figure 4 plot the benchmark wage gain explained by changes in premarket skills (Model 6) for men and women, respectively. The statistical significance of the benchmark wage gain for men and women is tabulated in columns (7)-(8) in Table 4. There exist substantial gender differences. First, accounting for race, gender, family background, AFQT, education, and early work experience in Model 6, the change in premarket skills substantially increases women's wage at all levels of the wage distribution, but it does not significantly increase men's wage across two cohorts (Column (7) in Table 4). Second, the wage gain associated with changes in premarket skills monotonically increases for men (although it is not significant), but it is hump-shaped for women along the wage distribution. For women, the wage gain associated with skill change explains the 0.053 log point wage increase at the 10th percentile of the wage distribution, the 0.134 log point wage increase at the 50th percentile, and the 0.116 log point wage increase at the 90th percentile of the wage distribution.[34]

Note that the estimated wage gains for men are significantly positive at the top end of the wage distribution if I use different sets of premarket skills when estimating the propensity weights. In Table 5, the estimates are significantly positive for men at the top and of the wage distribution across Model 1-5. In particular, the estimated gains are 0.048 log point at 75th percentile, 0.092 log point at 90th percentile, and 0.093 log point at 95th percentile. This finding is consistent with estimates in Altonji et al. (2012) who use skill measures as of age 22 in the estimation of the propensity weight. However, once I add working hours and occupation during age 18-22 in Model 6, the estimated wage gain becomes insignificant. Thus, increasing labor supply at low paying jobs during college periods is an important factor that explains why the recent male cohort do not gain significant wage gain over time.

To sum, there exists a substantial gender divergence in skill acquisition of men and women across two cohorts at all levels of the wage distribution. Women at all levels of wage distribution significantly gain premarket skills, whereas men do not accumulate more premarket skills in recent years except at the top end of the wage distribution. By estimating the wage gain at each percentile of the wage distribution, the above finding shows that it is the middle range of the wage distribution where the gender divergence in premarket skill changed mostly across the two cohorts.

Marginal Effect of Different Premarket Skills

Altonji et al. (2012) find that most of skill gains across the NLSY79 and NLSY97 cohorts can be explained by the change in characteristics determined before the college period (precollege skills) across the two cohorts, arguing that American youth do not acquire significantly greater skills despite the dramatic increases in earnings premium for skilled-labor recent decades (Heckman et al. (2008)). In this section, I re-evaluate the marginal contribution of skill components (i) determined before the college period, (ii) acquired through education, and (iii) related to the early work expe-

[34] Altonji et al. (2012) find that the wage gain at the prime age (23 years of experience) explained by skill changes monotonically increases for White men and women, but has a hump-shaped pattern for Black and Hispanic women.

rience in explaining the wage gain across the two cohorts.

The dashed line in Panel (a) of Figure 4 plots the counterfactual wage gain explained by the change in the precollege skills of the whole population when I use skill measures specified in Model 2 —race, gender, family background, and the AFQT score. The statistical significance of the estimates for the wage gain is tabulated in Table 3. Consistent with Altonji et al. (2012), the wage gain associated with changes in those precollege skills is significantly positive and large, accounting for 0.03-0.65 log points increases. The solid line with plus marks in Panel (a) of Figure 4 plots the estimated wage gain in Model 3 which additionally includes skills acquired through college education as measured by the final degree attainment and the field of study when estimating the propensity weights. The sequential marginal effect (SME) of skills acquired through college education is estimated by subtracting the estimated wage gain in Model 2 (dashed line) from that in Model 3 (solid line with plus marks). Adding educational attainment to precollege premarket skills significantly raises the wage gain of the NLSY97 cohort by 0.02-0.05 log points. The SEM of college education is greatest at the top end of the wage distribution; 0.024 log wage points at the 10th percentile of the wage distribution, 0.035 log wage points at the 50th percentile, and 0.043 log wage points at the 90th percentile. Comparing the estimates in Model 2 and the SME of college education, the wage gain explained by the change in college education is about 70% of the wage gain explained by changes in precollege characteristics. As a comparison, Altonji et al. (2012) finds changes in college education explains a much smaller marginal gain in the wage across two NLSY cohorts once changes in precollege skills are accounted. Different findings in two papers can be explained by different measures for college education; the educational attainment measured at age 35 in this paper is significantly greater than that measured at age 22 in Altonji et al. (2012), and the delayed school-to-work transition are more pronounced among the recent cohorts (Section 2.1).

The SME of skills acquired through the early work experience is estimated by subtracting the wage gain in Model 3 from the wage gain the benchmark counterfactual in Model 6 (solid line in Panel (a) of Figure 4) which further includes the age completing schooling, labor supply during ages 18-22, and occupation at ages 22 and 25. The marginal contributions of the school-to-work transition variables have significantly negative impacts on the wage gain of the NLSY97 cohort, reducing the wage by about 0.02 log point at all levels of the wage distribution. The negative impact of these school-to-work transition variables is largest around the 70th-80th percentiles of the wage distribution.

Why do American youth work more during their early 20s at low-paying jobs, even though it would reduce their skill components in the labor market? Increasing demand for low-skilled jobs could provide more opportunities for college students to work in part-time jobs. Another potential explanation is the increased cost of a college education. As discussed in Belley and Lochner (2007) and Lochner and Monge-Naranjo (2011), the NLSY 97 cohort may face more credit constraints in financing their college education than the NLSY79 cohort. As discussed in Bound et al. (2012), the increased financial burden for college education could have raised the need for self-financing during college by working at low-paying jobs.[35]

The marginal effect of different premarket skills also exhibits substantial gender differences (Panels (b) and (c) of Figure 4). First, the wage gain associated with the change in precollege characteristics in Model 2 monotonically increases for men along the wage distribution (dashed line in panel (b)), but it is hump-shaped for women (dashed line in panel (c)). Second, the marginal contribution of acquired skills through education and early work experience is much greater for women than men. For men, the change in degree attainments and field of study explains the 0.01 to 0.03 log point wage gain of the NLSY97 cohort compared with the NLSY79 cohort, but accounting for additional school-to-work transition variables cancels out the gain associated with higher educational attainment. As a result, for men, the benchmark counterfactual wage gain is almost the same as the case in which I only include precollege characteristics. For women, because the marginal effect of college education is much greater than men, thus, the wage gain in the benchmark model is still positive and statistically significant after accounting for negative

[35]Comparing slightly older cohorts, between high school graduates in 1972 and 1992, Bound et al. (2012) find that increasing time to get a BA degree is more pronounced among low-income students at selective public universities where the resources for each student decreased substantially over time.

impacts of early work experience on the wage gain.

5.2 Decomposition of the Wage Gain across the Two Cohorts

Figure 5 shows the decomposition results of the actual wage gain across the NLSY79 and NLSY97 cohorts. The decomposition shows to what extent changes in premarket skills can explain the observed evolution of the wage distribution across the two cohorts. The solid lines in Figure 5 plot the wage gain of the NLSY97 cohort compared with the NLSY79 cohort explained by changes in the skill correlates in Model 6 (part (ii) in equation (1)). The dashed lines in Figure 5 plot the wage gain across two cohorts that is not explained by changes in premarket skills ((i) in equation (1)). Thus, the dash lines reflect wage gains explained by changes in demand side factors, such as skill prices, over time.

