Task Requirements, Time Requirements, and the Gender Gap in Jobs and Pay*

Chinhui Juhn†

University of Houston

Yona Rubinstein‡

London School of Economics

July 25, 2022

Preliminary - Please do not cite without permission

Abstract

Non-routine cognitively demanding jobs require both the capacity to cope with complexities as well as long and unusual hours. Women's comparative *dis*advantage in providing hours might offset their comparative advantage in performing cognitively demanding jobs. Using DOT and O*NET data and measures of time requirements at the occupational level we find that non-routine cognitive task requirements are highly correlated with time requirements. Using Census and American Community Survey data from 1960 to 2018 we find that women are less likely to sort into non-routine cognitive tasks, and more likely to sort into routine tasks. In National Longitudinal Survey of Youth (NLSY) data we find that women with high cognitive skills are especially under-represented in non-routine cognitive tasks while they are over-represented in routine tasks. These mismatch effects play an important role in accounting for the slowdown in the convergence of occupational sorting and wages.

*Jeonghyeok Kim and provided excellent research assistance.

†University of Houston. E-mail: cjuhn@uh.edu

‡London School of Economics. E-mail: y.rubinstein@lse.ac.uk

1 INTRODUCTION

There is now a large literature on the task-based approach to studying skill demand (Acemoglu and Autor 2011, Autor and Dorn 2013). Autor, Levy, and Murnane 2003 first introduced the approach to explain how technology complements or replaces workers of different skill type. Non-routine analytical tasks require workers with cognitive skills who are flexible and can adapt to a changing environment. On the other hand, routine tasks require repetitive execution of a set of instructions which can be codified and performed by machines and computers. Autor, Levy, and Murnane 2003 show that computers have replaced workers in routine tasks leading to job polarization (see figure 1). Since then, many papers on a wide array of topics have used the framework.[1]

The growing importance of tasks which require cognitive skills and social skills suggests that women should have done well in this economy. In fact, Galor and Weil 1996, and Welch 2000 predicted that rising returns to cognitive skills should reduce the gender wage gap given women's comparative advantage in cognitive tasks. While this prediction was born out in the 1980s (see, for example, Mulligan and Rubinstein 2008) the wage convergence between men and women have largely stalled since the 1990s, especially among the college-educated. This is particularly puzzling since women have increased their educational attainment relative to men at almost every margin (Goldin, Katz, and Kuziemko 2006).

In this paper, we seek to resolve this puzzle by examining the extent to which women's differential labor supply affects the gender wage gap. The supply margin we focus on is not the extensive margin or even working full-time but rather, the *intensive* margin - the sorting into more demanding and rewarding occupations. In terms of employment and full-time work, women have made great strides over the past five decades. In 2019, they represented 47.0 percent of the work force and 43.5 percent of the full-time work force.[2] It is less clear, however, whether conditional on working full-time, they have entered the same high paying occupations as men.

Recent work by Goldin 2014 and others (Cha and Weeden 2014, Cortes and Pan 2019, Gicheva 2013, Erosa et al. 2017) have brought to attention the return to long hours within certain professional occupations such as medicine, business, and law. These studies posit that within these occupation, there is a non-linear relationship be-

1. For example, Acemoglu and Restrepo 2020 study the impact of automation with the introduction of industrial robots; Hershbein and Kahn 2018 study labor demand on online job postings; Deming 2017 and Borghans, Weel, and Weinberg 2008 study the rising importance of other types of skills such as social or interpersonal skills; and Peri and Sparber 2009, Hurst, Rubinstein, and Shimizu 2020, Beaudry and Lewis 2014 study wage gaps across different groups.

2. U.S. Bureau of Labor Statistics. April 2021. "Women in the labor force: a databook." <https://www.bls.gov/opub/reports/womens-databook/2016/home.htm > (accessed August 16, 2021).

tween hourly wages and hours. Since women are not able to supply the long hours due to family responsibilities, they are paid a lower wage.[3] Studying a recent policy change which capped hours worked by medical residents, Wasserman 2019 finds that women are now more likely to enter specialties which had previously required long hours, showing that hours requirements play an important role in women's occupational decisions.[4]

In this study we bring these two literatures together by examining both task and time requirements at the occupational level. The key idea is that non-routine cognitive tasks require not only cognitive skills but also long, and often non-routine, hours. The non-routine nature of the task makes it difficult to predict and plan the work schedule. Drawing on the original insight of Autor, Levy, and Murnane 2003, we also posit that non-routine nature of the task also makes it difficult to codify the task into a set of instructions so a substitute can easily fill in if a given worker is not available. In contrast, routine tasks do not have such hours requirements since another worker can readily fill in to complete the task. While women may have a comparative advantage in performing cognitively demanding jobs, their comparative *dis*advantage in providing long and unusual hours might offset their advantage. Consequently the time requirements serves as a wedge between able women and cognitively demanding high-paying jobs.

We find that the complementarity between non-routine cognitive task requirements and long hours requirements is robustly confirmed in the data. Using the Dictionary of Occupational Titles (DOT) and, its successor, the Occupational Information Network (O*NET), we find that non-routine cognitive ("Abstract") score of an occupation is strongly positively correlated with the measure of " long hours," which is typically measured as the fraction of males who worked 50 or more hours per week. In contrast, the routine task score is strongly negatively correlated with "long hours." We study the implication of this complementarity between task and hours requirements on the occupational sorting of men and women using the decennial Census and the American Community Surveys over the period 1960-2018. We find that while women converged on men from 1960 to 1990 in terms of entering occupations with a high "Abstract" score, convergence basically stopped since 1990. In 2018, female employment share was approximately 5 percentage points lower in an occupation with a one standard deviation higher "Abstract" score. Remarkably, this gender gap in sorting into "Abstract" tasks is entirely erased once we control for "long hours."

An important insight of the task-based approach is that there is a distinction be-

3. Cubas and Silos 2019 study the importance of the timing of work and wage penalties women suffer in occupations with coordinated work schedules. Adams-Prassl 2020 examines the cost of work interruptions among female workers on an on-line platform.

4. Using a different approach Wiswall and Zafar 2017 and Mas and Pallais 2016 elicit workers' willingness to pay and find that women have a higher willingness to pay for more flexible work schedules.

tween the tasks required on the job and the underlying skills of the workers matched to the job. This distinction is important in exploring which type of workers—in terms of skill level— are found in occupations with different task requirements. We use the 1979 and 1997 waves of the National Longitudinal Survey of Youth to study this question. A major advantage of the NLSY data is that it provides detailed information on proxies for cognitive, non-cognitive and social skills of individuals measured before they enter the labor market. We find that women and men sort differently into "Abstract" and "Routine" tasks based on their underlying abilities. In particular, the gender gap in sorting into "Abstract" tasks is higher among men and women with high cognitive skills which we proxy with the AFQT score. Cross-occupational differences in long work hours again accounts for most of this gap, pointing to the relevancy of time rather than skills constraints. One possibility is that women may just have preferences against "long hours." We find, however, that the gender gap in sorting is especially pronounced after the birth of the first child, which point to differential costs of providing long work hours when childcare responsibilities are present, rather than innate preferences, as the main source of the gender difference.

We find that the "mis-match" of tasks and skills come at a cost. In the last part of our analysis we explore the implications of the gender gap in sorting on wages. Using a sample of white males, we run a wage regression interacting measures of task requirements and underlying worker skills. Workers with AFQT scores that is one standard deviation higher enjoy a 10.4 percent wage premium. They enjoy an additional 3.8 percent premium if they work in occupations with one standard deviation higher "Abstract" task score. However, they lose 3.8 percent if they locate to occupations with 1 standard deviation higher "Routine" task score. Moreover, the size of the positive interaction between tasks and skills increases over time. This suggests that women with high cognitive ability may be penalized by not being able to work in demanding and rewarding occupations, and the size of this penalty has increased over time. Controlling for education and basic demographics, the gender wage gap fell approximately 23 percent from 1990-2018. Adjusting for the gender gap in sorting, the gender wage gap dropped by at least 40 percent, and as much as 60 percent if we make some adjustments for the likely sorting within occupations. The gender differences in sorting on task requirements is especially important among college educated. While the wage gap controlling for demographics and education barely changed among this group, we find that the gender wage gap dropped 18-40 percent once we take sorting into account.

Our paper makes a number of contributions to the existing literature on the gender gap in jobs and pay. First, by bringing time requirements into the task requirement literature, we provide an explanation as to why women were not fully able to take advantage of their underlying comparative advantage in cognitive tasks. Women are still restricted by their inability of provide the requisite hours, even when they are working full-time.

Second, we also advance the literature on flexible work arrangements and the gender pay gap by bringing the following insight: while previous papers including Goldin 2014 had argued that within skilled occupations, women were less likely to enter jobs demanding long-hours and hence suffered a wage penalty, we argue in this paper that *because* of the long-hours, women are actually less likely to enter skilled occupations. That is, we document a real trade-off between hours and the complexity of tasks women perform on their jobs. For high ability women, there is a real trade-off between hours and the amount of underlying skills they put to use in their jobs. While occupations are crude proxies for jobs, one can easily imagine this pattern of sorting into less demanding jobs occurring even within occupations.

Finally, ours is the first paper to explore the gender gap in selection into tasks based on underlying ability, bringing to the foreground the notion of mis-match. Hsieh et al. 2019 document a large increase in aggregate productivity resulting from the movement of women from the home sector to highly skilled occupations. We bring micro-level evidence of this mismatch. Our paper suggests that the amount of misallocation of talent may still be substantial in that the most able women are found in jobs that do not put their talents to the best use.

Our paper is organized as follows. Section 2 describes the data sets we employ as well as the task and skill measures. Section 3 describes the complementarity between task and time requirements. Section 4 reports our results on occupational sorting in the Census and ACS data. Section 5 reports our results on occupational sorting by worker skills and parental status using the NLSY data. Section 6 explores the implications of the occupational sorting on the gender wage gap. Section 7 concludes.

2 Data and Measures

In this section we describe the data sets we utilize in the study as well as our occupation-level measures of task requirements and time requirements. Our goal is not to invent novel new measures of tasks or hours demand. Rather, we view our contribution as bringing attention to the relationship between these well established measures originating from two different strands of research.

2.1 Census and American Community Survey

To examine long run changes in wages and sorting into occupations we use the U.S. decennial Census for 1960-2000, and the annual American Community Surveys (ACS) for 2010-2012 and 2016-2018 available from the Integrated Public Use Microdata Series

(IPUMS). We restrict our sample to employed men and women aged 25-54 who are civilians and do not live in group quarters. We exclude self-employed workers. We measure wages as annual earnings during the previous year divided by the product of weeks worked last year and usual hours per week. Much of our analysis centers on full-time, full-year workers who we define as those who worked 50 or more weeks last year and worked 40 or more hours per week. Finally, we weight the data using the survey weights provided by the Censuses and the ACS.

2.2 National Longitudinal Survey of Youth 1979 and 1997

We also utilize data from 1979 and 1997 waves of the National Longitudinal Survey of Youth (NLSY). A major advantage of the NLSY data is that it provides detailed information on proxies for cognitive, non-cognitive and social skills of individuals measured before they entered the labor market. The NLSY waves are representative surveys of 12,686 and 8,984 individuals who were 15 to 22 years old in 1979 or 13-17 years old in 1997 when they were first surveyed. The surveys were conducted either annually or bi-annually every year since for each cohort. The NLSY79 survey was conducted annually through 1993 while the NLSY97 was conducted annually through 2011.

We restrict the sample to individuals 25 years of age and older who report key labor market outcomes including wages, hours worked, weeks of work, occupation and industry. For our analysis samples, we also exclude respondents with missing values for education and skill measures discussed below. We compute hourly wages by dividing total salary and wages from the prior year by annual hours worked. We index hourly wages to 2010. Following Altonji, Bharadwaj, and Lange 2012 and Deming 2017 we also trim values of deflated hourly wage that are below $2 per hour and above $500 per hour (2010 CPI adjusted). As with the Census/ACS data, we exclude the self-employed.

We use measures of performance on cognitive test and psychometric assessments from the NLSY79 and the NLSY97 to generate a set of unified proxies for cognitive, non-cognitive and social skills. These skill measures were primarily collected before the individual entered the labor market. For a worker's level of Cognitive Skills (COG), we use the respondent's standardized scores on the Armed Forces Qualifying Test (AFQT). This measure was asked of all respondents in their initial wave of the survey and was measured in both the 1979 and 1997 waves. We follow Altonji, Bharadwaj, and Lange 2012 and Deming 2017 to generate age-adjusted AFQT scores. We follow Heckman, Stixrud, and Urzua 2006 to construct measures of non-cognitive skills in NLSY79 and follow Deming 2017 to construct measures of non-cognitive skills in NLSY97. We also use the measures of social skills developed by Deming 2017. Appendix A.1.3 provides more details.

2.3 American Time Use Survey

We use for our analysis the 2003-2018 American Time Use Survey (ATUS). One respondent per household is drawn from the Current Population Survey (CPS) samples and the interviews are conducted 2 to 5 months after the last CPS interview. The ATUS respondent is asked to fill out a time diary over the previous day, recording their activities and starting and ending times. There are 17 aggregate activities. We focus on "work and work-related activities". For each individual we calculate minutes spent on these activities for each hour of the day using information on starting and ending times.[5] The ATUS also contains demographic and labor force information including labor force status and usual hours worked.

2.4 Measures of Task Requirements

Following Autor, Levy, and Murnane 2003 and Autor and Dorn 2013 we use task requirement measures reported for each detailed occupation in the U.S. Department of Labor's Dictionary of Occupational Titles (DOT). The DOT was first constructed in 1939 and updated in subsequent years. We use the 1977 version in this study. Different questions have answers that range on different ordinal scales (e.g., 0-5, 1-7, 0-10, etc.). We follow Autor, Levy, and Murnane 2003 and re-scale average scores by occupation to 0-10 scale that reflects weighted percentile rank in the 1980 distribution of task inputs. We focus on five task measures used in Autor, Levy, and Murnane 2003 and also many others: 1) "GED Math" 2) "Direction, Control, and Planning (DCP)" 3) "Set Limits, Tolerance, or Standards (STS)" 4)"Finger Dexterity (FINGDEX)" and 5) "Eye, Hand, Foot Coordination (EYEHAND)". Higher levels of GED-Math are associated with higher quantitative abstract tasks. Higher levels of DCP are associated with higher levels of abstract thinking associated with management, organizational, and teaching tasks. The literature has equated these task requirements with non-routine analytical and problem solving requirements. We combine GED-Math and DCP by taking simple averages and label this measure "Abstract."

STS measures the adaptability to work in situations requiring setting of limits and measurements and serves as a proxy for routine cognitive tasks. FINGDEX measures the ability to move fingers and manipulate small objects with fingers and serves as a proxy for repetitive routine manual tasks. We again take a simple average of these two

5. We downloaded the ATUS from IPUMS using Create Variable from Scratch option, selecting "Work and Work-Related Activities" (050000-060000) and "Caring for and Helping Household Members" (030000-040000) by time of day (specifying beginning and ending times) and also by site (work place, home, other). We note that "work" does not include travel or commuting time.

measures and label it "Routine." Finally, we use EYEHAND as our measure of manual task requirements. In our analysis we label this measure "Manual."

In recent years the DOT has been replaced by Occupational Information Network (O*NET) sponsored by the U.S. Department of Labor/Employment and Training Administration (USDOL/ETA). While O*NET is continuously updated we use the 1998 O*NET to check the robustness of our results. We adopt the measures used in Acemoglu and Autor 2011 and Denning et al. 2019. Appendix A.1.1 and A.1.2 provide more detail on the construction of our task measures.

Deming 2017 provides the DOT and O*NET task measures by the harmonized 1990 occupation codes developed by David Dorn ("occ1990dd"). We make slight modifications to convert "occ1990dd" codes to harmonized 1990 occupation codes provided by IPUMS in all years starting in 1960.[6]

2.5 Measures of Time Requirements

We use the Current Population Surveys for 2003-2018 to construct occupational level measures of "long hours" following the method introduced by Cha and Weeden 2014, Cortes and Pan 2016 and Cortes and Pan 2017. More specifically, we keep men who are 18-65 years old and worked full-time. For each detailed census occupation, we calculate the share of this group who worked 50 or more usual hours per week on the main job.[7]

One issue with the above measure of "long hours" requirement on the job is that it may not necessarily reflect job requirements (demand) but rather reflect worker responses (supply). One way to address this issue is to use job requirements related to hours (in)flexbility in the O*NET dataset. This is the approach used by Goldin 2014. We downloaded the same 5 characteristics: "Time pressure", "Contact with others", "Establishing and maintaining interpersonal relationships," "Structured versus unstructured work," and "Freedom to make decisions." We create z-scores of each component measure and also create an index of (in)flexibility by averaging the components.[8]

For both measures we first aggregate to detailed 2010 Census occupation codes and convert to the harmonized 1990 occupation codes using the crosswalk provided by IPUMS.

