Essays on Labor and Personnel Economics

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Public Policy

have examined a dissertation entitled

"Essays on Labor and Personnel Economics"

presented by Madeleine Gelblum

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ______________________

Typed name: Professor David Deming, Chair

Signature ______________________

Typed name: Professor Claudia Goldin

Signature ______________________

Typed name: Professor Lawrence Katz

Signature ______________________

Typed name: Professor Amanda Pallais

Date: May 27, 2020

Essays on Labor and Personnel Economics

A dissertation presented

by

Madeleine Gelblum

to

The Department of Public Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Public Policy

Harvard University

Cambridge, Massachusetts

May 2020

Dissertation Advisors:
Professor David Deming
Professor Lawrence Katz

Author:
Madeleine Gelblum

Essays on Labor and Personnel Economics

Abstract

This dissertation examines labor market inequality and factors that affect job satisfaction, productivity and skill requirements in firms. Chapter 1 provides evidence that gender differences in how individuals value activities performed at work, termed job tasks, can help explain gender differences in job choices. I conduct a hypothetical choice experiment to elicit workers' willingness to pay for a set of tasks that are more frequently performed by one gender than the other. I find significant gender differences in willingness to pay for three of the five tasks that I examine, and document that these differences can account for a substantial portion of occupational segregation in the U.S. labor market.

Chapter 2, which is co-authored work with John Horton, examines the relationship between wages and productivity and how firms make decisions about which workers to hire, using data from an online labor market. If workers are paid their marginal product, then a higher-wage worker should be a more productive worker who finishes a discrete project more quickly, leaving the total wage bill unchanged. We find that higher-wage workers do work fewer hours, as expected, but increase the total wage bill, suggesting that employers may systematically overvalue these individuals.

Chapter 3 explores how skill requirements in two cognitive occupations—marketing managers and financial analysts—change when employers adopt technology that facilitates data-driven decision-making, termed algorithmic technology. Using data from online job postings, I find that the mention of algorithmic technology is positively associated with complementary technical skills but negatively related to many frequently-listed non-routine cognitive skills in both occupations. In addition, algorithmic technology is positively correlated with wages across geographic area and year. These results suggest that data from online job postings can be valuable in understanding how technology use is related to skill requirements and wages.

Contents

Acknowledgments

I am grateful to my advisors, David Deming, Claudia Goldin, Lawrence Katz and Amanda Pallais, for their guidance. This research has received funding from the Harvard University Multidisciplinary Program in Inequality & Social Policy, the Harvard Kennedy School Women and Public Policy Program, and the Lynde and Harry Bradley Foundation, and has benefited from feedback from seminar participants at Harvard University and the London School of Economics.

To Ravi, who believed in me and supported me at every step along the way.

Chapter 1

Preferences for Job Tasks and Gender Gaps in the Labor Market

Chapter Abstract

Women and men work in markedly different jobs, leading to persistent occupational segregation by gender. This paper provides evidence that gender differences in how individuals value activities performed at work, termed job tasks, can help explain these sorting patterns. I conduct a hypothetical choice experiment to elicit workers' willingness to pay for a set of tasks that are more frequently performed by one gender than the other. The experimental scenarios ask participants to choose between two hypothetical jobs that differ in terms of pay and the amount of time spent on a gender-typical task, but are otherwise the same. I find significant gender differences in willingness to pay for three of the five tasks that I examine. Willingness to pay is significantly higher among participants who report spending more time on a task in their current job, suggesting that estimates are correlated with actual sorting behavior. I show that gender differences in preferences for the tasks that I investigate can account for a substantial portion of occupational segregation in the U.S. labor market.