Panel (a) in Figure 5 is the decomposition result for the entire population consisting of both men and women. As discussed in the previous section, changes in premarket skills (solid line) cannot explain the U-shaped wage gain across the two cohorts. The U-shaped wage gain is largely explained by the residual (dashed line). This finding is consistent with Autor et al. (2008), who demonstrate that changing skill prices is the main driving factor that explain increasing (decreasing) upper-tail (lower-tail) wage inequality.[36] However, premarket skills still play an important role in explaining the overall increase in the wage across the two cohorts. In particular, changes in premarket skills (in particular, a large increase in educational attainment) cancel out the substantial wage loss (7%-8%) in the middle range of the wage distribution driven by changes in the skill

[36] Altonji et al. (2012) find a greater contribution of skill changes to the increasing upper-tail wage inequality; they find that the change in premarket skills between the NLSY79 and NLSY97 cohorts account for the 5% increase in the lower-tail (50th-10th) wage inequality and the 10% increase in the upper-tail (90th-50th) wage inequality. Although I closely follow the sample construction and estimation procedures of Altonji et al. (2012), two important differences exist. First, different from Altonji et al. (2012), who use wage data at prime ages (with 23 years of potential experience), I use wage data at around age 35 (13 years of potential experience). Thus, the different finding on upper-tail wage inequality between this paper and Altonji et al. (2012) can be related to the fact that the wage inequality fans out over the life-cycle Lemieux (2006). Second, I use a different set of premarket skills to estimate the benchmark propensity score from Altonji et al. (2012). While they include the same set of pre-college premarket skills, I include different variables for education and early work experience. As discussed in Section 2.1, there are substantial differences in the educational attainment measure between age 22 and age 35.

prices. Also, changes in premarket skills explain more than half of the wage gain experienced by the top 10th percentile of the wage distribution.

Panel (b) and (c) in Figure 5 plot decomposition results for men and women, respectively. For both men and women, the residual (the estimate for part (i)) presents a U-shaped pattern, having negative values in the middle ranges of the wage distribution. The wage loss associated with changes in the skill prices is greater for men. For men, the change in skill prices has a negative impact on the wage gain for most of the individuals below the 85th percentile of the wage distribution. For women, the skill prices increased at the bottom and top quintiles of the wage distribution, and decreased in the middle. At the 50th percentile, changes in premarket skills explain 0.134 log point increase in wage across two cohorts, whereas changes in skill prices explain 0.065 log point decrease in wage. At the 90th percentile, changes in premarket skills increase the wage by 0.116 log points, whereas changes in skill prices increases the wage by 0.066 log points. For women, changes in premarket skills play quantitatively more important roles in determining the wage gain across two cohorts than changes in skill prices.

Focusing on the wage inequality for men and women, the change in premarket skills increases the upper-tail wage inequality for men (0.041 log point increase in 90th-50th inequality and 0.033 log point increase in 95th-50th wage inequality), and it decreases the upper-tail wage inequality for women (0.018 log point decrease in 90th-50th inequality and 0.033 log point decrease in 95th-50th wage inequality).[37] Therefore, recent changes in premarket premarket skills have different implications on the recent trends in wage inequality for men and women. For men, changing the premarket skills contributes positively to the increasing upper-tail wage inequality over the two cohorts, explaining 2/3 of the trend in the data. For women, changes in premarket skills played an important role in reducing the upper-tail wage inequality, thus cannot explain the increasing upper-tail wage inequality observed in the data.

[37] Appendix A presents to what extent the change in premarket skills explains the difference between the wage at each percentile of the wage distribution from the median.

5.3 Changes in Premarket Skills and the Gender Wage Gap

In this section, I decompose the change in the gender wage gap between the NLSY79 and NLSY97 into two parts: (i) changes in premarket skills and (ii) all other factors (residual) including changes in skill price. The dashed line in Figure 6 plots the change in the gender wage gap observed in the data. The solid line with plus marks in Figure 6 plots the change in the gender wage gap explained by changes in premarket skills. To calculate the counterfactual trend in the gender wage gap explained by skill changes, I first calculate the gender wage gap of the NLSY79 cohort if I reweight the characteristics to represent that of the NLSY97 cohort in Model 6 (the benchmark counterfactual gender wage gap). Then I subtract the benchmark counterfactual gender wage gap from the gender wage gap in the data for the NLSY79 cohort.

The counterfactual decrease in the gender wage gap associated with changes in premarket skills is hump-shaped and smallest at the top end of the wage distribution.[38] This implies that the reduction in the gender wage gap associated with the gender specific change in premarket skills is greater in the middle of the wage distribution and smallest at the top end. Changing the premarket skills explains almost all of the reduction in the gender wage gap across the two cohorts in the middle. However, changes in the premarket skills alone cannot explain the decrease in the gender wage gap at the top and bottom ends of the wage distribution. As discussed in Cortes et al. (2020), changes in skill price, therefore, seem to also contribute significantly to explaining the narrowing gender wage gap at the top and bottom ends of the wage distribution.

[38] In Appendix C, I document the gender-specific SME of different premarket skills in explaining the relative wage gain of women to men across two cohorts.

6 Demand Side Factors of Gender-Specific Changes in Skill Acquisition

Given the pronounced divergence between men and women in their skill acquisition and wage distribution across two cohorts, I further investigate the source of gender specific change in parmarket skill accumulation. It can complement previous literature that takes education as exogenously given factors when evaluating the role of demand side shifts in closing the gender wage gap (Galor et al., 1996; Welch, 2000; Weinberg, 2000; Black and Spitz-Oener, 2010; Bacolod and Blum, 2010; Cortes et al., 2020).[39]

Can Trends in Gender-Specific Returns to Education Explain the Overall Gender Gap in Educational Attainment?

Table 6 documents the estimates from OLS regressions of the log hourly wage of men and women on various premarket skills. To facilitate comparison with previous findings based on the NLSY79 and NLSY97 cohorts, I closely follow Castex and Kogan Dechter (2014) regarding sample construction and estimation equation for the wage regression. In order to use recent wage data, I include wage data between age 18 and 35 instead of 18 and 28, as in Castex and Kogan Dechter (2014). First, as Castex and Kogan Dechter (2014) show, the returns to education increased, whereas the returns to cognitive skill, as measured by the standardized AFQT score, decreased substantially for both men and women. For instance, the returns to a BA degree increased from 0.375 to 0.562 for men and from 0.472 to 0.659 for women across the two cohorts. On the other hand, the returns to AFQT score decreased from 0.109 to 0.054 for men and from 0.123 to 0.075 for women. The changing skill price for education and AFQT score can explain the general increase

[39] Galor et al. (1996) and Welch (2000) claim that gender differences in brains and brawn are important in explaining the narrowing gender gap in the labor market. Weinberg (2000) provides empirical evidence that a decline in emphasis on physical skills following computerization has increased the demand for female workers. Black and Spitz-Oener (2010) and Bacolod and Blum (2010) also find evidence that the shift in demand from manual tasks to analytical tasks plays an important role in changes in the gender wage gap. Cortes et al. (2020) show that since the 1980s, the probability of working in a cognitive/high-wage occupation decreases for men but increases for women, and the main factor driving this divergence is the increase in social skills in cognitive/high-wage occupations.

in educational attainment for both men and women. Second, the return to education is higher for women than for men at all levels of academic degrees in both the NLSY79 and NLSY97 cohorts. However, the gender gap in the return to education is greater in the NLSY97. The gender gap in the return to a BA degree doubled across the two cohorts, increasing from 0.097 to 0.187, and the gender gap in the return to a MA or above increased from 0.223 to 0.275. Therefore, the skill price for men and women changes differently across the two cohorts, which can explain the intensified gender difference in educational attainment for the NLSY97 cohort.

Why Men in the Middle Range of the Wage Distribution Did Not Accumulate as much Skills as Comparable Women?

In Section 5.3, I show that the gender-specific change in premarket skills can explain most of the decrease in the gender wage gap in the middle of the wage distribution, but its contribution is smaller at the top and bottom ends of the wage distribution. In following, I argue that (i) substantial shifts in gender-specific occupational sorting in the middle range of occupation and (ii) heterogeneous trends in skill requirement across occupations could explain why the educational attainment of women increased much faster than that of men in the middle range of the wage distribution.