6. We have also checked the robustness of our results using "occ1990dd" codes and found very similar results.

7. Cortes and Pan 2016 restrict their analysis to college educated workers while Cortes and Pan 2017 examine workers with some tertiary education. Accordingly, their measure of incidence of long hours is calculated for men with college degree or with tertiary education. We check the robustness of our results using "long hours" among men with tertiary education in table A.3 in the appendix and find very similar results.

8. We downloaded from the O*NET database on January 15, 2020.

3 Correlation of Task Requirements and Time Requirements

In this section we report our first empirical result showing that jobs which require non-routine analytical skills also require hours to complement those skills. All measures are at the detailed 3-digit 1990 harmonized occupation level. There are 389 occupations overall although the number of unique occupations with non-missing data varies by year.[9]

We first regress task requirement measures on the "long hours" measure to investigate the correlation of these measures across occupations. Figure 2 shows the fitted lines from the regression with the associated 95-percent confidence intervals. The left panel shows the relationship between abstract task and hours as well as the relationship between routine task and hours. The right panel again shows abstract task and hours but also shows the relationship between manual tasks and hours.

The figure shows that abstract task requirements are strongly positively related to "long hours" while routine task requirements are strongly negatively related to "long hours." The right panel shows that manual task requirements are also positively related to "long hours" although the relationship is not very strong.

Table 1 reports the pair-wise correlations. The first row shows that the correlation between "long hours" and "Abstract" is 0.506 while the correlation between "long hours" and "Routine" is -0.397. The bottom panel reports the correlation of the task requirements with the various O*NET characteristics associated with hours inflexibility. Almost all of the measures are positively correlated with "Abstract" and negatively correlated with "Routine." The only exception is "Time Pressure" which seems to be more closely associated with routine tasks.

While "long hours" is a useful proxy, figure 2 showed that manual tasks are also associated with "long hours" suggesting that it does not fully capture the type of hours required by the non-routine abstract tasks. In the following we further probe the hours associated with non-routine abstract tasks using the American Time Use Survey. The ATUS provides a convenient snap shot of a workday which can be compared to the CPS survey question "usual hours worked per week." We run the following regression of hours worked on a workday in the time diary data on task measures:

$$H_{ijt} = \delta_t + \sum_{k=1}^{3} \delta_k Z_{kijt} + \delta_x X_{it} + \delta_h H^{u}_{ijt} + \epsilon_{ijt} \tag{1}$$

where H_{ijt} is hours worked on workday in occupation j. Z_{kijt} is job task requirement k in occupation j observed for person i in period t, and H^{u}_{ijt} is usual hours worked per

9. The number of occupations 1960-2018 are as follows: 233, 308, 383, 385, 339, 333, 322.

week. We also include δ_t which are day of week and year controls, and X_{it} which are dummies for single years of age, detailed education categories and detailed race categories. We define the workday as a diary day where the respondent worked positive hours at either the work site or offsite other than the home.[10] We run the regression on a sample of 25-54 year old men who are full-time wage and salary workers and also non multiple-job holders.

Table 2 shows the results. The first column does not control for "usual hours" while the second column adds it as an additional control. The table shows that "Abstract" task requirements are positively related to hours worked on the diary day, even controlling for "usual hours" while "Routine" exhibits no relationship. The bottom row tests the difference in coefficients between "Abstract" and "Routine" which are significant at conventional levels. If we think of "usual hours" as the contracted number of hours, "Abstract" tasks require hours above and beyond the contracted hours. Our interpretation is that it requires *unusual* hours which are not entirely predictable and for which it is difficult to find a substitute. In our view, these are the type of time demands that women with family responsibilities find particularly difficult to fulfill.

4 SORTING ACROSS OCCUPATIONS BY GENDER AND TASK: ACS

In section 3 we showed that abstract tasks also require "long hours" and even further, *unusual* hours which are hard to anticipate. In this section we investigate the implication of this complementarity of task and time requirements for occupational choice. In particular, since women are more constrained in terms of their ability to supply the requisite hours demanded, this will inhibit their ability to enter occupations which not only require the ability to perform abstract tasks but also demand these long, and often non-routine, hours.

To examine how men and women differentially sort into occupations with various task and time requirements we estimate the following equation:

$$\mathsf{FEMALE}_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt} \tag{2}$$

where FEMALE_{ijt} is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, race dummies, and region dummies.

10. An alternative definition of the workday as a weekday yields very similar results.

The above regression should be thought of as describing the nature of the match between a worker i and the job j in year t. The regression addresses the following question: when we find a worker in a particular occupation, what is the probability that the worker is a female? The specification allows us to examine female shares in occupations with various task requirements at the micro level while controlling for other worker characteristics X. Note that the task measures at the occupation level do not vary by time. However, the sorting in terms of which type of workers are found in particular occupations does change over time which is reflected in the fact that γs have t subscripts. γ_t reflects the female employment share among full-time, full-year wage and salary workers in occupations with the average score on each task measure. Our focus is γ_{kt} which shows how task requirements affect the female employment share and reflects the gender gap in sorting into tasks. We run the regression separately by each census year, pooling 2010-2012 and 2016-18 in the later years.

It is also worth noting that each occupation j is not identified with a single task requirement such as "Abstract," "Routine" or "Manual" but is rather associated with a *vector* of scores on each task requirement k. Also, these task scores are not constrained to sum to one. In other words, an occupation which scores high on "Abstract" task could also score high on "Routine" and "Manual" task measures.

Another issue is whether we should control for education when we consider the gender gap in sorting into occupations with various task requirements. As is now well known, women have made great strides in terms of educational attainment (Goldin, Katz, and Kuziemko 2006), with recent cohorts surpassing men at all levels. One of the ways that women have been able to gain on men in terms of entering high paying occupations has been through educational attainment. Autor and Price 2013 find that from 1970 to 2000, the male-female gap in analytic and interpersonal tasks had shrunk substantially but they do not control for women's gain in educational attainment relative to men. Since the focus of our paper is on gender differences in sorting and pay *conditional* on the skills men and women bring to the labor market, we control for education in the analysis below.[11]

Figure 3 plots the coefficients, γ_{kt}, for different years, which show the effect on female employment shares per standard deviation units of a given task measure, controlling for the other task requirements, and controlling for worker characteristics. Again, our sample consists of men and women who worked full-time, full-year the previous year and we exclude self-employed workers. The figure shows that in 1960 there was an 8 percentage point gap in employment share per z-score of the 'Abstract'

11. In tables A.2 in the appendix, we investigate the robustness of our results when we do not control for education. We find that the gender gap in sorting into "Abstract" tasks is somewhat muted when we do not control for education (-0.034), indicating that women to some extent compensate by acquiring more education. However, we find very similar results in that the gap entirely disappears when we control for "long hours."

task requirement. From 1970 to 1990, women made progress by entering occupations which require 'Abstract' tasks more rapidly than men but this progress slowed down considerably since 1990. The gap stands at approximately 5 percentage points in 2018.

The trends are reversed in terms of "Routine" tasks. In 1980, there was a near 5 percentage points gap in favor of women in employment share per z-score of "Routine" task requirement. This advantage has eroded although the pace of the erosion has slowed in recent years. In 2018, men and women were about equally likely to be in "Routine" tasks. Finally, women are far less likely than men to be found in manual tasks with the gap hovering around -15 percentage points.

Figure 4 explores the extent to which controlling for other task requirements matter. We find that controlling for other task measures *does* make a considerable difference. For example, without other task controls, full-time, full-year working women are nearly at parity relative to men, while with other task controls, they are approximately 5 percentage points behind per z-score in abstract task measure. What explains this difference? The difference lies in the fact that women are far more likely to be in occupations with high "Routine" scores and far less likely to be in occupations with high "Manual" scores. Since occupations with high "Routine" scores generally score higher on "Abstract" than those with high "Manual" scores, women are unconditionally more likely to be in high "Abstract" occupations. However, conditional on these other task measures, women are still considerably behind in terms of locating in abstract occupations.[12]

4.1 Gender Gap in Sorting with Hours Control

We now examine the extent to which hours requirements play a role in explaining this gender gap in sorting into high "Abstract" occupations. The complementarity between abstract task and hours requirements we explored earlier suggests that if women are not able to supply the same type of hours as men, they would be placed at a disadvantage in terms of entering these occupations. To explore this possibility we amend the regression, equation 2, with the following:

$$\mathrm{FEMALE}_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{ht} \mathrm{HOURS}_{ijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt} \tag{3}$$

where HOURS_{ijt} is the "long hours" requirement in occupation j, expressed in z-scores.

12. Figure A5 in the appendix also explores the extent to which the gender gap in sorting into abstract tasks differs when we include all workers (including those who work part-year and part-time). Among all workers we see a more dramatic convergence of women into abstract tasks, suggesting that women's gains in labor force attachment was a contributing factor. However, the remaining gender gap is somewhat larger in this case.

The left panel of figure 5 plots the coefficients γ_{kt} we showed earlier without the hours control whereas the right panel shows the coefficient with the "long hours" control. The figure shows that controlling for hours increases female employment share in "Abstract" tasks while reducing female employment share in "Routine" tasks. Remarkably, the figure illustrates that the tendency of women to sort into "Routine" tasks versus "Abstract" tasks relative to their male counterparts is entirely due to the hours requirement.

5 Sorting Across Occupations by Gender and Skill: NLSY

The complementarity of "Abstract" task requirements and hours requirements implies not only that women will be placed at a disadvantage in terms of entering abstract occupations, it also implies that the margin of disadvantage would be even higher for highly skilled women relative to equally skilled men. This is true as long as the inability to supply hours does not vary by ability but is equally binding for high and low skilled women. Our analysis so far did not explore this heterogeneity by underlying skill level. For this exercise we turn to the NLSY data which has the advantage of containing detailed information on proxies for cognitive, non-cognitive and social skills of individuals measured before they entered the labor market.

5.1 Gender Gap in Sorting with Hours Control: NLSY

Before investigating the heterogeneity by underlying skills, we first examine whether our previous results using the Census and ACS data can be duplicated in the NLSY data. Namely, we examine whether controlling for "long hours" significantly reduces the gender gap in sorting into abstract tasks.

For the NLSY we pool across both the 1979 and 1997 waves and estimate equation 3 but now pooling across all the years. We keep men and women who are 25+ years old when we observe them employed in a particular occupation. This essentially results in using data over the period 1981-2016. We again select full-time, full-year workers who are not self-employed. To be included in the regression, individuals also have to have non-missing cognitive (AFQT), non-cognitive, and social skill measures. We weight by sampling weights and cluster standard errors at the person level.

Table 3 reports the results. Column (1) reports the results without controlling for worker skills or "long hours." Similar to our results using the Census and ACS, the gender gap in abstract tasks is approximately 5 percentage points per z-score in "Abstract" task requirement. Again, women are slightly more likely to be in "Routine" tasks and significantly less likely to be in "Manual" tasks. The constant term reflects the fact that the female share among full-time, full-year working population with average task scores is approximately 42 percent. Column (2) adds worker skills, AFQT, non-cognitive skills, and social skills, as additional controls. The coefficients on task measures do not change much which shows that skills are not the reason women are under-represented in "Abstract" and over-represented in "Routine." Column (3) adds the "long hours" control only without worker skills. With the addition of "long hours" the coefficient on "Abstract" is reduced to -.011 and is no longer significant while the coefficient on "Routine" turns negative. The coefficient on "long hours" is significantly negative as we expected based on the previous literature (Goldin 2014 and Cortes and Pan 2019). The final column also adds the worker skills, AFQT, non-cognitive skills, and social skills, as additional controls. The coefficient on "Abstract" is further reduced.

Figure 6 provides a plot of the coefficients (with the associated 95% confidence intervals) reported in table 3. The top panel reports the coefficients including the constant term while the bottom panel reports the coefficients without the constant term.

5.2 Gender Gap in Sorting by Worker Cognitive Skill: NLSY

So far our results indicate that women are less likely to sort into "Abstract" tasks and more likely to sort into "Routine" tasks, a pattern which appears to be due to the complementarity of task and hours requirements. In the previous tables we established that simply controlling for worker skills does not erase the gender gap. In this section we explore whether gender gap in sorting into "Abstract" tasks varies by worker skill– that is, we explore the *interaction* between the "Abstract" task requirement and worker cognitive skill level.

We simplify by categorizing job requirements and skills into above and below average and estimate the following regression:

$$\begin{aligned} \mathrm{FEMALE}_{ijt} &= \gamma_0 + \gamma_{kA} D^{Abs}_{ijt} + \gamma_{kR} D^{Rou}_{ijt} + \gamma_{kM} D^{Man}_{ijt} \\ &\quad + \sum_{l=1}^{3} \gamma_{A'} \mathrm{SKILL}_{ilt} + \gamma_h \mathrm{HOURS}_{ijt} \\ &\quad + \gamma_x X_{it} + \varepsilon_{ijt} \end{aligned} \tag{4}$$

where D^{Abs}_{ijt}, D^{Rou}_{ijt}, D^{Man}_{ijt} are dummy variables equal to one if the standardized "Abstract" task score, "Routine" score, or "Manual" score is > 0 in occupation j. Table 4 reports the results. The first three columns report results for workers with below average AFQT scores, while the last three columns report results for workers with above average scores. In column (1) we find that among low ability workers, women are somewhat over-represented in "Abstract" occupations. As we found in figure 4, controlling for other task measures matters. Among lower ability workers, women are highly unlikely to locate in "Manual" tasks, which have low "Abstract" task scores. Once we condition for "Manual" tasks in column (2) women's advantage in "Abstract" disappears. In column (3) we include hours controls, and the coefficient on D_{ijt}^{Abs} again becomes strongly positive and now is significant. Among low ability workers, women have a comparative advantage in "Abstract" tasks which becomes more apparent when we control for "long hours." The sorting of low ability women into cognitively demanding tasks however may not necessarily produce a return however. We investigate the returns to sorting into tasks by ability in the next section.

The story is different among high ability workers as shown in columns (4) to (6). In column (4) we find that high ability women are under-represented in "Abstract" tasks by 5.3 percentage points. Controlling for other tasks does not change the story. In fact, when we include "Routine" and "Manual" tasks, the coefficient on "Abstract" becomes even more negative and the coefficient on "Routine" is positive. When we control for long hours in column (6), the under-representation of women in "Abstract" tasks and over-representation of women in "Routine" tasks disappears. The table underscores the point that high ability women do not sort into "Abstract" tasks because of the long hours associated with "Abstract" tasks. We examine in the next section the implications of this mismatch of skills and job task requirements for women in terms of wages.

5.3 Gender Gap in Sorting by Parental Status: NLSY

A key part of our argument is that women have a higher cost of providing long and unusual hours due to their larger share of childcare and household responsibilities. An implication of this is that we should observe a larger gender gap in sorting into "Abstract tasks" among men and women with children. A large literature has documented the presence of child penalties. For example, Kleven, Landais, and Søgaard 2019 estimate that 10 years after the birth of the first child the impact of children on the labor market outcomes of women relative to men in Denmark is about 20 percent in earnings, a third of which is due to lower hourly wages.[13] In this section we explore the extent to which the gender gap in sorting into "Abstract" tasks is associated with the presence of children.

We modify our main sorting equation as laid out in equation 2 by introducing two variables to indicate parental status. One is whether the individual i ever has a child, PARENT_i, and another is whether the individual i currently has a child in period t,

13. Among others, Angelov, Johansson, and Lindahl 2016 document similar wage penalties associated with children in Sweden; Kuziemko et al. 2018 in U.S. and U.K.