1.1 Introduction

Women and men work in markedly different jobs, leading to persistent occupational segregation by gender (Blau *et al.* 2013). Figure 1.1 shows that the median woman is employed in an occupation in which 70 percent of workers are female, while the median man works in an occupation that is 71 percent male.[1] Occupational segregation contributes to the gender wage gap (Blau and Kahn 2017) and may reflect an inefficient allocation of workers to jobs. Indeed, women remain under-represented in many high-paying professional occupations despite having higher levels of education than men (Goldin *et al.* 2006).[2]

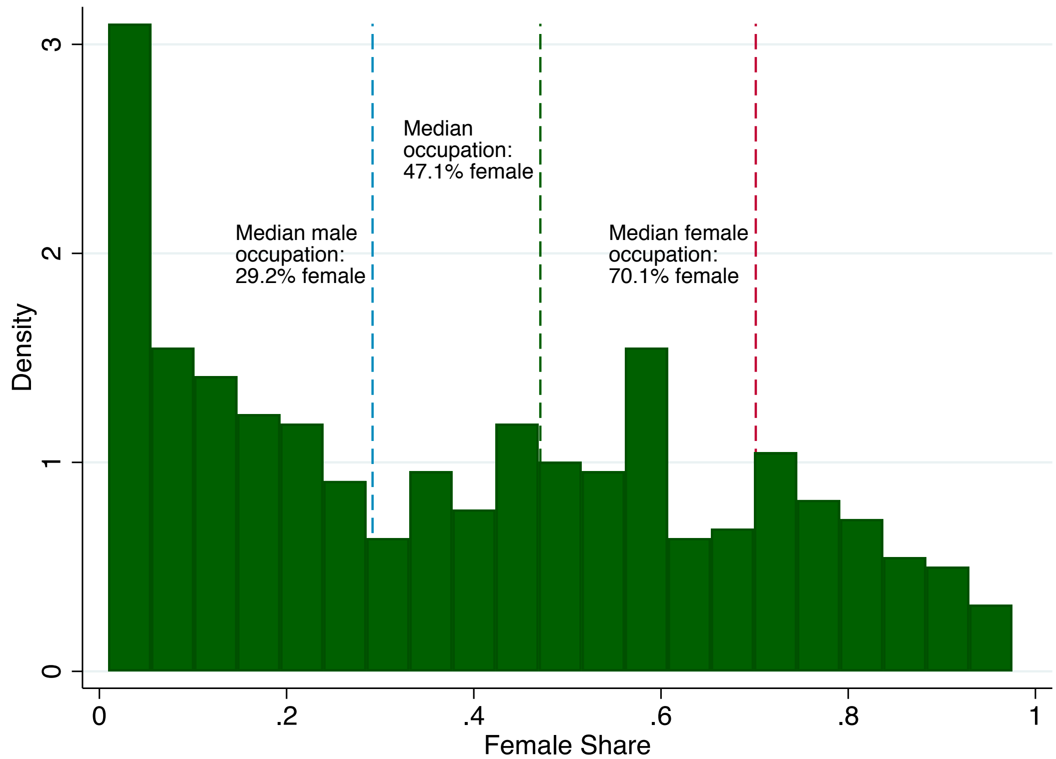
Gender differences in how individuals value activities performed at work, termed *job tasks*, may help explain these sorting patterns. Research documents that women and men work in jobs that involve different activities (e.g. Lordan and Pischke 2018; Cortes *et al.* 2018). In addition, measures of tasks can account for a large fraction of the variation in the share of workers in an occupation who are female, termed the *female share*.[3] Workers may have preferences over tasks as in Rosen (1986), such that they are willing to accept lower wages in jobs that involve activities they enjoy, and must be compensated extra to perform activities they dislike. If women and men have different preferences over tasks, these valuations may contribute to occupational segregation and other gender gaps.

This paper examines whether preferences over job tasks differ by gender. I conduct a hypothetical choice experiment embedded in a survey to elicit workers' willingness to pay (WTP) for a set of *gender-typical* tasks that women perform more frequently than men, or vice versa. While it is likely that preferences for a task will affect sorting, women and men

[1]Figure 1.1 displays the distribution of the share of workers in an occupation who are female for currently employed individuals aged 18 and older, using data from the 2012-2016 American Community Survey (ACS).

[2]Goldin (2014) shows that gender differences in occupation can account for approximately a third of the gender wage gap even among college graduates employed full-time.