Figure 2 plots the change in the share of men and women in each occupation at ages 18, 22, and 30 between the NLSY79 and NLSY97 cohorts.[40] On the x-axis, I plot occupations according to the mean wage for the NLSY97 cohort, which is documented in Table A4. Consistent with the polarization literature, employment share increased at the top end (manager) and bottom end of the wage distribution (manual) (Autor et al. (2003), Autor and Dorn (2013)). However, the increase in the share of employment at both ends of the job distribution does not exhibit significant gender differences. Focusing on managerial occupations, the employment share at age 30 increased by 2.3 and 2.4 percentage points for men and women, respectively.[41]

[40] The share is calculated for all individuals with valid information on current/most recent job, including those who do not have a job.

[41] STEM occupations present an opposite trend in the employment share by gender, but the magnitude is small. Changes in STEM jobs are small, such that the share increases by 0.5 percentage point for men and decreases by 0.6 percentage point. As documented by Deming and Kahn (2018), growth in the employment share is slower for STEM

On the other hand, the share of workers in sales/clerical occupations (lower-middle) increased by 2.2 percentage points for men, but it decreased by 8.5 percentage points for women at age 30. The share of workers in health care, education, and social occupations (upper-middle) increased by 2.3 percentage points for men and 7.1 percentage points for women at age 30. Thus, changes in the employment share across the two cohorts in the middle range of the occupation distribution are greater, and more importantly, present substantial gender differences. The rapid growth in female-dominant occupations such as heath care, education, social occupations plays an important role in creating gender-specific shifts in the middle range of the occupation.

To examine to what extent educational attainment is associated with shifts in employment share across occupations, I use job-posting data from Atalay et al. (2020) and examine changes in the college degree requirement across different occupations.[42]

Figure 7 plots changes in the share of job postings that require a four-year college degree between 1975 and 1998.[43] First, consistent with the literature, more jobs require a college degree over time. In 1975, only 2% of job ads in managerial, STEM, social occupations explicitly mention their requirement for a four-year college degree; in 1998, it is about 8%. Low-paying jobs (sales, operative, manual) also increasingly require a four-year college degree; this increases from 1% to 5% between 1975 and 1998. Importantly, the gap in college degree requirement between high-paying and low-paying jobs also increased from 1% to 3%. Second, the share of jobs that require a college degree increased rapidly in occupations with increasing share of females. Comparing sales/clerical and health care/education/social occupations, the latter increases faster in the college degree requirement. Therefore, changes in occupational sorting and in educational requirement jointly explain the gender differences in educational attainment, especially in the middle range of the wage distribution.[44]

---

major as technological progress slows down.

[42] Atalay et al. (2020) construct a dataset from text content of 7.8 million job ads from the *Boston Globe*, *New York Times*, and *Wall Street Journal* and report the skill requirements/task contents of each job across 1945-2000. I use their public data, which aggregate skill requirements and other information at 3-digit SOC level.

[43] Due to lack of data, I could not document the trend after 2000. I use the 3-year moving average to calculate the share of jobs with a four-year college degree requirement.

[44] As well documented in the literature (Acemoglu and Autor (2011), Autor et al. (2006), Deming (2017), Deming

7 Conclusion

Based on the NLSY79 and NLSY97, I find substantial changes in educational attainment and the school-to-work transition across 1957-1964 and 1980-1984 birth cohorts in the U.S. I claim that changes in premarket skills have different implications on the trend of wage inequality for men and women. A large increase in the educational attainment of women plays an important role in reducing the growing wage inequality driven by changes in skill prices. For men, acquired skills positively contributes to the widening upper-tail wage inequality. I argue that gender-specific responses to skill-biased technological changes in their skill acquisition process is important in explaining the trend.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier.

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.

Ashworth, J., Hotz, V. J., Maurel, A., and Ransom, T. (2017). Changes across cohorts in wage returns to schooling and early work experiences. Technical report, National Bureau of Economic Research.

Atalay, E., Phongthiengtham, P., Sotelo, S., and Tannenbaum, D. (2020). The evolution of work in the united states. *American Economic Journal: Applied Economics*, 12(2):1–34.

and Kahn (2018), Atalay et al. (2020)), there has been substantial change in the employment share of each occupation over the last few decades. While the literature on job polarization highlights the increasing demand for cognitive skills and decreasing demand for routine/manual skills between the 1980s and 2000s with computerization and automation, recent studies document that the growth of the demand for high-skilled jobs has slowed down since 2000 (Acemoglu and Autor, 2011; Beaudry et al., 2016,?; Deming, 2017). Deming (2017) finds that high-paying jobs increasingly require social skills, whereas the share of STEM jobs decreased substantially between 2000 and 2012.

Autor, D., Katz, L. F., and Kearney, M. S. (2005). Rising wage inequality: the role of composition and prices. Technical report, National Bureau of Economic Research.

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.

Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in us wage inequality: Revising the revisionists. *The Review of economics and statistics*, 90(2):300–323.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333.

Autor, H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–97.

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.

Beaudry, P., Green, D. A., and Sand, B. M. (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, 34(S1):S199–S247.

Belley, P. and Lochner, L. (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human capital*, 1(1):37–89.

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. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Bound, J., Lovenheim, M. F., and Turner, S. (2012). Increasing time to baccalaureate degree in the united states. *Education Finance and Policy*, 7(4):375–424.

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

Castex, G. and Kogan Dechter, E. (2014). The changing roles of education and ability in wage determination. *Journal of Labor Economics*, 32(4):685–710.

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. and Kahn, L. B. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1):S337–S369.

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.

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

Goldin, C. D. and Katz, L. F. (2009). *The race between education and technology*. harvard university press.

Heckman, J. J., Lochner, L. J., and Todd, P. E. (2008). Earnings functions and rates of return. *Journal of human capital*, 2(1):1–31.

Hendricks, L. and Schoellman, T. (2014). Student abilities during the expansion of us education. *Journal of Monetary Economics*, 63:19–36.

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.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, 96(3):461–498.

Lochner, L. J. and Monge-Naranjo, A. (2011). The nature of credit constraints and human capital. *American economic review*, 101(6):2487–2529.

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

Taber, C. and Vejlin, R. (2020). Estimation of a roy/search/compensating differential model of the labor market. *Econometrica*, 88(3):1031–1069.

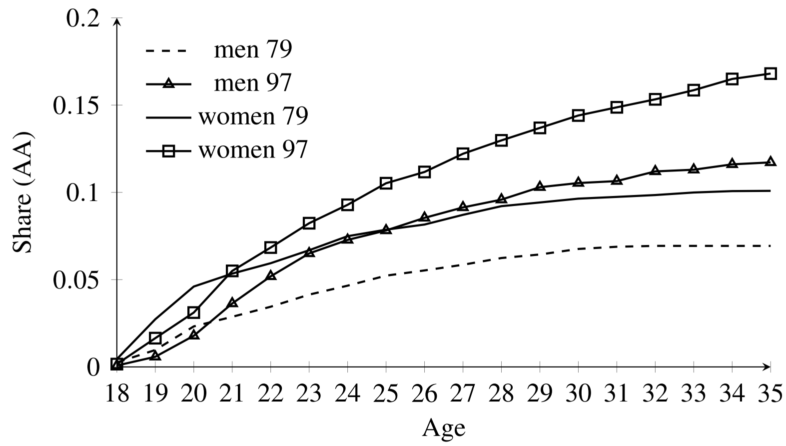
Weinberg, B. A. (2000). Computer use and the demand for female workers. *ILR Review*, 53(2):290–308.

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.