AFTER_{it}, indicating that t is after the birth of the first child. We interact these variables with the task measures. We run the following regression:

$$\mathrm{FEMALE}_{ijt} = \gamma_0 + \sum_{k=1}^{3} \gamma_k Z_{kijt} + \gamma_c \mathrm{PARENT}_i + \gamma_{c'} \mathrm{AFTER}_{it}$$
$$+ \sum_{k=1}^{3} \gamma_{kc} Z_{kijt} * \mathrm{PARENT}_i + \sum_{k=1}^{3} \gamma_{kc'} Z_{kijt} * \mathrm{AFTER}_i t + \gamma_x X_{it} + \varepsilon_{ijt} \quad (5)$$

To explore the extent to which "long hours" requirement in occupation j can account for the gender gap in sorting, we again add HOURS_{ijt} as well as its interaction with PARENT_i and AFTER_{it}.

Table 5 reports the results. Column (1) repeats the results shown in column (2) of table 3 indicating that women are 0.045 percentage points behind in "Abstract" tasks. In column (2) we add the indicator for PARENT_i as well as the interaction with "Abstract". The column shows that the effect is almost all driven by mothers relative to fathers. Among parents, women are -0.047 percentage points behind in "Abstract" tasks. Among non-parents, the gap is only -0.012 and not statistically significant. In column (3) we investigate the effect of current parental status by adding the indicator AFTER_{it} and interactions. Note that the omitted group in this case consists of individuals who are never parents as well as those who will eventually become parents. Among those who are currently non-parents, there is a marginally significant gap of -0.018. However, this effect is much smaller than the effect of current parental status, -0.048, which is strongly negative and statistically significant. This indicates that while there may be some anticipatory effect among future mothers not entering occupations with "Abstract" task requirements due to anticipated discrimination or career interruptions, it is the actual motherhood itself that mainly drives the sorting. Column (4) provides a horse-race between the two and we find that it is again the current parental status that dominates. Column (5) and column (6) add controls for "long hours", both the main effect as well as the interaction with AFTER_{it} and PARENT_i. Similar to the results in table 3 the gender gap in sorting into "Abstract" tasks completely disappears. We view this as evidence that hours requirements on the current job, particularly after the arrival of the first child, is the main driver of the gender gap in sorting into "Abstract" tasks.

6 Sorting into Jobs and the Gender Gap in Observed Wages

A large body of research analyzed the growing presence of women in labor markets and their increasing wages relative to men over the past half century. Several puzzles emerged in the overall story of gender convergence. First, during the 1980s and 1990s when wage inequality was rising both across and within demographic groups, there was growing gender equality. Blau and Kahn 1997 suggested that women were "swimming upstream." Mulligan and Rubinstein 2008 argued that women's observed wages improved relative to men's wages *because* of, rather than despite of, growing wage inequality. They pointed to the impact of rising returns to skills on the change in sorting of women into the labor market – selection at the "extensive margin"– and its implications on the latent skills of the typical working women (relative to men) during this period.

Another puzzle, as pointed out earlier in the introduction, is the stalled wage convergence between men and women since the 1990s, especially among the college educated. In explaining this puzzle, we think that selection or sorting again played an important role. Rather than selection at the "extensive margin," however, what is important for the later period is selection at the "intensive margin" – the differential sorting of men and women into jobs among full-time, full-year workers. In this section we study the implication of the gender gap in sorting into jobs on wages.

6.1 Empirical Setting

Perhaps the main takeaway from our findings so far is the mis-allocation of women's talent. High ability women sort into less demanding occupations than their male counterparts. To identify the role of selection at the intensive margin on the measured gender wage gap, we introduce a simplified, reduced form sorting and wage model to be taken to the data.

Workers are paid hourly wages based on their supply of skills (cognitive, non-cognitive, and social), S_i', the task requirements they perform in job j, T_j', and the quality of matching between tasks requirements and skills, $(T_j S_i)$. Hourly wages (in logs) can be approximated with the following linear equation:

$$W_{ij} = \beta_0 + \beta_F Female_i + S_i' \beta_S + T_j' \beta_T + T_j S_i' \beta_{TS} + \mu_{ij} \quad (6)$$

Wages also vary with gender ($Female_i$) and i.i.d. mean-zero worker-specific shock (μ_{ij}).

Up to now we have used the words "job" and "occupation" interchangeably. In this section we explicitly distinguish between these two concepts and introduce new notation to keep track. The wide range of J jobs in the economy is aggregated in available data into O occupational bins ($O < J$). Consequently, the task requirements performed by worker i in occupational bin o is equal to (i) the average task requirement in occupation o, (T_o) and (ii) a mean-zero within occupational bin component (t_{oi}):

$$T_{ji} = T_o + t_{oi}.$$

The aggregation into O occupational bins implies that only the first term is observed by the econometrician. If all workers within a bin perform the same job then $t_{oi} = 0$ and $T_{oi} = T_o$. Otherwise, the wage of worker i in occupational bin o reflects observed and unobserved task requirements in addition to worker specific shocks:

$$W_{io} = \beta_0 + \beta_F Female_i + S_i^{'}\beta_S + T_o^{'}\beta_T + T_o S_i^{'}\beta_{TS} + \eta_{io}. \tag{7}$$

The error term in this case is no longer gender neutral:

$$\eta_{io} = t_{io}^{'}\beta_T + t_{io} S_i^{'}\beta_{TS} + \mu_{ij}.$$

Altonji, Elder, and Taber 2005 introduce a method for correcting for selection on unobservables using the selection on observable characteristics. In our setting we use the gender gap in sorting *between* occupations, (T_o) and ($T_o S_i$), to correct for the gender gap in sorting *within* occupational bins, (t_o) and ($t_o S_i$). When gender gap in sorting is proportional to wage rewards, then the gender gap in sorting on unobserved task requirements is proportional to the gender gap in sorting on observed task requirements. Following Altonji, Elder, and Taber 2005, consider a linear projection of workers' gender on the task requirements they perform, weighted by labor market returns:

$$\hat{F}_{io} = \alpha_F(S_i^{'}\beta_S + T_o^{'}\beta_T) + \alpha_F(t_{oi}^{'}\beta_T + t_{oi}S_i^{'}\beta_{TS}) = \alpha_F \hat{W}_i + \alpha_F \hat{w}_i. \tag{8}$$

Selection on unobserved tasks, weighted by labor market prices ($\hat{w}_i$), is governed by the same parameter as selection on observed tasks requirements ($\hat{W}_i$). In this case the gender gap in sorting within occupational bins is proportional to the gender gap in sorting between occupational bins:[14]

$$\Delta\hat{w}_F = \Delta\hat{W}_F \frac{var(\hat{w}_i)}{var(\hat{W}_i)} = \Delta W_F M. \tag{9}$$

The upper bound of the multiplier M is the ratio of the residual variation over the

14. The condition corresponds to condition 4 in Altonji, Elder, and Taber 2005.

variation explained by observed tasks, $\frac{(1-R_W^2)}{R_W^2}$.

$$0 \leqslant M \leqslant \frac{(1-R_W^2)}{R_W^2} \tag{10}$$

One issue with the above as pointed out by Altonji, Elder, and Taber 2005 is that when R_W^2 is relatively low, this results in a large correction for unobservables for a minor gap in observables. With this in mind, we take a conservative approach and bound the multiplier. For any $R_W^2 \leqslant 0.5$ we truncate the adjustment multiplier at 1:

$$\hat{M} = \min{(1-R_W^2)/R_W^2, 1}$$

By imposing this inner bound we allow selection on observed variables to project selection on unobserved variables when the observed variables account for most of the variation in wages within gender groups. Yet, we truncate the adjustment multiplier at 1 when the observed variables do not account for most of the variation in wages within gender groups. Practically, in this paper the adjustment multiplier in all specifications equals 1, that is $\hat{M} = 1$.

There are three main takeaways to take to the data. First, occupational fixed effects may be insufficient in accounting for gender differences in sorting across jobs given the aggregate nature of the occupational bins we observe in the data. Second, gender gaps in occupational task requirements point to gender gaps in the same direction on task requirements within occupational bins. Third, the sorting bias increases with gender gaps in the quality of matching, which in our simple model is captured by the interaction of skills and tasks,$(T_j S_i)$.

6.2 The Gender Wage Gap

Employing our empirical setting we quantify the female-male gap in log hourly wages among workers by estimating the following equation:

$$W_{iot} = \beta_{Ft} Female_i + \sum_{k=1}^{3} \beta_{kt} T_{ko} + \sum_{s=1}^{3} \beta_{st} S_{is} + \sum_{k=1}^{3} \sum_{s=1}^{3} \beta_{kst} T_{ko} S_{is} + X_{it}' \beta_{xt} + \mu_{iot}. \tag{11}$$

Skills (S_i) include measures of cognitive, non-cognitive and social skills at the individual level. Each occupation is characterized by the intensity of "Abstract," "Routine," and "Manual" task requirements (T_o). Demographics such as race, age, education and location are captured by the vector (X_i). Skills and tasks measures are normalized to z-scores. The quality of matching supply to demand is captured by the interactions between skills and tasks. The impact of skills, tasks, demographics, and gender might vary over time.

We preview the gender gap analysis with estimation of the parameters of the wage equation among prime-aged full-time full-year working white males, which we refer to hereafter as "male prices."[15] We begin with NLSY data. Once again, the advantage of the NLSY data is that we directly observe individual skills, which is important for estimating the interaction of skills and tasks (i.e. the match quality).

Table A.4 reports the results from the NLSY. We estimate the equation for all years (1981-2016) as well as separately for the 1990s and the 2000s. The first column reports our estimates for all years, the second column for the 1990s, and the third column reports the change in the coefficients between the 1990s and the 2000s. We highlight three main findings. First, individual skills and occupational task requirements predict wages. A one standard deviation higher AFQT score is associated with 10.4 percent higher earnings. Higher task measures also command higher wage premiums. A one standard deviation higher "Abstract" score, holding other scores constant, leads to an 8.9 percent premium. Second, the *quality of the match* matters. We find a positive interaction between cognitive skills and "Abstract" (0.037). Interestingly, the interaction between higher cognitive score and higher "Routine" score is negative (-0.038). This suggests that the mis-match of skills to tasks is likely to be especially costly for high ability women. Finally, the returns to the quality of the match has increased over time. The interaction effect of cognitive skill and "Abstract" mattered relatively little in the early period but increased to 0.061 in the 2000s. These findings suggest that the same gender gap in sorting is associated with a larger gender gap in wages in the 2000s.

While the NLSY may be reliable at the cross-sectional level, it may not be adequate for studying trends. To study trends, we turn to ACS-Census data. When we turn to the ACS-Census data, we impute individual skills by occupation (3-digits), gender, race and decade using the NLSY waves. The imputation of skills limits the extent to which we can examine trends prior to 1990. In the following analysis we examine the impact of the gender gap in sorting on the gender wage gap beginning in 1990. Table A.5 in the appendix reports the coefficients from the Mincerian wage regression on the sample of white males. Table A.6 also reports the coefficients from the full sample including women and both races.

In the following section we compare the raw gender wage gap to various residual wage gaps after correcting for differences in observed characteristics evaluated at male prices. Namely, we report residuals $(W^f - W^m) - \beta^m(X^f - X^m)$ where superscripts f and m refer to female and male respectively. We progress in steps, adding additional

15. This corresponds to estimating the selection effects under the null of no Catholic school effect in Altonji, Elder, and Taber 2005. They use the 8th grade sample for this purpose. In our case, we estimate the selection effects under the null of no female effect which corresponds to using the male sample.

observed characteristics with each specification: demographic controls including education; demographics plus skills; demographics plus tasks; demographics plus tasks interacted with cognitive skills; fully interacted model interacting all skills with tasks. In the final specification we also correct for sorting within occupations as laid out in section 6.1.

Figure 7 shows our results. The top panel plots the log differences in female and male wages while the bottom panel plots changes since 1990. Our results show that taking into account the gender gap in sorting had little impact in 1990, but notable effects later. Note, however, that the addition of either skills or tasks alone does not explain much. Rather, it is the *interaction* of the two that explains why the gender gap did not reduce further since 1990. Controlling for demographics and education, the gender gap in 2018 was approximately 23 log points. Taking into account the gender gap in sorting, the wage gap is 17 log points, or about 25 percent lower. In our view this is a conservative estimate since this does not take into account the likely gap in sorting within occupations. If we take into account the sorting within occupations, the gender gap is 12 log points in 2018, more than 40 percent lower. In terms of changes, the gender gap decreased from 30 log points to 23 log points when controlling for demographics and education, a drop of 23 percent. Accounting for gender differences in sorting, the gender wage dropped by at least 40 percent, from approximately 30 log points in 1990 to 17 log points in 2018. If we take into account the sorting within occupations, the wage gap dropped 60 percent from 28 log points in 1990 to 12 log points in 2018. Women gained relative to men as they became more educated. However, much of this gain was reversed due to the gap in sorting, and in particular, due to the fact that wage penalty associated with mis-match appears to have increased over time.

Figure 8 and 9 show results separately for less than college and college graduate populations. The gender differences in sorting on task requirements is especially important among college educated. Accounting for education, the female-male wage gap among college graduate workers barely changed between 1990 and 2018, in contrast to the drop of approximately one-third among the less educated workers. Accounting for sorting offers a different perspective. The female-male wage gap among college graduate workers dropped 18-40 percent taking into account the sorting, largely due to the fact that sorting was already important for this group in 1990 and the raw gap was considerably larger (30 log points) than the adjusted gaps (22 to 17 log points). Our analysis also highlights the fact that while the raw gap is larger among college educated compared to less educated, this is reversed once we take the gender gaps in sorting into account.

7 Conclusion

Using the DOT and O*NET task measures as well as standard measures of the "long hours" requirement, we find that non-routine "Abstract" tasks require not only cognitive skills but long hours. We hypothesize that women's comparative *dis*advantage in supplying these hours, offsets their comparative advantage in performing cognitively demanding tasks. We find robust empirical evidence supporting our hypothesis. Using Census and ACS data, we find that women are less likely to sort into "Abstract" tasks and more likely to sort into "Routine" tasks. Remarkably, the gap disappears once we control for occupation level hours requirements. Using NLSY data, we find that women with high AFQT scores are particularly under-presented in cognitively demanding "Abstract" tasks while they are over-represented in "Routine" tasks. This mis-match between ability and tasks is costly in terms of wages, and played an important role in slowing down wage convergence.

The trade-off between hours and complexity of tasks we document in this paper is reminicent of recent papers which show that women make constrained choices in terms of job search and labor supply (Barbanchon, Rathelot, and Roulet 2021, Bolotnyy and Emanuel 2019, Cook et al. 2018). While women seemingly "leave money on the table" in these papers, the emphasis is on constraints. Alternatively other papers have emphasized underlying differences in bargaining ability or aversion to competition (Card, Cardoso, and Kline 2015, Roussille 2021, Biasi and Sarsons 2020, Niederle and Vesterlund 2007). Ultimately both channels likely owe their origins to slowly evolving gender norms which dictate greater household responsibilities as well as more passive behaviors for women.

In this paper we have put forward the idea that the complementarity between task and time requirements are inherently driven by the production function. An alternative interpretation of the "long hours" requirements is that it is a sorting mechanism to identify the most committed workers (Landers, Rebitzer, and Taylor 1996) or possibly even an artificial barrier to keep women away. Wasserman 2019 documents that capping hours requirements among medical residents did not lead to ostensibly worse patient health outcomes, suggesting that not all time requirements are dictated by productivity concerns. We view uncovering the reasons for these time requirements as a fruitful avenue for future research.

REFERENCES

Acemoglu, D., and D. Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." Chap. 12, 1st ed., 4:1043–1171. Handbook of Labor Economics. Elsevier.

Acemoglu, Daron, and Pascual Restrepo. 2020. "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy* 128 (6): 2188–2244.

Adams-Prassl, Abigail. 2020. *The Gender Wage Gap on an Online Labour Market: The Cost of Interruptions.* CEPR Discussion Papers 14294. Centre for Economic Policy Research.

Altonji, Joseph, Todd Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–184.

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange. 2012. "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics* 30 (4): 783–828.

Angelov, Nikolay, Per Johansson, and Erica Lindahl. 2016. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34 (3): 545 –579.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics* 118 (4): 1279–1333.

Autor, David H., and Brendan Price. 2013. *The Changing Task Composition of the US Labor Market: An Update of Autor, Levy, and Murnane (2003).* Technical report.

Barbanchon, Thomas Le, Roland Rathelot, and Alexandra Roulet. 2021. "Gender Differences in Job Search: Trading off Commute against Wage." *The Quarterly Journal of Economics* 136 (1): 381–426.

Beaudry, Paul, and Ethan Lewis. 2014. "Do Male-Female Wage Differentials Reflect Differences in the Return to Skill? Cross-City Evidence from 1980-2000." *American Economic Journal: Applied Economics* 6 (2): 178–194.

Biasi, Barbara, and Heather Sarsons. 2020. *Flexible Wages, Bargaining, and the Gender Gap.* NBER Working Papers 27894. National Bureau of Economic Research, Inc.

Blau, Francine D, and Lawrence M Kahn. 1997. "Swimming Upstream: Trends in the Gender Wage Differential in 1980s." *Journal of Labor Economics* 15 (1): 1–42.

Bolotnyy, Valentin, and Natalia Emanuel. 2019. *Why Do Women Earn Less Than Men? Evidence from Bus and Train Operators.* Technical report.

Borghans, Lex, Bas ter Weel, and Bruce Weinberg. 2008. "Interpersonal Styles and Labor Market Outcomes." *Journal of Human Resources* 43 (4).

Card, David, Ana Cardoso, and Patrick Kline. 2015. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women*." *The Quarterly Journal of Economics* 131 (October): qjv038.

Cha, Youngjoo, and Kim A. Weeden. 2014. "Overwork and the Slow Convergence in the Gender Gap in Wages." *American Sociological Review* 79 (3): 457–484.

Cook, Cody, Rebecca Diamond, Jonathan Hall, John A. List, and Paul Oyer. 2018. *The Gender Earnings Gap in the Gig Economy: Evidence from over a Million Rideshare*

Drivers. NBER Working Papers 24732. National Bureau of Economic Research, Inc, June.

Cortes, Patricia, and Jessica Pan. 2016. *Prevalence of Long Hours and Skilled Women's Occupational Choices.* Technical report. Boston University, Questrom School of Business.

———. 2017. "Cross-Country Evidence on the Relationship between Overwork and Skilled Women's Job Choices." *AEA Papers and Proceedings* 107:105–09.

———. 2019. "When Time Binds: Returns to Working Long Hours and the Gender Wage Gap among the Highly Skilled." *Journal of Labor Economics* 37 (2): 351–398.

Cubas, amd Juhn Chinhui, German, and Pedro Silos. 2019. *Coordinated Work Schedules and the Gender Wage Gap.* NBER Working Papers 26548. National Bureau of Economic Research, Inc.

Deming, David J. 2017. "The Growing Importance of Social Skills in the Labor Market." *The Quarterly Journal of Economics* 132 (4): 1593–1640.

Denning, Jeffrey T., Brian A. Jacob, Lars Lefgren, and Christian vom Lehn. 2019. *The Return to Hours Worked Within and Across Occupations: Implications for the Gender Wage Gap.* Technical report.

Erosa, Andrés, Luisa Fuster, Gueorgui Kambourov, and Richard Rogerson. 2017. *Hours, Occupations, and Gender Differences in Labor Market Outcomes.* NBER Working Papers 23636. National Bureau of Economic Research, Inc, July.

Galor, Oded, and David N Weil. 1996. "The Gender Gap, Fertility, and Growth." *American Economic Review* 86 (3): 374–387.

Gicheva, Dora. 2013. "Working Long Hours and Early Career Outcomes in the High-End Labor Market." *Journal of Labor Economics* 31 (4): 785 –824.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–1119.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20 (4): 133–156.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411–482.

Hershbein, Brad, and Lisa B. Kahn. 2018. "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings." *American Economic Review* 108 (7): 1737–1772.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87 (5): 1439–1474.

Hurst, Eric, Yona Rubinstein, and Kazuatsu Shimizu. 2020. *Task Based Discrimination.* Technical report.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11 (4): 181–209.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington. 2018. *The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?* Working Paper, Working Paper Series 24740. National Bureau of Economic Research.

Landers, Renee M, James B Rebitzer, and Lowell J Taylor. 1996. "Rat Race Redux: Adverse Selection in the Determination of Work Hours in Law Firms." *American Economic Review* 86 (3): 329–348.

Mas, Alexandre, and Amanda Pallais. 2016. *Valuing Alternative Work Arrangements.* Technical report. Princeton University, Department of Economics.

Mulligan, Casey B., and Yona Rubinstein. 2008. "Selection, Investment, and Women's Relative Wages Over Time." *The Quarterly Journal of Economics* 123 (3): 1061–1110.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122 (3): 1067–1101.

Peri, Giovanni, and Chad Sparber. 2009. "Task Specialization, Immigration, and Wages." *American Economic Journal: Applied Economics* 1 (3): 135–169.

Roussille, Nina. 2021. *The Central Role of the Ask Gap in Gender Pay Inequality.* Technical report. University of California, Berkeley.

Wasserman, Melanie. 2019. *Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents.* Technical report.

Welch, Finis. 2000. "Growth in Women's Relative Wages and in Inequality among Men: One Phenomenon or Two?" *American Economic Review* 90 (2): 444–449.

Wiswall, Matthew, and Basit Zafar. 2017. "Preference for the Workplace, Investment in Human Capital, and Gender*." *The Quarterly Journal of Economics* 133, no. 1 (August): 457–507.

Table 1: Correlation of Task and Time Requirements

Time Requirements	Abstract	Routine
Long Hours	0.506*	-0.397*
Contact with Others	0.180*	-0.271*
Establish and Maintain Relationships	0.509*	-0.392*
Freedom to Make Decisions	0.529*	-0.242*
Structured vs. Unstructured	0.580*	-0.200*
Time Pressure	-0.070	0.216*
O*NET Index	0.558*	-0.282*

Source: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and Current Population Surveys 2003-2018. Task measures are from DOT 1977. We use the classifications used by Autor, Levy, and Murnane 2003 and define "Abstract" as the average of "GED Math" and "Direction, Control, and Planning (DCP)". "Routine" is the average of "Set Limits, Tolerance, or Standards (STS)" and "Finger Dexterity (FINGDEX)." "Manual" corresponds to "Eye, Hand, Foot Coordination (EYEHAND)". We use the 2003-2018 CPS to define "long hours" at the occupation level. We follow Cortes and Pan 2016 and define "long hours" as the share of men 18-65 years old who worked 50 or more usual hours per week at the main job. We follow Goldin 2014 and downloaded 5 characteristics from the O*NET database that measure job in-flexibility:"Time pressure", "Contact with others", "Establishing and maintaining interpersonal relationships," "Structured versus unstructured work," and "Freedom to make decisions." We create z-scores of each component measure and also create an index of (in)flexibility by averaging the components, which we label "O*NET Index."

Table 2: Hours Worked on a Workday, Wage and Salary Men 25-54

Outcome is Hours Worked on a Workday	(1)	(2)
Abstract	0.170***	0.094***
	(0.044)	(0.033)
Routine	-0.115***	-0.007
	(0.036)	(0.025)
Manual	0.139***	0.084**
	(0.052)	(0.038)
Usual Hours Worked		0.092***
		(0.003)
Abstract - Routine	0.285***	0.101**
Observations	16917	16917

Source: American Time Use Surveys 2003-2018. We regress hours worked during the workday on task requirements in the occupation. The sample consists of 25-54 year old men who are full-time wage and salary workers and also non multiple-job holders. We define the workday as a diary day where the respondent worked positive hours at either the work site or offsite other than the home. We report δ_k from the following regression $H_{ijt} = \delta_t + \sum_{k=1}^{3} \delta_k Z_{kijt} + \delta_x X_{it} + \delta_h H^u_{ijt} + \epsilon_{ijt}$ where H_{ijt} is hours worked on workday in occupation j. Z_{kijt} is job task requirement k in occupation j observed for person i in period t, and H^u_{ijt} is usual hours worked per week. We also include δ_t which are day of week and year controls, and X_{it} which are education, age, and race dummies. Task measures are from DOT 1977. We define "Abstract" as the average of "GED Math" and "Direction, Control, and Planning (DCP)". "Routine" is the average of "Set Limits, Tolerance, or Standards (STS)" and "Finger Dexterity (FINGDEX)."

Table 3: Female Share and Job Task and Time Requirements:
NLSY Salaried FTFY Workers 25+ 1981-2016

Outcome is Female Dummy	(1)	(2)	(3)	(4)
Abstract	-0.049***	-0.045***	-0.011	-0.008
	(0.006)	(0.006)	(0.007)	(0.007)
Routine	0.015***	0.016***	-0.008*	-0.007
	(0.005)	(0.005)	(0.005)	(0.005)
Manual	-0.150***	-0.151***	-0.136***	-0.137***
	(0.005)	(0.005)	(0.005)	(0.005)
Long hours			-0.072***	-0.070***
			(0.007)	(0.006)
Constant	0.419***	0.427***	0.421***	0.428***
	(0.007)	(0.007)	(0.007)	(0.007)
Worker Skills		X		X
Observations	65778	65778	65778	65778

Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1979 and 1997 waves of the National Longitudinal Surveys of Youth. DOT 1977 task measures by occupation are merged with NLSY79 and NLSY97. We keep men and women who are 25+ years old observed employed over 1981-2016 period. We again select full-time, full-year workers who are not self-employed. To be included in the regression, individuals also have to have non-missing cognitive (AFQT), non-cognitive, and social skill measures. We weight by sampling weights and cluster standard errors at the person level. The table reports γ_k from the following regression : $FEMALE_{ijt} = \gamma_0 + \sum_{k=1}^{3} \gamma_k Z_{kijt} + \gamma_h HOURS_{ijt} + \gamma_x X_{it} + \varepsilon_{ijt}$. The table reports coefficients with and without including "long hours" controls and controls for skill measures. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Female Share and Job Task and Time Requirements:
NLSY Salaried FTFY Workers 25+ 1981-2016

Outcome is $D_{(Female)}$	Cognitive⩽0			Cognitive>0		
	(1)	(2)	(3)	(4)	(5)	(6)
$D_{(Abstract>0)}$	0.038**	-0.023	0.088***	-0.053***	-0.062***	0.010
	(0.019)	(0.019)	(0.019)	(0.013)	(0.013)	(0.014)
$D_{(Routine>0)}$		-0.064***	-0.127***		0.022*	-0.037***
		(0.015)	(0.016)		(0.012)	(0.012)
$D_{(Manual>0)}$		-0.283***	-0.263***		-0.154***	-0.158***
		(0.016)	(0.015)		(0.013)	(0.012)
Noncognitive	-0.019	-0.023**	-0.018	-0.027***	-0.026***	-0.022**
	(0.013)	(0.012)	(0.011)	(0.009)	(0.009)	(0.009)
Social	-0.007	-0.007	-0.005	0.020**	0.020**	0.024***
	(0.012)	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)
Long Hours			-0.127***			-0.094***
			(0.009)			(0.007)
Constant	0.398***	0.578***	0.556***	0.463***	0.511***	0.518***
	(0.013)	(0.017)	(0.016)	(0.011)	(0.013)	(0.013)
Observations	23298	23298	23298	42480	42480	42480

Table 5: Female Share by Parental Status:
NLSY Salaried FTFY Workers 25+ 1981-2016

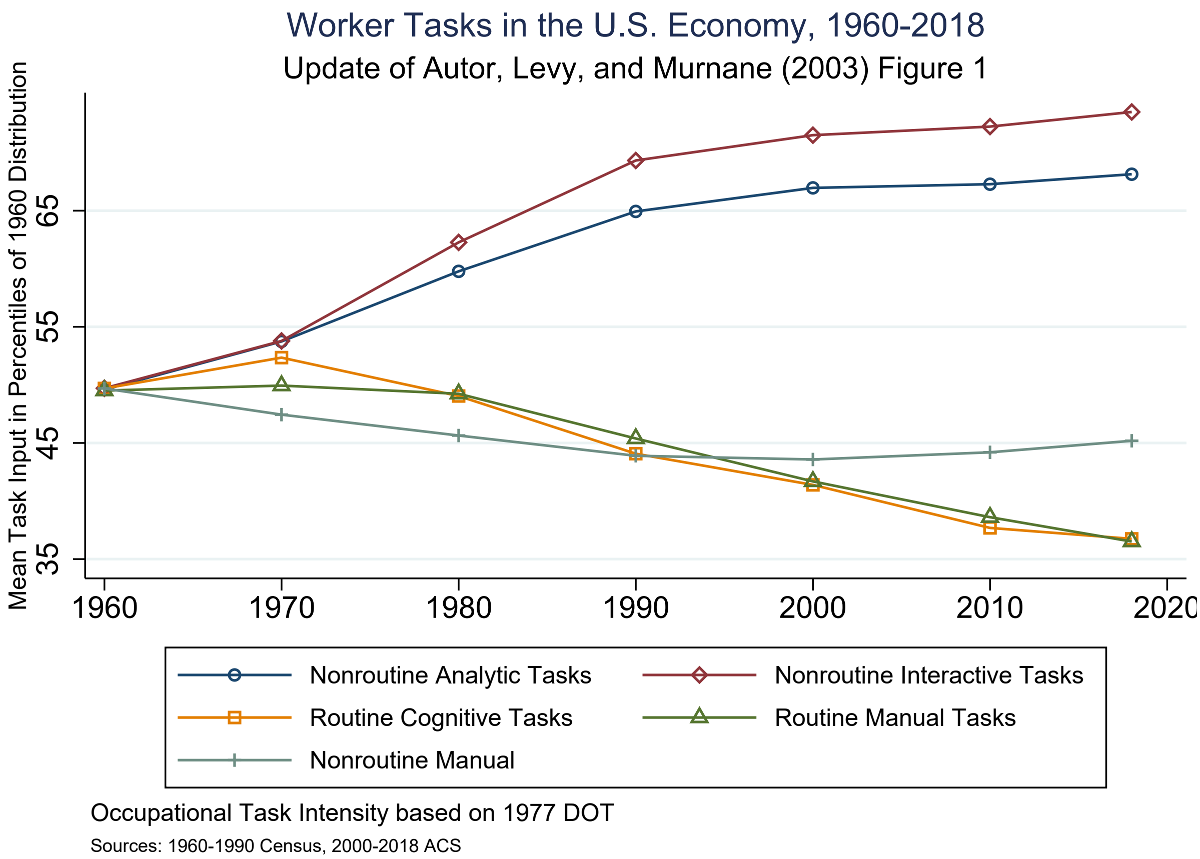
Outcome is $D_{(Female)}$	(1)	(2)	(3)	(4)	(5)	(6)
Abstract	-0.045***	-0.011	-0.015*	-0.011	0.024*	0.009
	(0.006)	(0.012)	(0.008)	(0.012)	(0.012)	(0.014)
Routine	0.016***	0.025***	0.025***	0.025***	0.003	0.012
	(0.005)	(0.010)	(0.007)	(0.010)	(0.010)	(0.011)
Manual	-0.151***	-0.123***	-0.128***	-0.123***	-0.111***	-0.116***
	(0.005)	(0.010)	(0.007)	(0.010)	(0.010)	(0.010)
After 1st Child			0.012	-0.003	-0.003	-0.001
			(0.013)	(0.013)	(0.013)	(0.013)
Parent		0.019		0.021	0.023	0.022
		(0.016)		(0.018)	(0.018)	(0.018)
After × Abstract			-0.050***	-0.042***	-0.034***	-0.008
			(0.010)	(0.011)	(0.011)	(0.012)
After × Routine			-0.017**	-0.015*	-0.015	-0.032***
			(0.008)	(0.009)	(0.009)	(0.010)
After × Manual			-0.038***	-0.029***	-0.025***	-0.015
			(0.009)	(0.009)	(0.009)	(0.010)
Parent × Abstract		-0.046***		-0.012	-0.016	-0.017
		(0.013)		(0.015)	(0.015)	(0.018)
Parent × Routine		-0.013		-0.002	-0.002	0.000
		(0.011)		(0.013)	(0.013)	(0.014)
Parent × Manual		-0.037***		-0.014	-0.016	-0.016
		(0.011)		(0.013)	(0.013)	(0.014)
Overwork					-0.069***	-0.041***
					(0.006)	(0.013)
After × Overwork						-0.053***
						(0.013)
Parent × Overwork						0.005
						(0.018)
Constant	0.427***	0.413***	0.420***	0.413***	0.413***	0.413***
	(0.007)	(0.014)	(0.011)	(0.014)	(0.014)	(0.014)
Observations	65778	65778	65778	65778	65778	65778

Table 6: Returns to Skills and Job Task Requirements:
NLSY Salaried White Male FTFY Workers 25+ 1981-2016