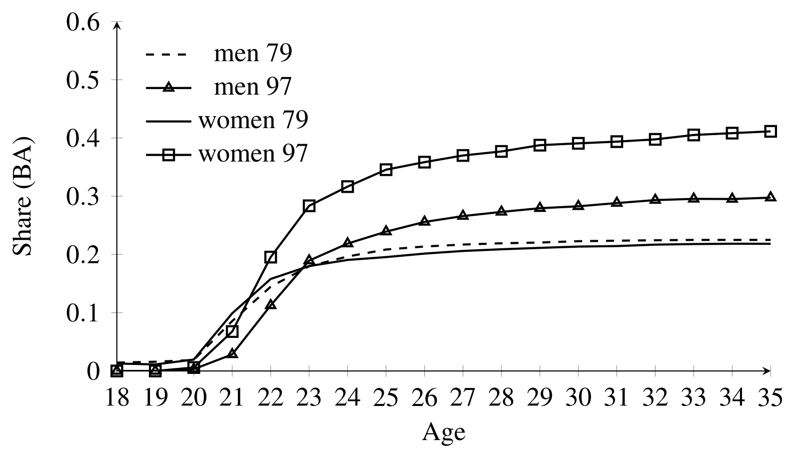
Yamaguchi, S. (2018). Changes in returns to task-specific skills and gender wage gap. *Journal of Human Resources*, 53(1):32–70.

Figure 1: Gender Difference in Changes in Educational Attainment

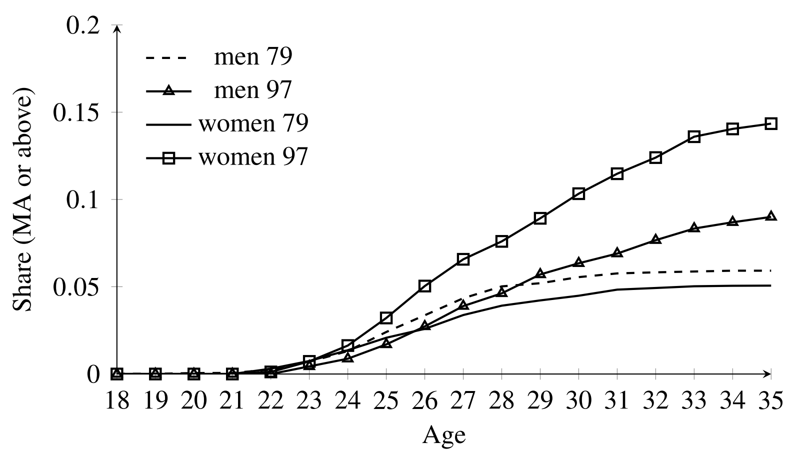
(a) Share of Individuals with AA



(b) Share of Individuals with BA



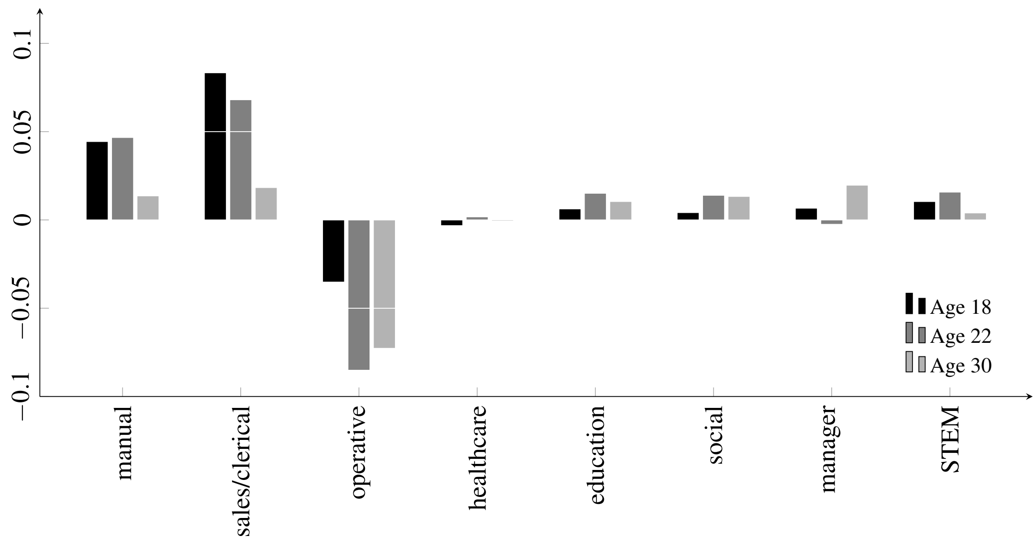
(c) Share of Individuals with MA or above



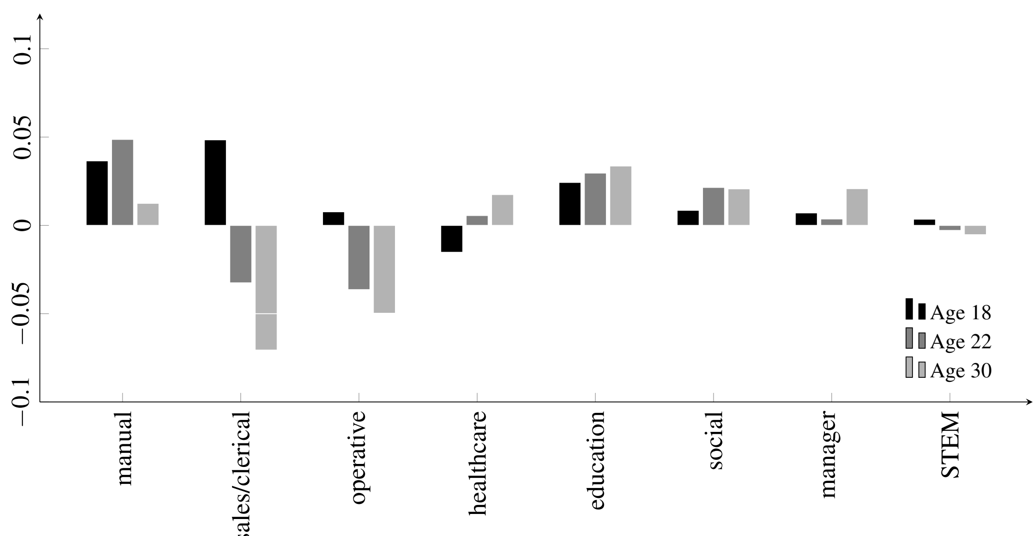
Note. The figure documents the share of the population whose highest grade completed is AA, BA, MA, or above by age. Data sources are from the NLSY79 and NLSY97.

Figure 2: Changes in the Employment Share across the NLSY79 and NLSY97 Cohorts

(a) Changes in Occupational Sorting of Men



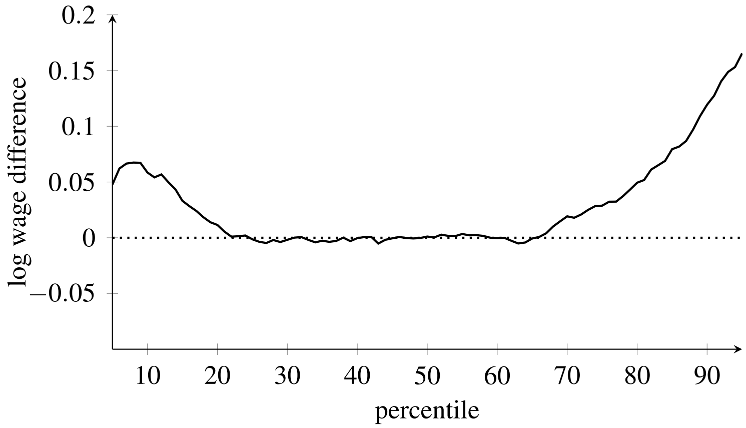
(b) Changes in Occupational Sorting of Women