Outcome is ln(*wage*)	(1) Pooled	(2) < 2000	(3) Change in 2000+
Abstract	0.089***	0.068***	0.041***
	(0.009)	(0.009)	(0.013)
Routine	0.050***	0.069***	-0.029***
	(0.006)	(0.007)	(0.010)
Manual	0.021***	0.005	0.036***
	(0.006)	(0.007)	(0.009)
Cognitive	0.104***	0.059***	0.087***
	(0.011)	(0.013)	(0.011)
Non-cognitive	0.041***	0.028***	0.034***
	(0.008)	(0.009)	(0.010)
Social	0.014*	0.011	0.009
	(0.008)	(0.009)	(0.010)
Cognitive × Abstract	0.037***	-0.003	0.061***
	(0.009)	(0.009)	(0.012)
Cognitive × Routine	-0.038***	-0.020***	-0.014
	(0.007)	(0.008)	(0.011))
Cognitive × Manual	-0.013*	0.008	-0.044***
	(0.007)	(0.008)	(0.010)
Observations	27729	27729	27729
R-Square	0.302	0.326	0.326

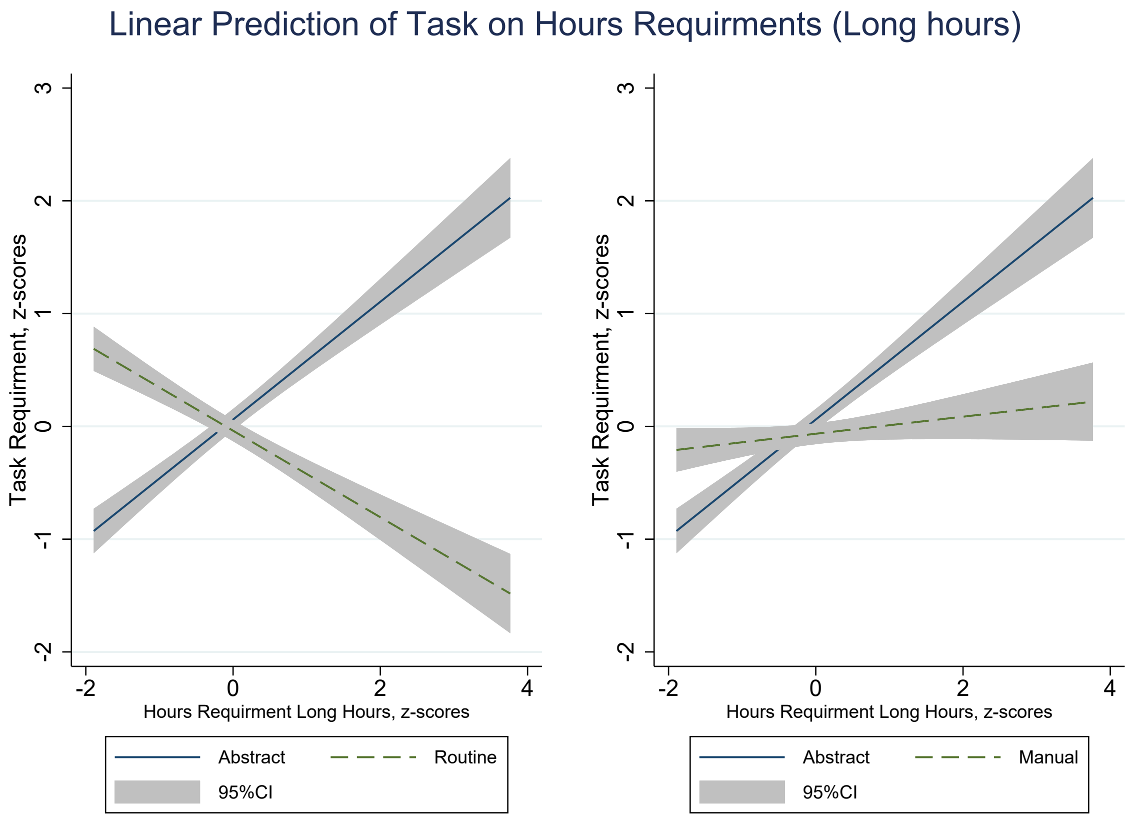
Sources: See notes to table 3. In addition to the variables reported in the table, additional controls include dummies for education categories, single years of age, region, urban, metro. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 1: Worker Tasks in the U.S. Economy, 1960 to 2018



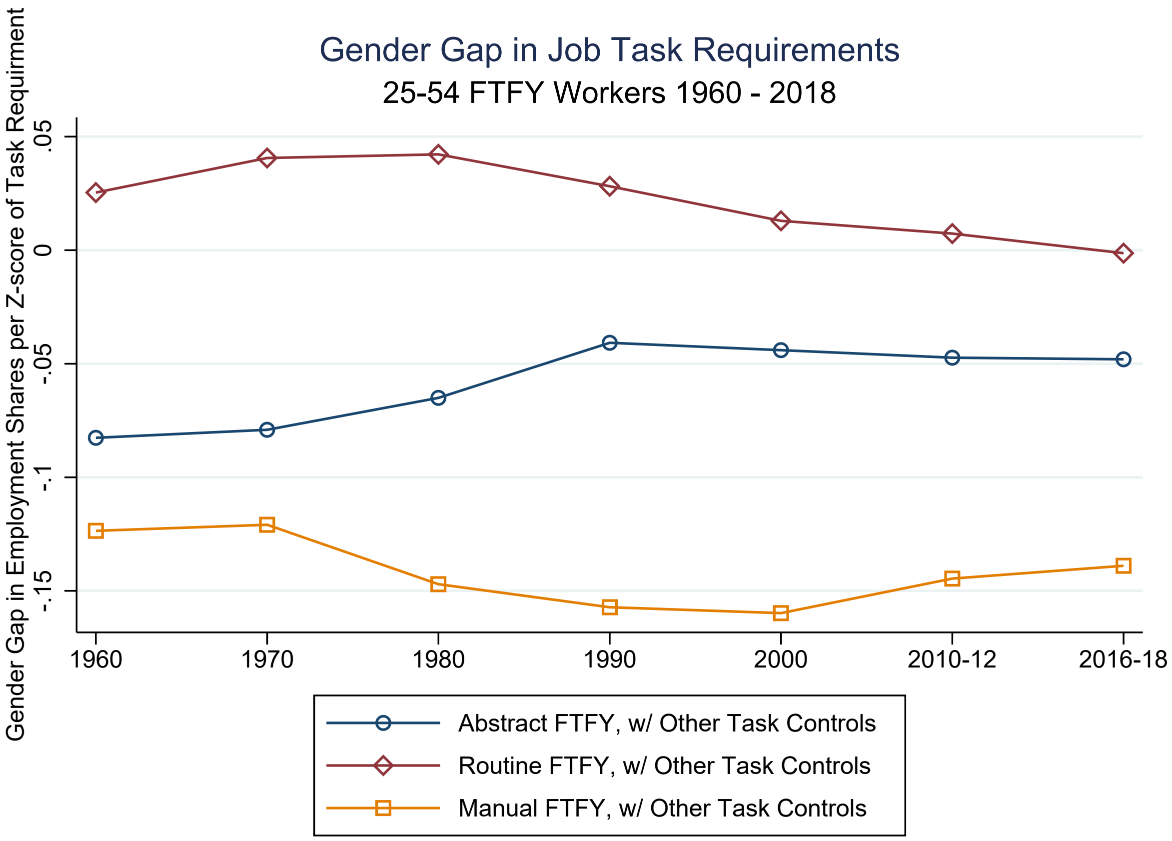
Source: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. We use harmonized 1990 occupation codes by IPUMS. Following Autor, Levy, and Murnane (2003) and Deming (2017), data are aggregated to industry-education-sex cells by year, and each cell is assigned a value corresponding to its percentile rank in the 1960 distribution of task input. The figure plots the employment-weighted mean of the percentile values in each year.

Figure 2: Task Requirements and Hours Requirements



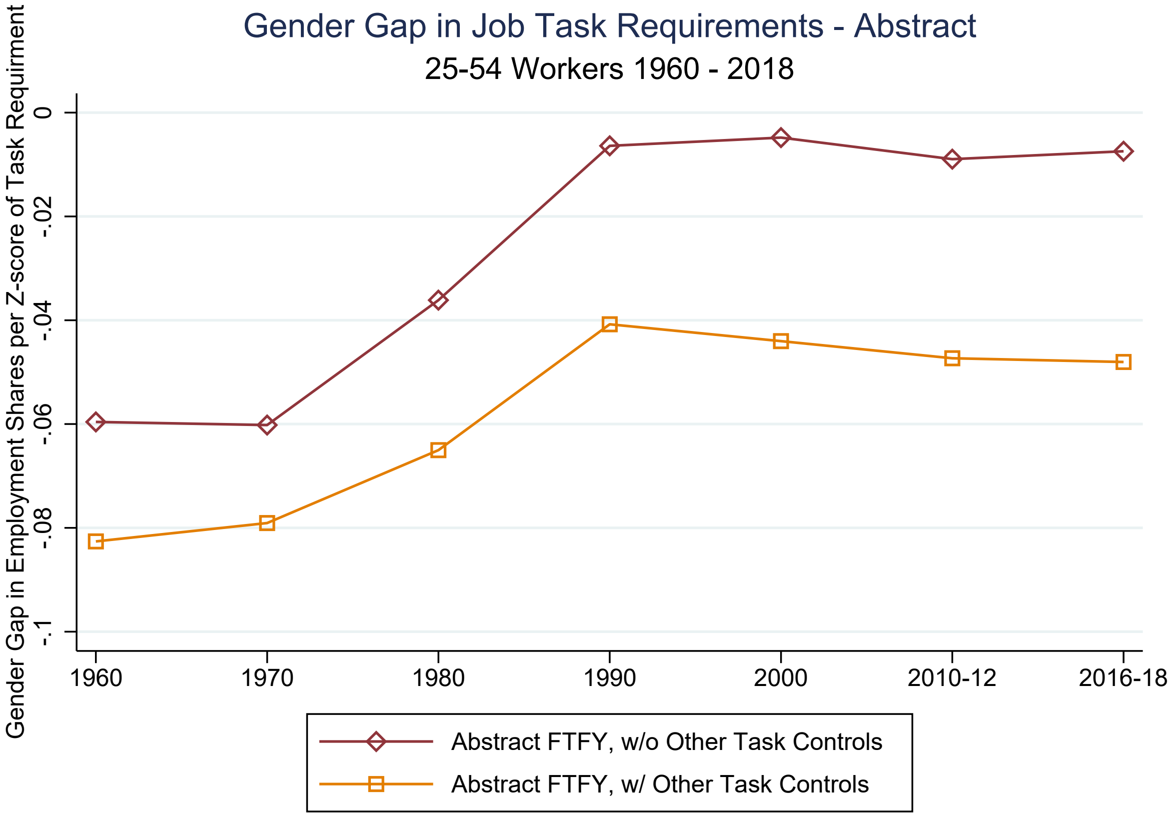
Source: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and Current Population Surveys 2003-2018. Task measures are from DOT 1977. We use the classifications used by Autor, Levy, and Murnane 2003 and define "Abstract" as the average of "GED Math" and "Direction, Control, and Planning (DCP)". "Routine" is the average of "Set Limits, Tolerance, or Standards (STS)" and "Finger Dexterity (FINGDEX)." "Manual" corresponds to "Eye, Hand, Foot Coordination (EYEHAND)". We use the 2003-2018 CPS to define "long hours" at the occupation level. We follow Cortes and Pan 2016 and define "long hours" as the share of men 18-65 years old who worked 50 or more usual hours per week at the main job. All measures are at the detailed 3-digit 1990 harmonized occupation level. There are 389 occupations overall although the number of unique occupations with non-missing data varies by year. The figures show the linear prediction plots with the associated 95-percent confidence intervals.

Figure 3: Gender Gap in Job Task Requirements



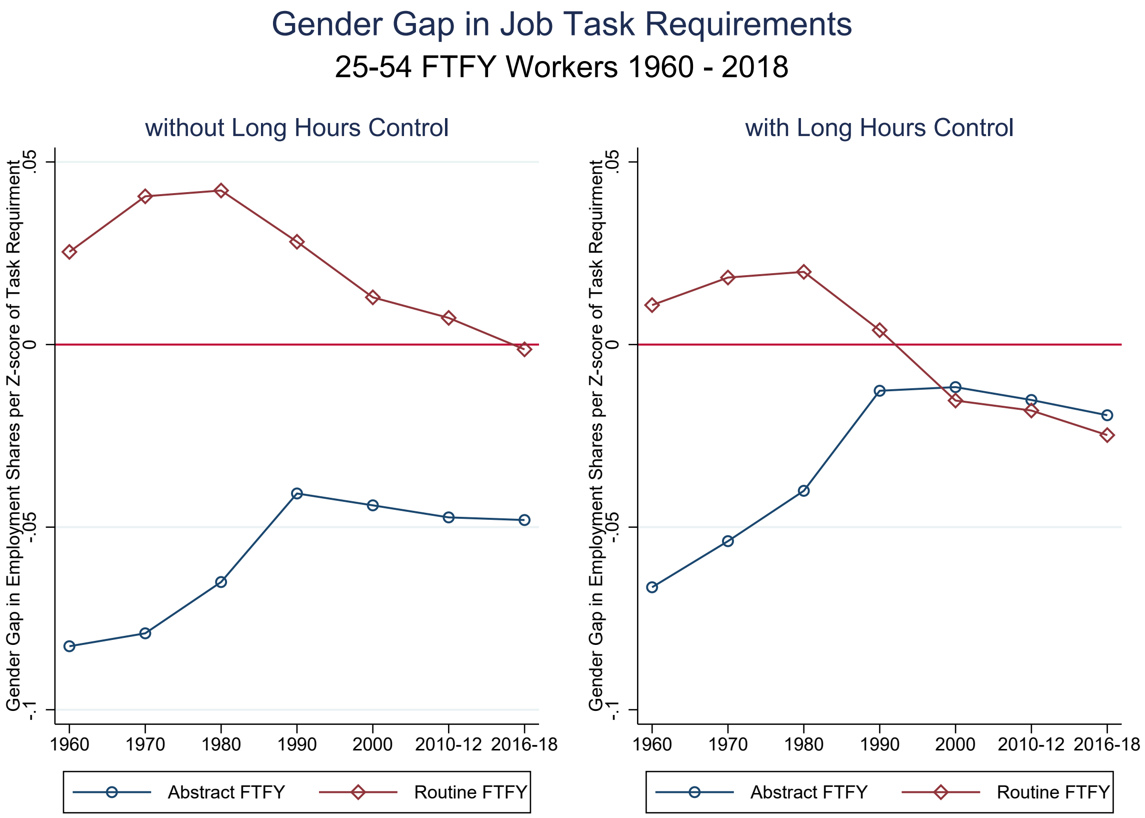
Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements. More specifically the figure plots γ_{kt} from the following regression: $\mathrm{FEMALE}_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where FEMALE_{ijt} is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, and region dummies.

Figure 4: Gender Gap in Job Task Requirements: Abstract



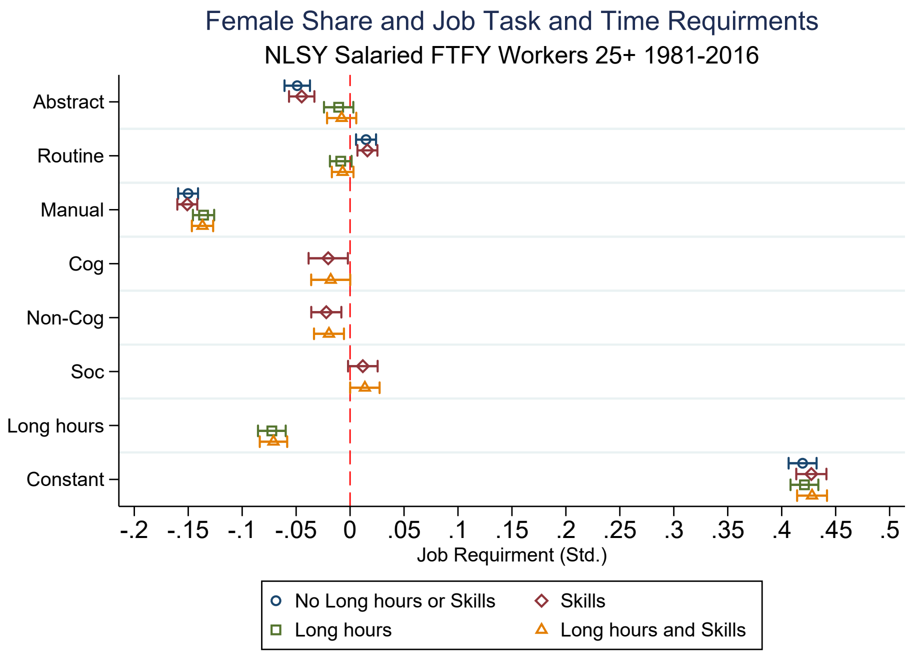
Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements. More specifically the figure plots γ_{kt} from the following regression: $FEMALE_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where $FEMALE_{ijt}$ is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, and region dummies.