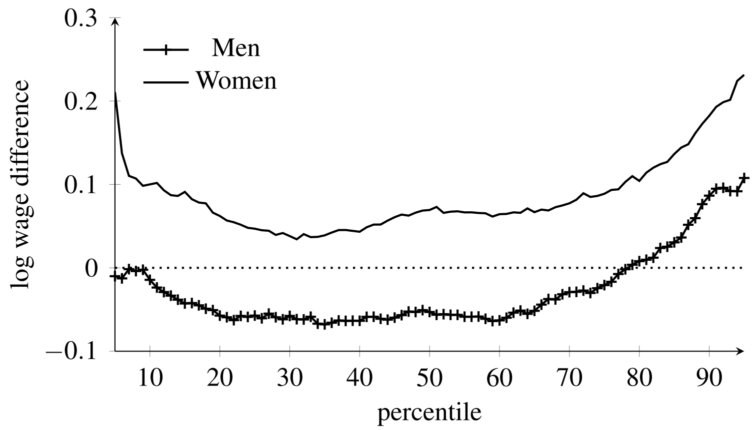
Note. Panels (a) and (b) plot the change in the share of men and women in each occupation at ages 18, 22, and 30 between the NLSY79 and NLSY97 cohorts. The share is calculated for all individuals with valid information on current/most recent job, including those who do not have a job. On the x-axis, I plot occupations according to the mean wage for the NLSY97 cohort documented in Table A4.

Figure 3: Log Wage Difference between NLSY97 and NLSY79 cohorts

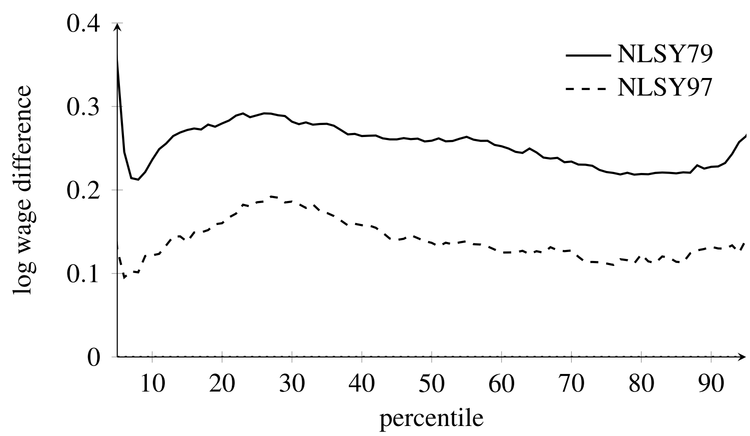
(a) Log Wage Difference between NLSY97 and NLSY79 cohorts



(b) Changes in Wage between NLSY79 and NLSY97 Cohorts



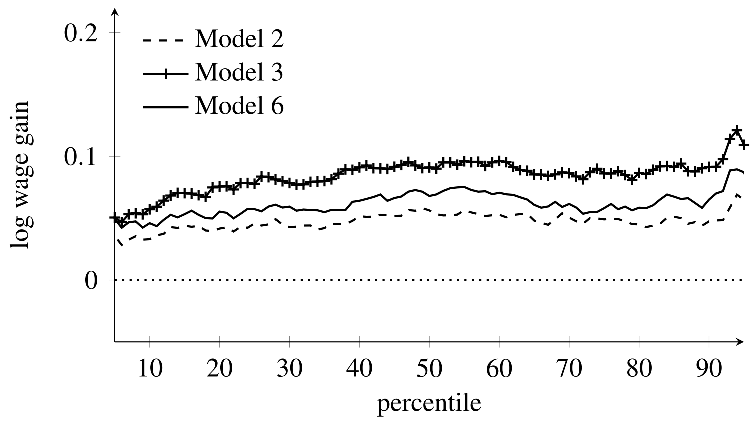
(c) Gender Wage Gap



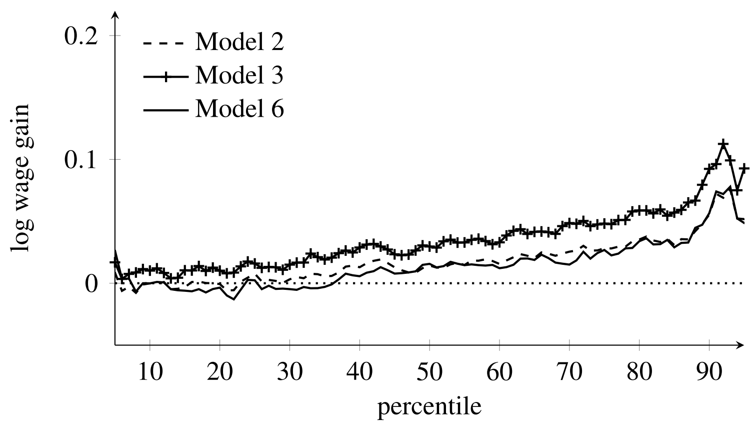
Note. Figure 2 (a) shows the log wage gain of the NLSY97 cohort compared with the NLSY79 cohort at each percentile of the wage distribution. The hourly wage is regression standardized with a potential experience of 13 years. Figure 2(b) shows the difference in the log wage gain of the NLSY97 cohort compared with the NLSY79 cohort by gender. Figure 2(c) shows the log wage difference between men and women at each percentile of the wage distribution for each NLSY cohort.

Figure 4: Counterfactual Wage Gain obtained from the DFL Method

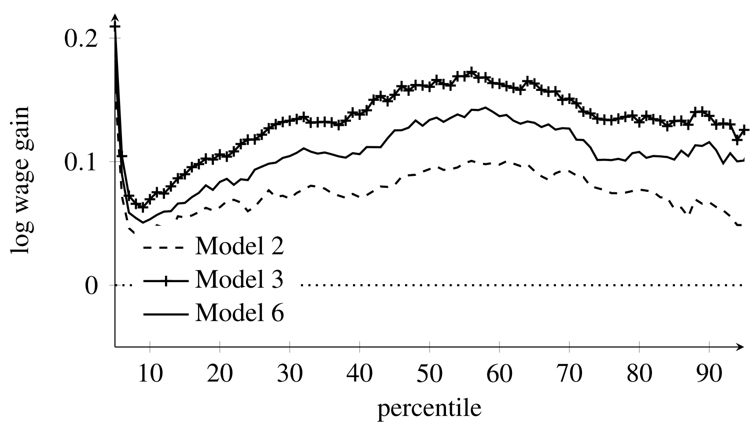
(a) Counterfactual Wage Gain for Whole Population



(b) Counterfactual Wage Gain for Men



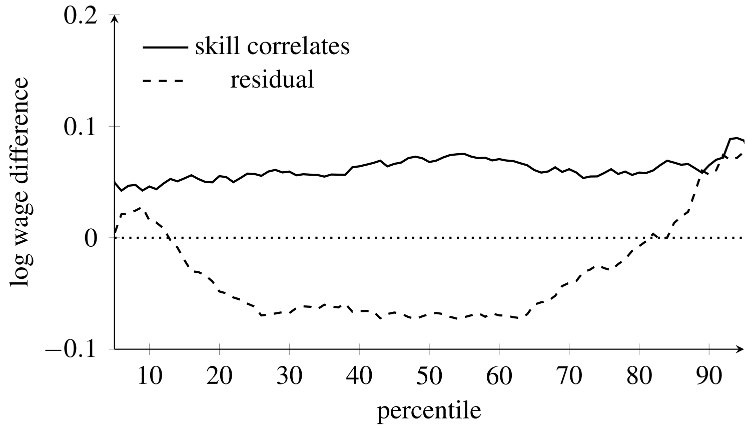
(c) Counterfactual Wage Gain for Women



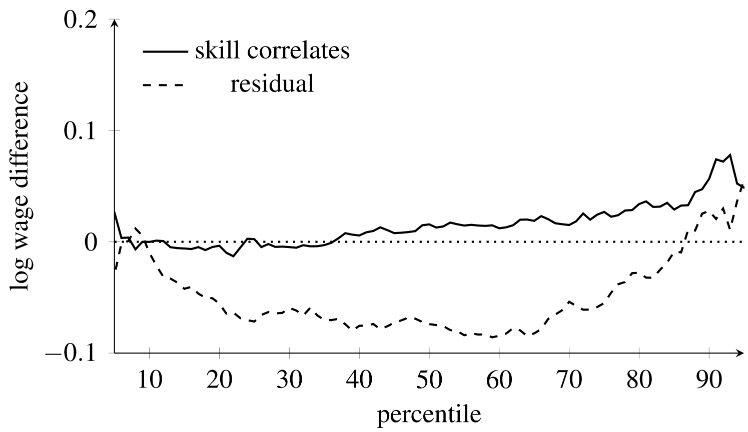
Note. Panels (a), (b), and (c) show the counterfactual wage gain for the whole population, for men, and for women, respectively. The dashed line matches the two cohorts based on race, gender, standardized AFQT score, parents' education, and family structure, which corresponds to Model 2. The solid line with plus marks also includes the field of study for college education and final degree attainment as of age 35, which corresponds to Model 3. The solid line is the benchmark counterfactual that also includes detailed school-to-work transition variables—weekly working hours during ages 18-22, occupation at ages 22 and 25, and age when completing education, which corresponds to Model 6.

Figure 5: Decomposition of the Wage Gain: Skill Correlates vs. Residual

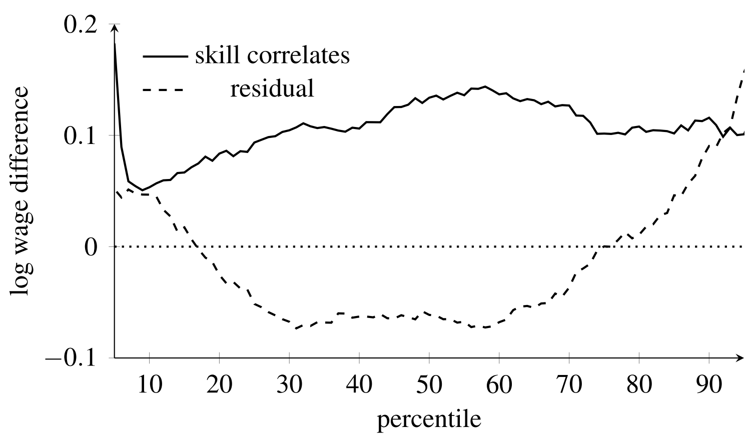
(a) Decomposition for Whole Population



(b) Decomposition for Men

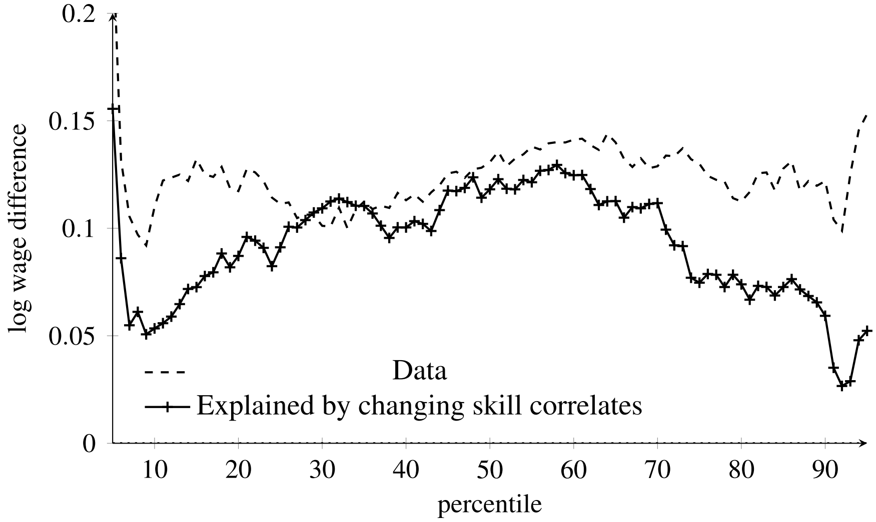


(c) Decomposition for Women



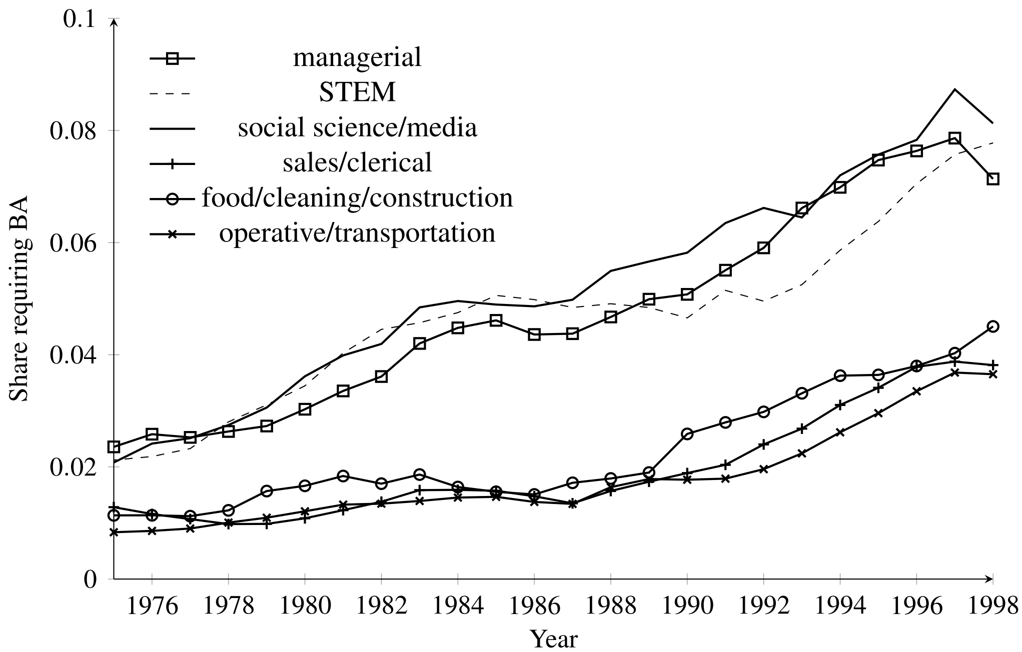
Note. The solid lines in Panels (a), (b), and (c) plot the benchmark (Model 6) wage gains of the NLSY97 cohort compared with the NLSY79 cohort explained by changes in the skill correlate (part (ii) in equation (1)) for the whole population, for men, and for women, respectively. The dashed lines plot the difference between the actual wage gain of the NLSY97 cohort relative to the NLSY79 cohort observed in the data and the counterfactual wage gain obtained from the DFL method (part (i) in equation (1)).

Figure 6: Decrease in the Gender Wage Gap over the Two Cohorts



Note. The dashed line plots the change in the gender wage gap observed in the data. The solid line with plus marks plots the change in the gender wage gap across the two cohorts by subtracting the gender wage gap in the data for the NLSY79 cohort from the benchmark counterfactual gender wage gap—i.e., the gender wage gap of the NLSY79 cohort if I reweight the characteristics to represent that of the NLSY97 cohort in Model 6.

Figure 7: Share of Job Postings Requiring a Four-Year College Degree



Note. The figure plots changes in the share of job postings that require a four-year college degree between 1975 and 1998. I use the 3-year moving average to calculate the share of jobs with a four-year college degree requirement. Data source is job-posting data from Atalay et al. (2020) and examine changes in the college degree requirement across different occupations. Atalay et al. (2020) construct a dataset from text content of 7.8 million job ads from the *Boston Globe*, *New York Times*, and *Wall Street Journal* and report the skill requirements/task contents of each job across 1945-2000. I use their public data, which aggregate skill requirements and other information at 3-digit SOC level.