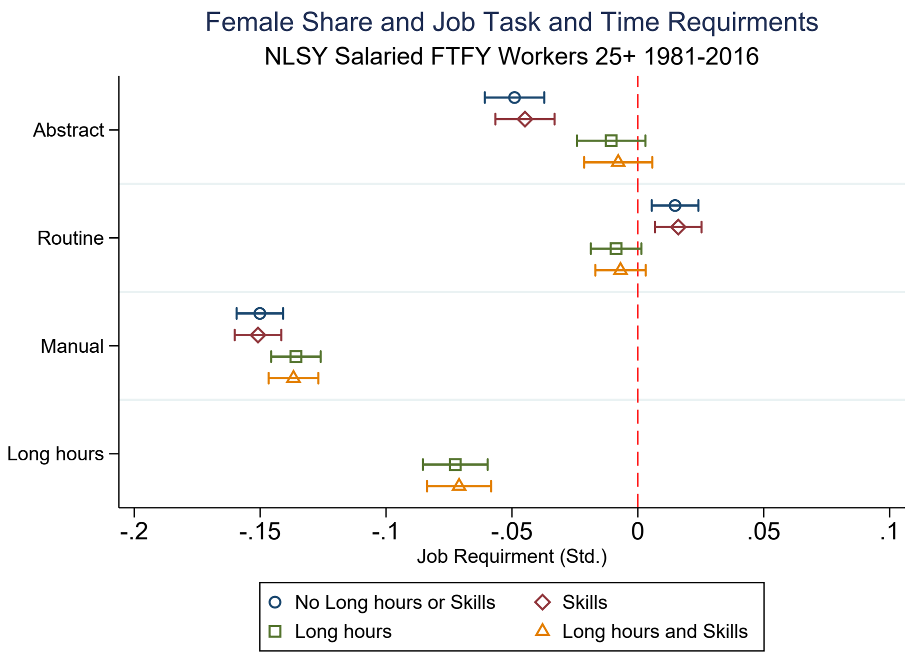
Figure 5: Gender Gap in Abstract Tasks with Hours Control



Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements, and with and without controlling for "long hours." More specifically the right panel plots γ_{kt} from the following regression: $FEMALE_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{ht} HOURS_{ijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where $HOURS_{ijt}$ is the "long hours" requirement in occupation j, expressed in z-scores.

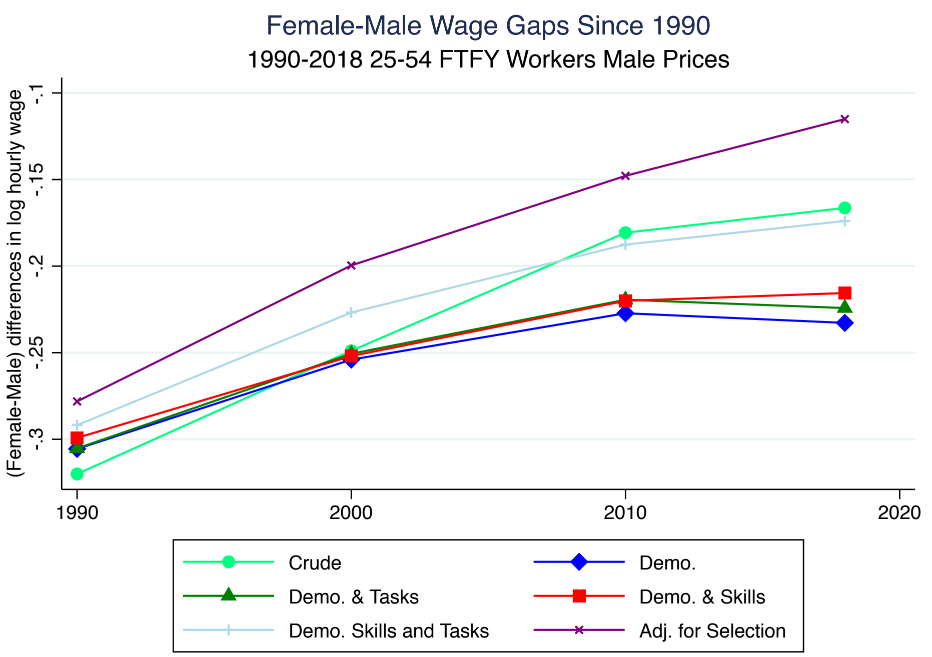
Figure 6: Gender Gap in Sorting into Tasks: w/ and w/o Hours Controls:NLSY





Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1979 and 1997 waves of the National Longitudinal Surveys of Youth. DOT 1977 task measures by occupation are merged with NLSY79 and NLSY97. We keep men and women who are 25+ years old observed employed over 1981-2016 period. We again select full-time, full-year workers who are not self-employed. To be included in the regression, individuals also have to have non-missing cognitive (AFQT), non-cognitive, and social skill measures. We weight by sampling weights and cluster standard errors at the person level. The figure plots γ_{kt} from the following regression : $\text{FEMALE}_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{ht}\text{HOURS}_{ijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$. The figure plots coefficients with and without including "long hours" controls and controls for skill measures.

Figure 7: Gender Wage Gap and Changes: 1990-2018 ACS-Census



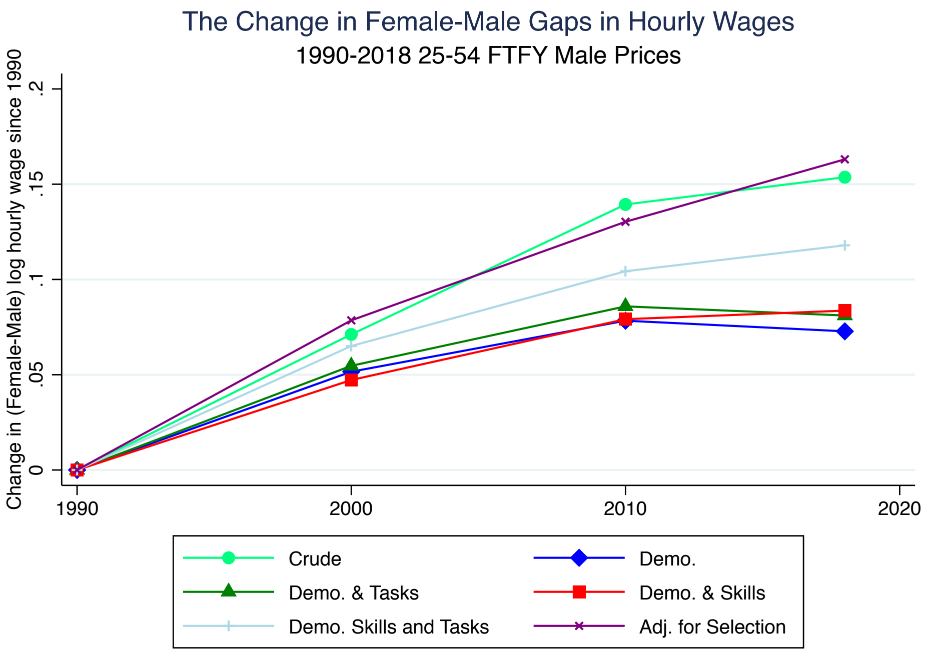
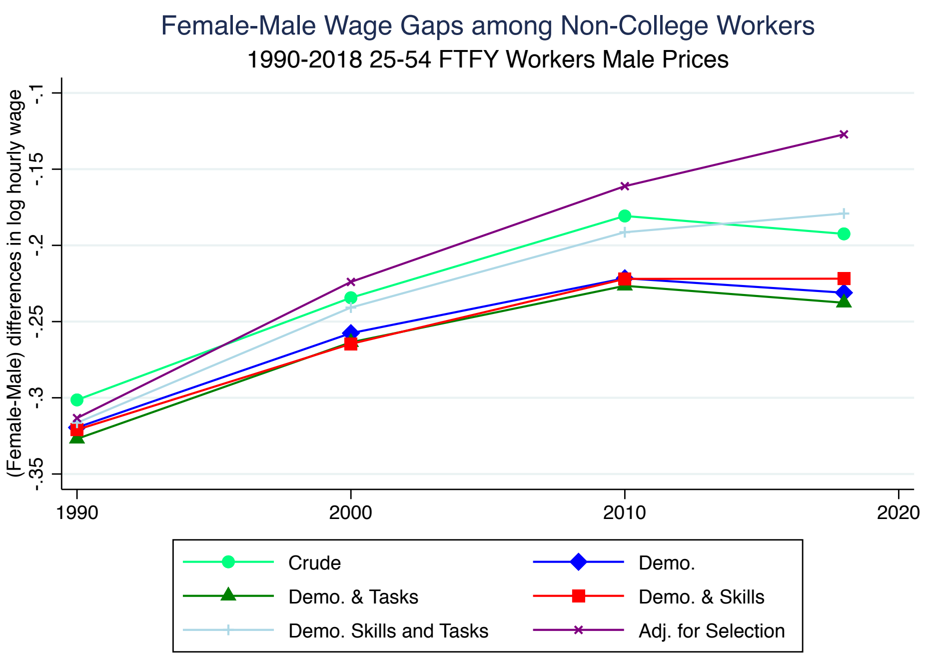


Figure 8: Gender Wage Gap and Changes Among Less than College: 1990-2018



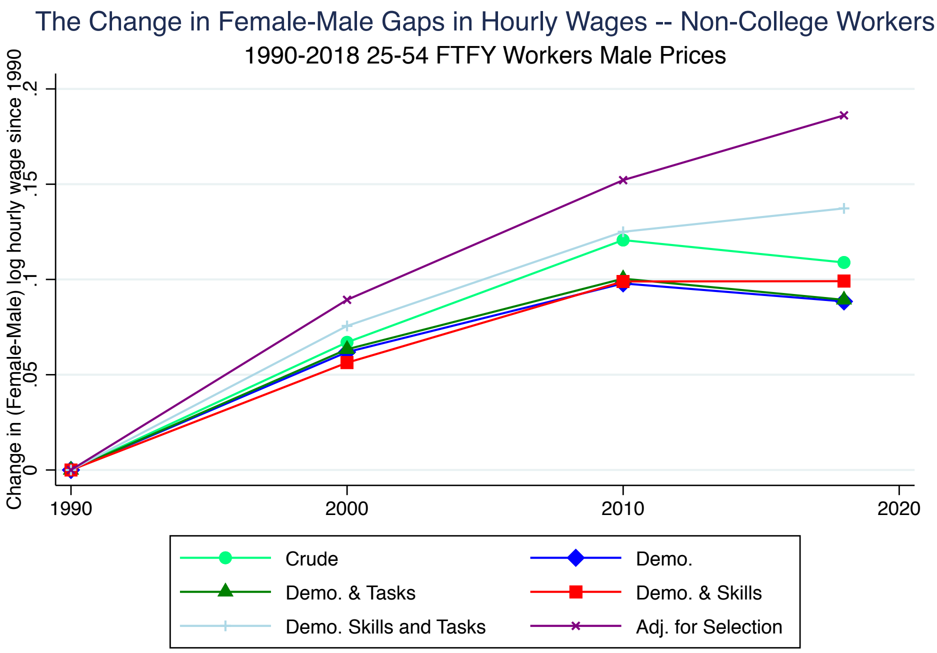
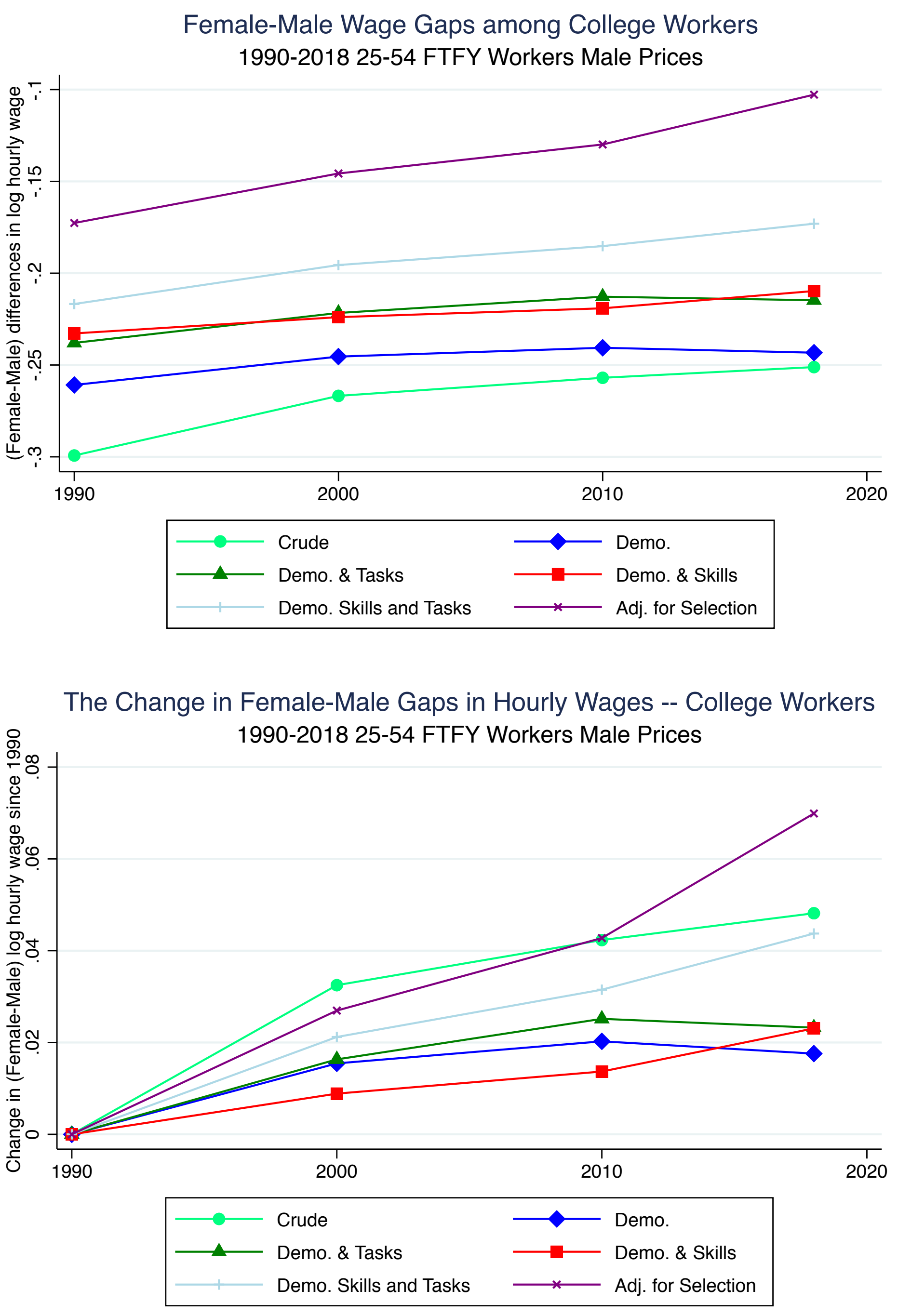


Figure 9: Gender Wage Gap and Changes Among College+: 1990-2018



Appendix

A.1 Data Appendix

A.1.1 Task Measures: DOT

We use the following 1977 DOT job characteristics measures used by Autor, Levy, and Murnane 2003 and Autor and Dorn 2013:

- Non-routine Analytical (GED Math): The degree to which the task demands analytical flexibility, creativity, reasoning, and generalized problem-solving
- Non-routine Interactive (Direction, Control, Planning - DCP): The degree to which the task demand complex interpersonal communications such as persuading, selling, and managing others
- Non-routine Manual (Eye Hand Foot Coordination) The degree to which the task demands eye, hand, and foot coordination
- Routine Analytical (Set Limits, Tolerances, or Standards - STS): The degree to which the task requires the precise attainment of set standards
- Routine Manual (Finger Dexterity): The degree to which the task requires repetitive manual tasks

Higher levels of GED-Math are associated with higher quantitative abstract tasks. Occupations with high measures of GED Math include various medical professionals, various engineers, accountants, and software developers. Higher levels of DCP are associated with higher levels of abstract thinking associated with management, organizational, and teaching tasks. Occupations with high measures of DCP include various managers, high school teachers, college professors and judges. The literature has equated these task requirements with non-routine analytical and problem solving requirements. We take a simple average of these two measures and label it "Abstract."