Table 1: Education and School-to-Work Transition

Sample	NLSY79		NLSY97		Difference	
	mean	s.d.	mean	s.d.	mean	s.e.
A. Degree received as of age 35						
HS	0.66	0.47	0.54	0.50	-0.12***	0.09
AA	0.09	0.28	0.14	0.35	0.06***	0.005
BA	0.22	0.42	0.35	0.48	0.13***	0.008
MA or above	0.05	0.23	0.12	0.32	0.06***	0.005
B. Degrees obtained by age						
AA by age 22	0.05	0.21	0.06	0.24	0.013***	0.004
BA by age 22	0.15	0.36	0.15	0.36	0.002	0.007
AA by age 27	0.07	0.26	0.11	0.31	0.033***	0.005
BA by age 27	0.21	0.41	0.32	0.46	0.11***	0.008
MA by age 27	0.04	0.19	0.05	0.22	0.013***	0.004
Dropout	0.48	0.50	0.37	0.48	-0.11***	0.011
Age received AA	23.33	3.75	25.24	4.54	1.91***	0.22
Age received BA	22.90	2.38	24.21	3.05	1.31***	0.099
Age received MA or above	26.73	2.63	28.49	3.22	1.75***	0.20
C. Highest grade completed						
as of age 22	12.70	2.06	13.09	2.28	0.39***	0.04
as of age 35	13.45	2.47	14.35	2.83	0.89***	0.04
D. College major						
Pure STEM	0.03	0.16	0.03	0.18	0.008***	0.003
Applied STEM	0.09	0.29	0.07	0.26	-0.018***	0.005
Business	0.17	0.37	0.13	0.34	-0.038***	0.006
Social/Humanity	0.05	0.22	0.12	0.33	0.069***	0.005
Education/Health	0.13	0.34	0.14	0.35	0.014**	0.006
E. Working during college						
hrs/week before 22	25.77	14.37	26.14	11.07	0.37	0.22
hrs/week before 22, college	18.42	13.42	21.77	11.09	3.35***	0.30
weeks/year, college, age 18	29.63	19.02	31.48	19.34	1.85***	0.38
weeks/year, college, age 20	33.09	19.67	35.75	19.56	2.66***	0.37
weeks/year, college, age 22	36.27	19.18	37.54	19.13	1.27***	0.35
F. Occupation at age 18						
Business	0.005	0.067	0.011	0.11	0.006***	(0.002)
STEM	0.001	0.031	0.008	0.087	0.007***	(0.002)
Manual	0.23	0.42	0.27	0.44	0.039***	0.011
Sales and Clerical	0.26	0.44	0.33	0.47	0.069***	0.012
Social/Teachers/Health	0.29	0.17	0.41	0.20	0.013**	0.005
Not working	0.20	0.40	0.11	0.31	-0.09***	0.009
G. Occupation at age 22						
Business	0.04	0.20	0.04	0.21	0.0006	0.004
STEM	0.02	0.13	0.02	0.15	0.006**	0.016
Manual	0.18	0.38	0.23	0.41	0.047***	0.008
Sales and clerical	0.27	0.44	0.28	0.45	0.018**	0.009
Social/Teachers/Health	0.03	0.17	0.04	0.20	0.043***	0.005
Not working	0.14	0.35	0.08	0.27	-0.06***	0.006
Number of obs.	8,848		6,028			

Note. The table tabulates educational attainment and school-to-work transition variables of the NLSY79 and NLSY97 cohorts. I use cross-sectional weights adjusted to account for attrition and nonresponse for the AFQT score to obtain summary statistics.

Table 2: Weights Used to Produced Counterfactual Wage Distribution based on the DFL Method

Model	Skill Correlates Included in z
ψ_0	$\psi_{NLSY79} \times \psi_{ATTR-AFQT79}$
Model 1	ψ(race, gender; parental education, intact family)
Model 2	ψ(race, gender; parental education, intact family, AFQT score)
Model 3	ψ(race, gender; parental education, intact family, AFQT score, degree attainments as of age 35, field of study)
Model 4	ψ(race, gender; parental education, intact family, AFQT score, degree attainments as of age 35, field of study, age completing schooling)
Model 5	ψ(race, gender; parental education, intact family, AFQT score, degree attainments as of age 35, field of study, age completing schooling, occupation at age 22 and 25)
Model 6	ψ(race, gender; parental education, intact family, AFQT score, degree attainments as of age 35, field of study, age completing schooling, labor supply during 18-22, occupation at age 22 and 25)

Note: The table presents weights used in estimating the counterfactual wage distribution based on the DFL method. ψ_0 is the weight I obtained by multiplying the cross-sectional smaple weight for the NLSY79 cohort (ψ_{NLSY79}) in the data and the weight to adjust attrition at age 22 and AFQT nonresponse ($\psi_{ATTR-AFQT}$). The weight used to calculate the counterfactual wage distribution for each specification (Model1-Model6) is calculated by multiplying ψ_0 with the weight provided in each cell depending on the model used.

Table 3: Comparison of Actual Wages of 1979 Cohort with Counterfactual Wage Distribution Based on Characteristics of 1997 Cohort

per-centile	Observed Wage Distribution in NLSY79 (1)	NLSY97 (2)	Observed Wage Gain (2) - (1) (3)	Counterfactual Minus Actual Wage Model 2 (4)	Model 3 (5)	Model 4 (6)	Model 5 (7)	Model 6 (8)
5%	6.253	6.306	0.053***	0.036**	0.051***	0.044***	0.046**	0.050***
	(0.011)	(0.007)	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)
10%	6.412	6.474	0.061***	0.033**	0.057***	0.051***	0.051***	0.046***
	(0.008)	(0.002)	(0.009)	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)
25%	6.748	6.744	-0.004	0.046***	0.078***	0.066***	0.069***	0.057***
	(0.009)	(0.003)	(0.009)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
50%	7.141	7.140	-0.001	0.056***	0.091***	0.084***	0.083***	0.068***
	(0.008)	(0.004)	(0.009)	(0.013)	(0.015)	(0.015)	(0.015)	(0.016)
75%	7.527	7.558	0.031***	0.049***	0.086***	0.079***	0.074***	0.058***
	(0.009)	(0.004)	(0.009)	(0.015)	(0.015)	(0.016)	(0.017)	(0.018)
90%	7.850	7.972	0.122***	0.048***	0.091***	0.081***	0.078***	0.065***
	(0.011)	(0.006)	(0.013)	(0.018)	(0.020)	(0.020)	(0.022)	(0.022)
95%	8.071	8.235	0.164***	0.065***	0.109***	0.101***	0.095***	0.087***
	(0.014)	(0.005)	(0.015)	(0.025)	(0.026)	(0.026)	(0.025)	(0.026)
Mean	7.129	7.159	0.030***	0.051***	0.088***	0.077***	0.076***	0.067***
	(0.008)	(0.004)	(0.009)	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)