STS measures the adaptability to work in situations requiring setting of limits and measurements and serves as a proxy for routine cognitive tasks. Occupations with high measures of STS include meter readers, pilots, drafters, auto mechanics, and various manufacturing occupations. FINGDEX measures the ability to move fingers and manipulate small objects with fingers and serves as a proxy for repetitive routine manual tasks. Occupations with high measures of FINGDEX include secretaries, dental hygienists, bank tellers, machinists, textile sewing machine operators, dressmakers, and x-ray technology specialists. We take an average of these two measures and label it "Routine." Occupations which score high on EYEHAND demand eye, hand, and foot coordination, and include athletes, police and fire fighters, drivers (taxi, bus, truck), skilled construction (e.g, electricians, painters, carpenters) and landscapers/-groundskeepers. In our analysis we label this measure "Manual."

A.1.2 Task and Time Measures: O*NET

We downloaded the following O*NET skill and job characteristics measures used by Acemoglu and Autor 2011 and Denning et al. 2019 from the 1998 O*NET (4.0):

- Abstract: Average of 3 measures: "Interpreting the Meaning of Information for Others," "Thinking Creatively," and "Analyzing Data or Information"
- Routine: Average of 5 measures: "Controlling Machines and Processes," "Spend Time Making Repetitive Motions," "Pace Determined by Speed of Equipment," "Importance of Being Exact or Accurate," and ""Importance of Repeating Same Tasks"
- Manual: Average of 4 measures: "Spend Time Using Your Hands to Handle, Control, or Feel Objects, Tools, or Controls," "Manual Dexterity," "Operating Vehicles, Mechanized Devices, or Equipment," and "Spatial Orientation"

We downloaded the following characteristics related to time requirements following Goldin 2014:

- Time pressure: How often does this job require the worker to meet strict deadlines?
- Contact with others: How much does this job require the worker to be in contact with others (face-to-face, by telephone, or otherwise) in order to perform it?
- Establishing and maintaining interpersonal relationships: Developing constructive and cooperative working relationships with others, and maintaining them over time.
- Structured versus unstructured work: To what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals?
- Freedom to make decisions: How much decision making freedom, without supervision, does the job offer.

A.1.3 NLSY Skill Measures

- Cognitive Skills: Age-adjusted AFQT scores made comparable across 1979 and 1997 cohorts following Altonji, Bharadwaj, and Lange 2012 and Deming 2017

- Non-cognitive skills: For the NLSY79, our measure of non-cognitive skills uses the individual's survey responses to questions measuring the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale. The Rotter scale measures the degree of control individuals feel they possess over the life. The Rosenberg scale measures perceptions of self-worth. Higher values of both are interpreted as high levels of non-cognitive skills. For our measure of non-cognitive skills in 1979, we convert the two measures to the same scale, average them together and then convert into z-score units. The Rotter and Rosenberg questions were not asked in the NLSY97. Instead, respondents answered 8 personality questions derived from Goldberg's Big Five Personality Assessment: (i) extraversion, (ii) agreeableness, (iii) openness (iv) conscientiousness, and (v) neuroticism. Deming 2017 utilizes the conscientious related question from the Big-5 – "How much do you feel that conscientious describes you as a person?" – to approximate respondents' non cognitive skill. We use his definition - which he converted to z-score units - for our measure of non-cognitive skills in the NLSY97.

- Social skills: We follow Deming 2017 to develop measures of social skills. Specifically, for the NLSY79, we use self-reported measures of sociability in childhood and sociability in adulthood. Individuals were asked to assess their current sociability (extremely shy, somewhat shy, somewhat outgoing, or extremely outgoing) and to retrospectively report their sociability when they were age 6. For the NLSY97, we proxy for social skills using the two questions that were asked to capture the extroversion factor from the commonly used Big 5 personality inventory. For both waves, we normalize the two questions so they have the same scale and then average them together (separately for each survey). We then convert the measures into z-score units.

Table A.1: Correlation of Task and Time Requirements

Time Requirements	Abstract	Routine
Long Hours: 1960 Census/ACS	0.228*	-0.392*
Long Hours: 1970 Census/ACS	0.327*	-0.356*
Long Hours: 1980 Census/ACS	0.457*	-0.380*
Long Hours: 1990 Census/ACS	0.506*	-0.335*
Long Hours: 2000 Census/ACS	0.464*	-0.278*
Long Hours: 2010 Census/ACS	0.444*	-0.282*
Long Hours: 2018 Census/ACS	0.352*	-0.217*
Long Hours: 2003-2019 CPS	0.506*	-0.397*

Table A.2: Female Share and Job Task and Time Requirements:
NLSY Salaried FTFY Workers 25+ 1981-2016, W/O Controlling for Education

Outcome is Female Dummy	(1)	(2)	(3)	(4)
Abstract	-0.034***	-0.034***	0.006	0.005
	(0.006)	(0.006)	(0.007)	(0.007)
Routine	0.014***	0.014***	-0.011**	-0.010*
	(0.005)	(0.005)	(0.005)	(0.005)
Manual	-0.151***	-0.150***	-0.136***	-0.135***
	(0.005)	(0.005)	(0.005)	(0.005)
Long hours			-0.076***	-0.076***
			(0.007)	(0.007)
Constant	0.418***	0.418***	0.420***	0.419***
	(0.007)	(0.007)	(0.007)	(0.007)
Worker Skills		X		X
Observations	65776	65776	65776	65776

Table A.3: Female Share and Job Task and Time Requirements:
NLSY Salaried FTFY Workers 25+ 1981-2016, Overwork Measure From Tertiary Educated

Outcome is Female Dummy	(1)	(2)	(3)	(4)
Abstract	-0.049***	-0.045***	-0.006	-0.003
	(0.006)	(0.006)	(0.007)	(0.007)
Routine	0.015***	0.016***	-0.014***	-0.012**
	(0.005)	(0.005)	(0.005)	(0.005)
Manual	-0.150***	-0.151***	-0.131***	-0.132***
	(0.005)	(0.005)	(0.005)	(0.005)
Long hours (Tertiary)			-0.091***	-0.089***
			(0.006)	(0.006)
Constant	0.419***	0.427***	0.421***	0.429***
	(0.007)	(0.007)	(0.007)	(0.007)
Worker Skills		X		X
Observations	65776	65776	65776	65776

Table A.4: Returns to Skills and Job Task Requirements:
NLSY Salaried White Male FTFY Workers 25+ 1981-2016

Outcome is ln(*wage*)	(1) Pooled	(2) < 2000	(3) Change in 2000+
Abstract	0.054***	0.057***	-0.009**
	(0.005)	(0.004)	(0.003)
Routine	0.059***	0.049***	0.014***
	(0.003)	(0.004)	(0.003)
Manual	0.029***	0.012	0.030***
	(0.008)	(0.007)	(0.003)
Cognitive	0.099***	0.061***	0.079***
	(0.009)	(0.008)	(0.006)
Non-cognitive	0.081***	0.058***	0.036***
	(0.003)	(0.003)	(0.005)
Social	0.015	0.029***	-0.025***
	(0.011)	(0.007)	(0.006)
Cognitive × Abstract	0.060***	0.049***	0.013***
	(0.004)	(0.003)	(0.002)
Cognitive × Routine	-0.048***	-0.044***	-0.005
	(0.005)	(0.005)	(0.005)
Cognitive × Manual	-0.024***	0.014**	-0.065***
	(0.006)	(0.005)	(0.005)
Observations	3749391	3749391	3749391
R-square	0.287	0.293	0.293

Sources: See notes to table 3. In addition to the variables reported in the table, additional controls include dummies for education categories, single years of age, region, urban, metro. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Top 10 occupations of each categories

i. Low Abstract & Low Overwork	ii. Low Abstract & High Overwork
Machine operators, n.e.c.	Truck, delivery, and tractor drivers
Secretaries	Salespersons, n.e.c.
General office clerks	Primary school teachers
Laborers outside construction	Police, detectives, and investigators
Janitors	Farm workers
Assemblers of electrical equipment	Fire fighting and prevention,
Customer service reps	Heavy equipment and mechanics
Nursing aides	Real estate sales occupations
Mechanics and repairers, n.e.c.	Operating engineers-construction
Production checkers and inspectors	Hairdressers and cosmetologists
iii. High Abstract & Low Overwork	**iv. High Abstract & High Overwork**
Bookkeepers and accounting clerks	Managers and administrators, n.e.c.
Registered nurses	Production supervisors or foremen
Computer systems analysts	Supervisors and proprietors of sales
Office supervisors	Accountants and auditors
Computer software developers	Marketing managers and specialists
Electrical engineer	Financial managers
Social workers	Supervisors of construction work
Material production clerks	Personnel, and HR specialists
n.e.c engineers	Other financial specialists
Electrical and electronic technicians	Insurance sales occupations

Notes:

Figure A1: Task Requirements and Hours Requirements: Goldin's Measure

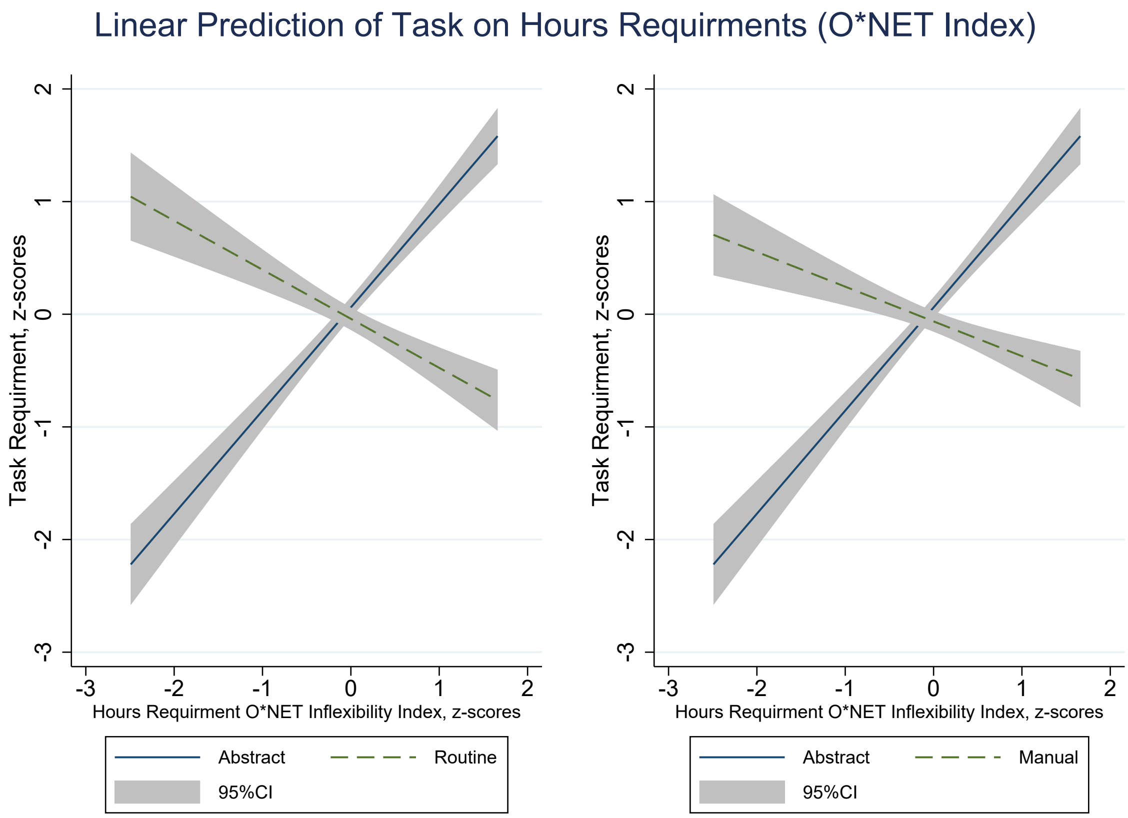


Figure A2: Task Requirements and Hours Requirements: Denning's Measure

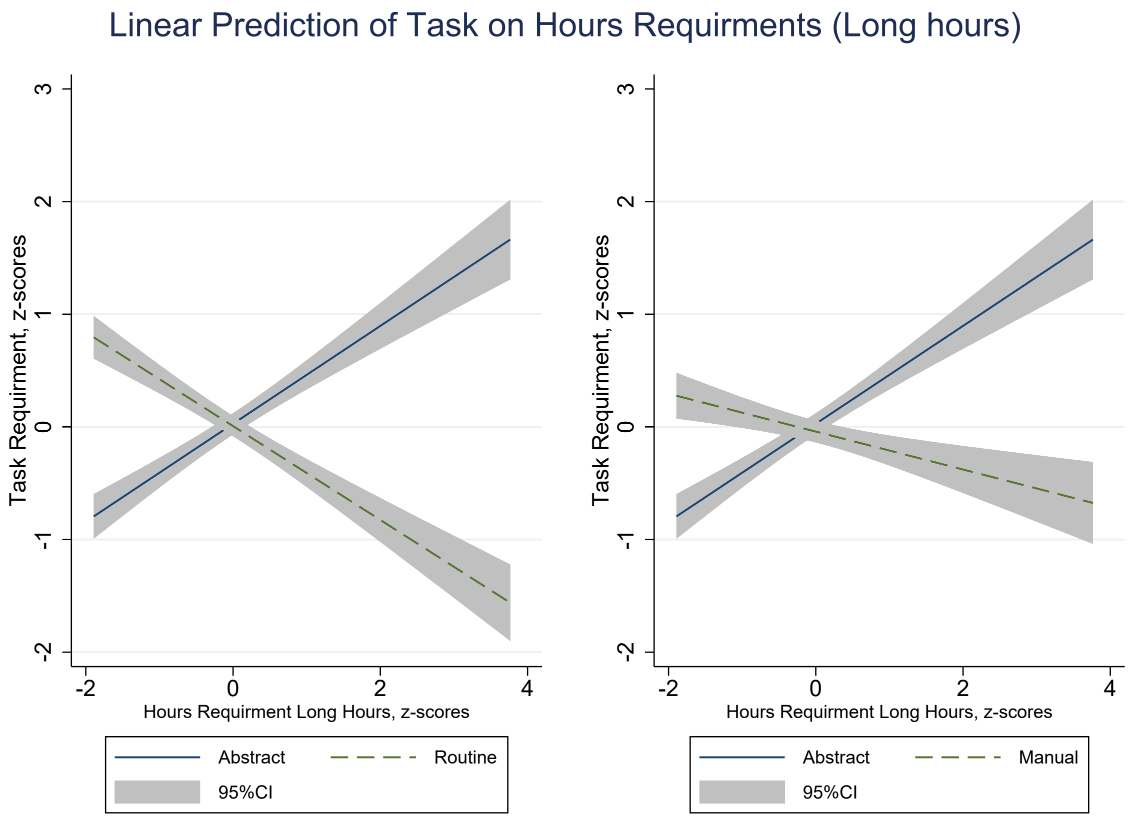
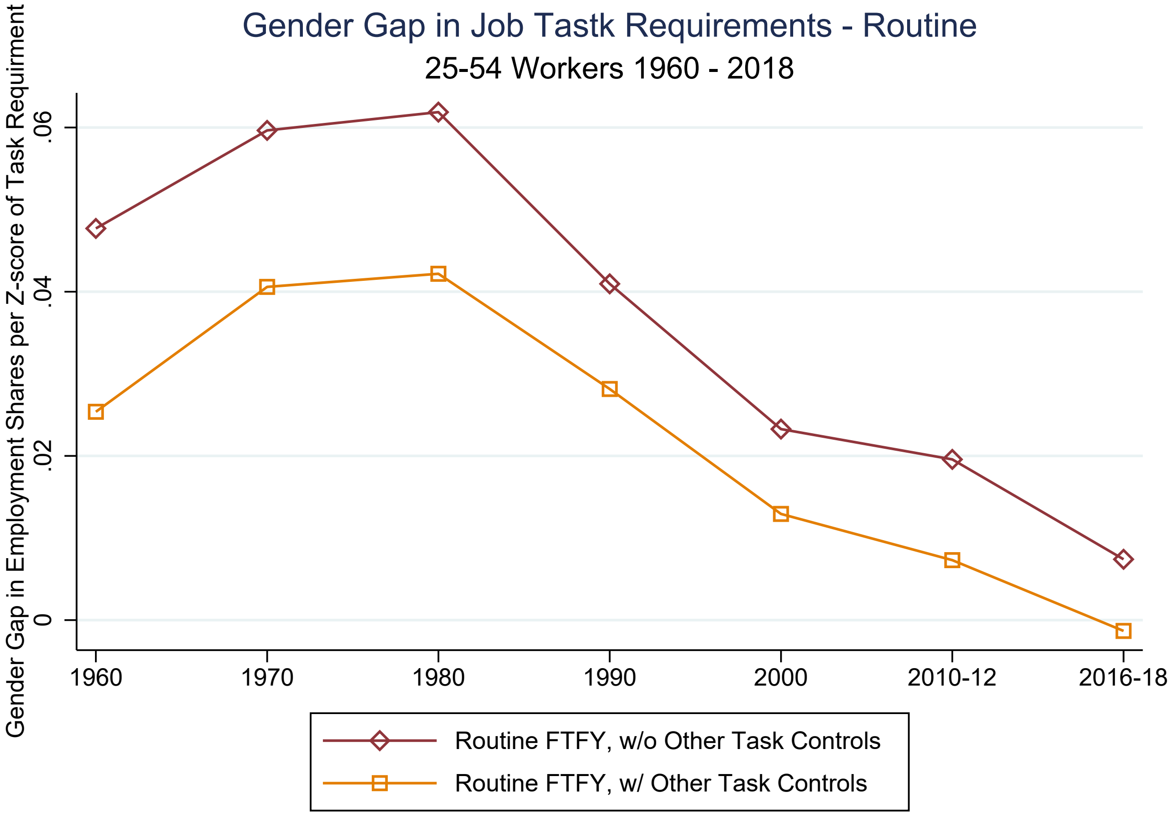
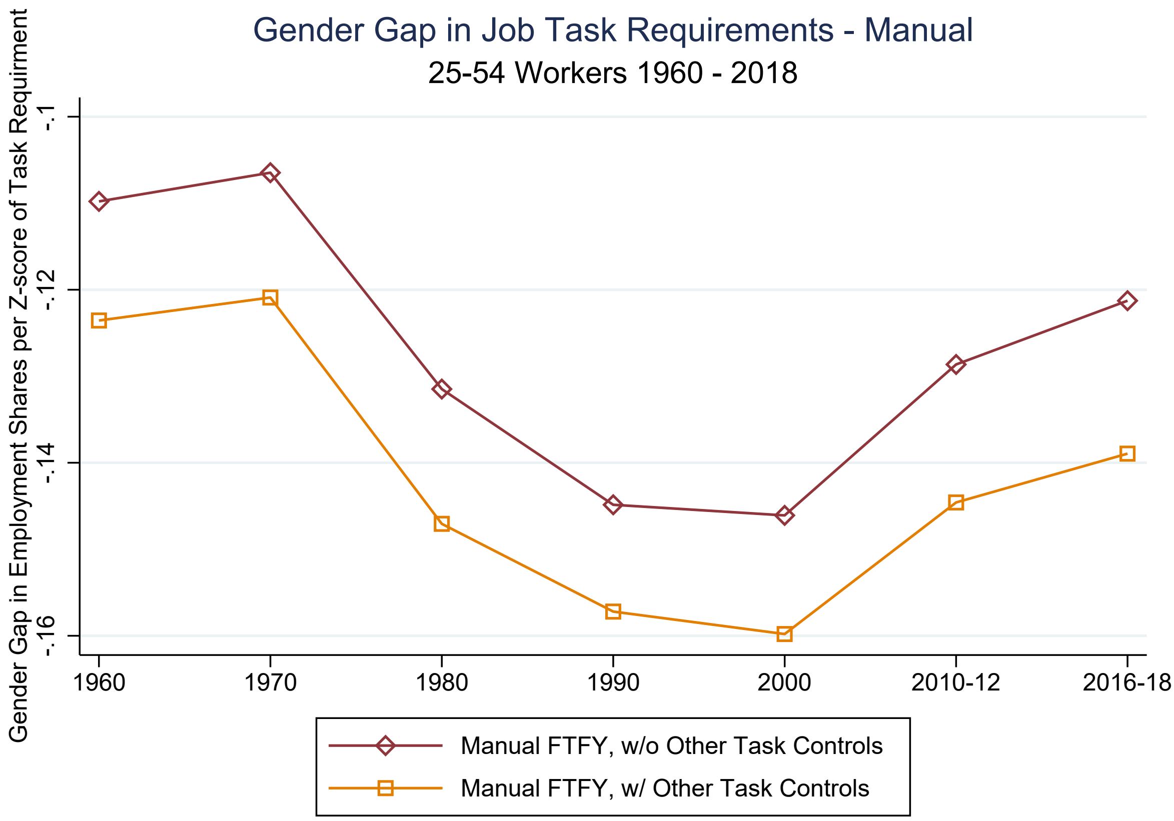


Figure A3: Gender Gap in Job Task Requirements: Routine



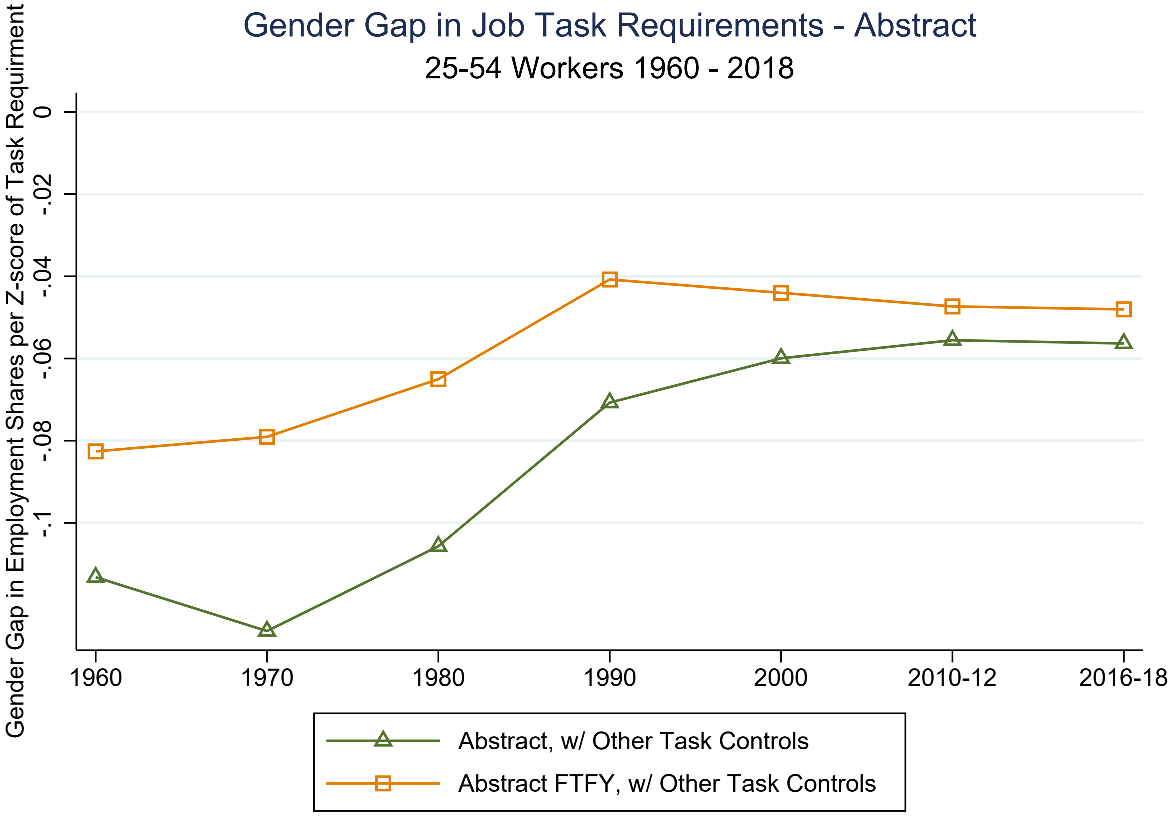
Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements. More specifically the figure plots γ_{kt} from the following regression: $FEMALE_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where $FEMALE_{ijt}$ is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, and region dummies.

Figure A4: Gender Gap in Job Task Requirements: Manual



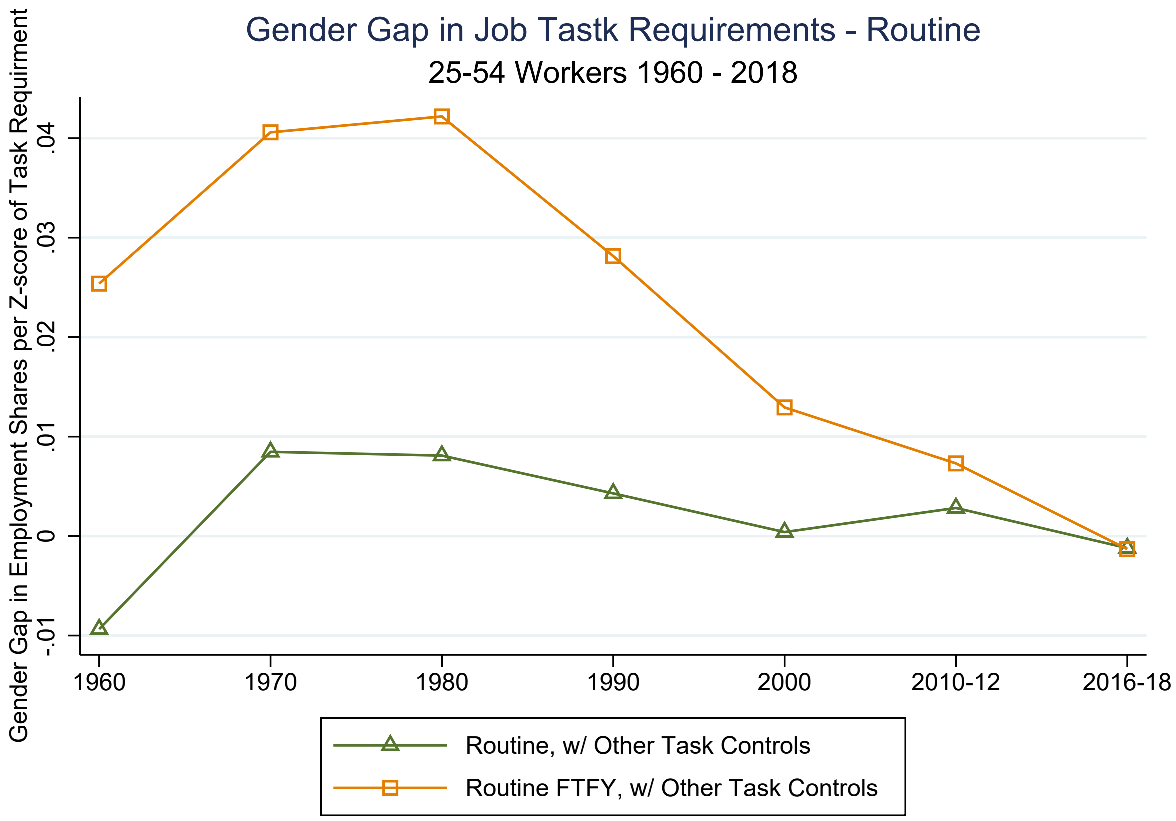
Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements. More specifically the figure plots γ_{kt} from the following regression: $FEMALE_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where $FEMALE_{ijt}$ is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, and region dummies.

Figure A5: Gender Gap in Job Task Requirements: Abstract



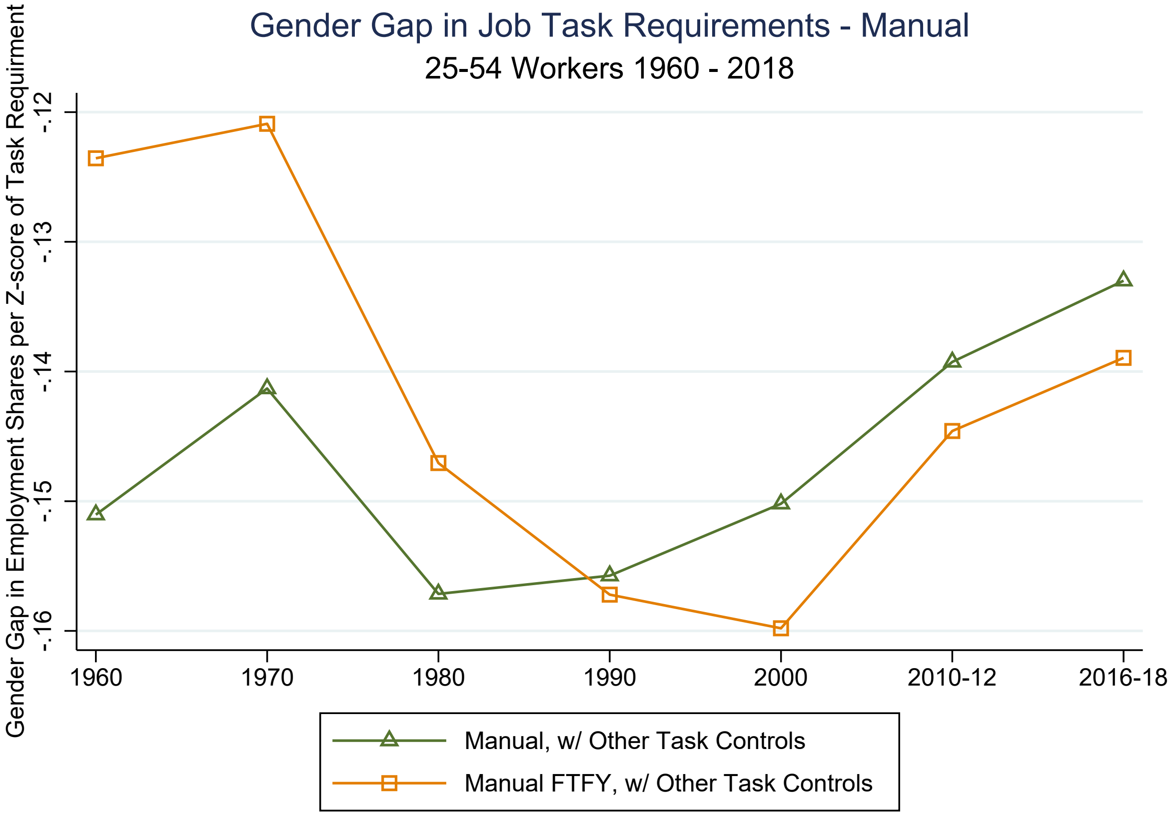
Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements. More specifically the figure plots γ_{kt} from the following regression: $\text{FEMALE}_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where FEMALE_{ijt} is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, and region dummies.

Figure A6: Gender Gap in Job Task Requirements: Routine



Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements. More specifically the figure plots γ_{kt} from the following regression: $\mathrm{FEMALE}_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where FEMALE_{ijt} is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, and region dummies.

Figure A7: Gender Gap in Job Task Requirements: Manual



Sources: U.S. DOL 1977 Dictionary of Occupation Titles (DOT) and 1960-2000 Census, 2010-2018 American Community Surveys. DOT 1977 task measures by occupation are merged with IPUMS 1960-2000 Censuses and the 2010-2018 American Community Survey samples. The sample includes men and women aged 25-54 who are wage and salary workers. "FTFY" refers to workers who worked 50 or more weeks the previous year and worked >=40 usual hours per week. The figure shows the gender gap in sorting into occupations with various task requirements. More specifically the figure plots γ_{kt} from the following regression: $FEMALE_{ijt} = \gamma_t + \sum_{k=1}^{3} \gamma_{kt} Z_{kijt} + \gamma_{xt} X_{it} + \varepsilon_{ijt}$, where $FEMALE_{ijt}$ is an indicator of whether individual i in job j in period t is female. Z_{kijt} is job task requirement k in occupation j observed for person i in period t. X_{it} are additional controls such as education dummies, age dummies, and region dummies.