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 experience = 13. Monetary value is adjusted to 2002 year USD by using CPI-U. All statistics are weighted by the cross-sectional weight, accounting for attrition by age 22 and AFQT nonresponses. 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 4: Comparison of Actual Wages of NLSY79 Cohort with Counterfactual Wage Distribution Based on Characteristics of NLSY97 Cohort by Gender

per-centile	Observed Wage Distribution in NLSY79		Observed Wage Distribution in NLSY97		Observed Wage Gain NLSY97-NLSY79		Counterfactual Minus Actual Wage Benchmark (Model 6)	
	Men (1)	Women (2)	Men (3)	Women (4)	Men (5)	Women (6)	Men (7)	Women (8)
5%	6.367	6.017	6.356	6.227	-0.010	0.211***	0.026	0.182**
	(0.014)	(0.057)	(0.005)	(0.010)	(0.015)	(0.058)	(0.022)	(0.071)
10%	6.560	6.324	6.546	6.424	-0.014	0.100***	-0.000	0.053***
	(0.013)	(0.009)	(0.004)	(0.004)	(0.013)	(0.010)	(0.026)	(0.020)
25%	6.901	6.611	6.843	6.658	-0.057***	0.047***	0.002	0.094***
	(0.012)	(0.011)	(0.004)	(0.003)	(0.012)	(0.011)	(0.023)	(0.025)
50%	7.261	7.002	7.208	7.072	-0.053***	0.069***	0.016	0.134***
	(0.011)	(0.011)	(0.004)	(0.003)	(0.012)	(0.012)	(0.023)	(0.024)
75%	7.627	7.406	7.607	7.495	-0.021	0.089***	0.027	0.102***
	(0.012)	(0.013)	(0.005)	(0.003)	(0.013)	(0.013)	(0.021)	(0.022)
90%	7.944	7.716	8.030	7.898	0.087***	0.182***	0.057	0.116***
	(0.015)	(0.015)	(0.005)	(0.005)	(0.015)	(0.016)	(0.035)	(0.027)
95%	8.180	7.915	8.287	8.146	0.108***	0.231***	0.049	0.102**
	(0.019)	(0.019)	(0.010)	(0.010)	(0.022)	(0.022)	(0.040)	(0.040)
Mean	7.261	6.992	7.226	7.086	-0.035***	0.095***	0.014	0.107***
	(0.010)	(0.011)	(0.006)	(0.005)	(0.012)	(0.012)	(0.019)	(0.032)

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 experience = 13. Monetary value is adjusted to 2002 USD by using CPI-U. All statistics are weighted by the cross-sectional weight, accounting for attrition by age 22 and AFQT nonresponses. 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. Benchmark counterfactual matches race, gender; parental education, intact family, AFQT score, final degree, field of study, age completing schooling, labor supply during 18-22, occupation at age 22 and 25.

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

Men	Observed Wage Gain	Counterfactual Wage Gain Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
5%	-0.004	0.027	0.009	0.017	0.015	0.019	0.026
	(0.015)	(0.019)	(0.023)	(0.023)	(0.022)	(0.023)	(0.022)
10%	-0.010	0.025	0.000	0.010	0.006	0.007	-0.0001
	(0.013)	(0.022)	(0.023)	(0.023)	(0.024)	(0.025)	(0.026)
25%	-0.069***	0.026	0.008	0.016	0.010	0.010	0.002
	(0.012)	(0.018)	(0.020)	(0.019)	(0.019)	(0.022)	(0.024)
50%	-0.058***	0.039**	0.014	0.030	0.025	0.023	0.015
	(0.012)	(0.020)	(0.021)	(0.021)	(0.021)	(0.022)	(0.023)
75%	-0.028	0.051***	0.028	0.048**	0.035*	0.031	0.027
	(0.013)	(0.019)	(0.019)	(0.020)	(0.020)	(0.021)	(0.021)
90%	0.084***	0.074**	0.055*	0.092***	0.072**	0.066**	0.057
	(0.015)	(0.029)	(0.029)	(0.032)	(0.032)	(0.033)	(0.035)
95%	0.104***	0.073*	0.052	0.093**	0.076*	0.069*	0.049
	(0.022)	(0.040)	(0.036)	(0.044)	(0.044)	(0.039)	(0.040)
Mean	-0.035***	0.013	0.014	0.014	0.015	0.016	0.016
	(0.012)	(0.017)	(0.017)	(0.018)	(0.018)	(0.019)	(0.019)

Women	Observed Wage Gain	Counterfactual Wage Gain Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
5%	0.211***	0.166**	0.161**	0.209***	0.186***	0.189**	0.182**
	(0.058)	(0.073)	(0.072)	(0.072)	(0.071)	(0.071)	(0.071)
10%	0.100***	0.048***	0.042***	0.070***	0.058***	0.058***	0.053***
	(0.010)	(0.016)	(0.016)	(0.017)	(0.018)	(0.019)	(0.020)
25%	0.047***	0.079***	0.063***	0.118***	0.105***	0.108***	0.094***
	(0.011)	(0.023)	(0.022)	(0.023)	(0.023)	(0.024)	(0.025)
50%	0.069***	0.102***	0.094***	0.161***	0.152***	0.152***	0.134***
	(0.012)	(0.020)	(0.021)	(0.024)	(0.023)	(0.023)	(0.024)
75%	0.089	0.077***	0.077***	0.134***	0.122***	0.121***	0.102***
	(0.013)***	(0.020)	(0.021)	(0.023)	(0.023)	(0.023)	(0.022)
90%	0.182***	0.071***	0.067***	0.137***	0.124***	0.128***	0.116***
	(0.015)	(0.026)	(0.026)	(0.026)	(0.026)	(0.027)	(0.027)
95%	0.231***	0.052*	0.049*	0.118***	0.104***	0.105***	0.100**
	(0.022)	(0.028)	(0.029)	(0.038)	(0.038)	(0.040)	(0.040)
Mean	0.095***	0.043***	0.078***	0.137***	0.124***	0.125***	0.112***
	(0.021)	(0.018)	(0.020)	(0.020)	(0.021)	(0.021)	(0.012)

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

Table 6: Returns to Education by Gender for the NLSY79 and the NLSY97 Cohorts

VARIABLES	Men		Women	
	NLSY79	NLSY97	NLSY79	NLSY97
AFQT score	0.109***	0.0540***	0.123***	0.0752***
	(0.00887)	(0.00794)	(0.00771)	(0.00748)
High School	0.125***	0.181***	0.142***	0.204***
	(0.0179)	(0.0197)	(0.0162)	(0.0173)
AA	0.261***	0.393***	0.359***	0.460***
	(0.0357)	(0.0365)	(0.0284)	(0.0323)
BA	0.375***	0.562***	0.472***	0.659***
	(0.0273)	(0.0272)	(0.0257)	(0.0241)
MA or above	0.585***	0.808***	0.689***	0.964***
	(0.0445)	(0.0392)	(0.0442)	(0.0313)
Experience	0.0559***	0.0623***	0.0485***	0.0467***
	(0.00362)	(0.00301)	(0.00367)	(0.00300)
Experience Square	-0.00179***	-0.00215***	-0.00182***	-0.00170***
	(0.000182)	(0.000177)	(0.000202)	(0.000182)
Black	-0.110***	-0.148***	0.0150	-0.0305**
	(0.0150)	(0.0155)	(0.0136)	(0.0137)
Unemployment Rates	-2.182***	-1.008***	-2.336***	-0.698***
	(0.343)	(0.194)	(0.346)	(0.204)
Metro	0.0496***	-0.00377	0.0555***	0.0368***
	(0.0153)	(0.0138)	(0.0131)	(0.0134)
Constant	2.336***	2.183***	2.119***	1.985***
	(0.0427)	(0.0234)	(0.0421)	(0.0220)
Observations	38,069	24,298	31,846	21,487
R-squared	0.210	0.239	0.271	0.367

Note. Table tabulates the estimates from OLS regression of log hourly wage of men and women on various skill correlates. I closely follow Castex and Kogan Dechter (2014) when constructing the sample for the wage regression and include the same set of skill correlates. AFQT score is normalized to account for the difference in the mapping between a pencil-and-paper and computer assisted tests and the age effect (Altonji et al. (2012),Castex and Kogan Dechter (2014)). High School, AA, BA, and MA or above represent the highest degree completed as of age 35. Experience is calculated as age-years of schooling - 6. Data for unemployment rate is from Castex and Kogan Dechter (2014). The sample consists of wage between age 18-35.