[3]Table 1.2 displays the R^2 and adjusted R^2 statistics from regressions of the occupational female share in the ACS on a set of occupation-level explanatory variables. Column 1 includes the mean log hourly wage, mean log hours of work per week, and share of workers with a college degree. Column 2 includes all *generalized work activity* and *work style* variables from the O*NET, a database of occupational characteristics. The adjusted R^2 of 0.77 based on the task measures in Column 2 is nearly twice as large as the adjusted R^2 of 0.40 based on the variables in Column 1.



Notes: This figure shows the distribution of the share of workers in an occupation who are female using data from currently employed individuals aged 18 and older in the 478 occupations available in the 2012-2016 ACS. The dashed lines indicate the median occupational female share for (from left to right) male workers, all workers and female workers.

Figure 1.1: *Distribution of Occupational Female Share*

may differ in how frequently they perform a task for other reasons. In particular, women and men may have a comparative advantage in different activities (Baker and Cornelson 2018). Tasks may also be correlated with gender-based discrimination (Kuhn and Shen 2013), or with other amenities that women and men value differently, such as flexible work arrangements (Goldin 2014). Using observational data on wages and job choices to disentangle preferences for tasks from these other factors presents a challenge.[4]

By contrast, the hypothetical choice experiment allows me to identify task valuations by randomly assigning wage offers and specifying worker choice sets. Specifically, the experimental scenarios ask participants to choose between two hypothetical jobs that differ in terms of pay and the amount of time spent on a focal task, but are otherwise the same. Thus the difference in wage offers between the two jobs is known and is unrelated to worker skills or other attributes. In addition, participants are told that the two jobs are exactly the same in terms of schedule, co-workers, benefits and all other characteristics aside from the wage and time spent on the focal task. Therefore, participants are unlikely to view the two jobs as differing in terms of factors such as discrimination that are not directly connected to work activities.

The experiment elicits preferences for five conceptual categories of tasks that are performed in a broad range of jobs and are not tied to specific credentials. The task measures are based on variables from the O*NET, a database of occupational characteristics, that are correlated with the occupational female share in the American Community Survey (ACS). The variables that I incorporate into the experiment explain more than half of the variation in the female share, and capture nearly as much variation as the full set of task measures available in the O*NET. The selected measures include two female-typical tasks related to interpersonal activities - *helping and caring for others* and *working and communicating with others*. A third female-typical task, *documenting and recording information*, is important in

[4] A large literature documents the difficulty of estimating compensating differentials due to unobserved human capital and amenities (e.g. Brown 1980; Bell 2019). Similarly, measuring labor market discrimination remains an empirical challenge (Blau and Kahn 2017; Altonji and Blank 1999). In both cases, a key issue is that a worker's outside options are not observed in conventional labor market data.

many female-dominated jobs in health, education and social services. The male-typical task of *operating and repairing equipment* is consistent with the notion that men enjoy working with their hands or with machinery. A second male-typical task, *making decisions and solving problems*, is essential in many majority-male professional occupations.

I use participant choices from the experiment to estimate WTP for gender-typical tasks as a share of the wage, guided by a simple discrete choice model. I find that women are willing to pay significantly more than men for the female-typical tasks of *helping and caring for others* and *documenting and recording information*. In my preferred specification, women are willing to forgo 3.3 percentage points more than men as a share of their wage to work in a job in which they spend more time *helping and caring for others*, and are willing to give up 2.6 percentage points more than men to spend more time *documenting and recording information*. In addition, men have a significantly higher WTP than women - by 8 percentage points - for the male-typical task of *operating and repairing equipment*. I find no significant gender differences in WTP for the female-typical task of *working and communicating with others* or the male-typical task of *making decisions and solving problems*.

Furthermore, I find that for all tasks, WTP is significantly higher among those who report devoting a larger share of working hours to that activity in their current job, consistent with preferences for tasks affecting sorting decisions. In addition, WTP results are similar when I weight the experiment sample by gender, race, education and major occupation to match the nationally representative ACS. These findings suggest that the WTP estimates are predictive of real-world sorting decisions and may reflect preferences for a broader share of the labor market despite the fact that the experiment sample is not randomly selected.

In the final section of the paper, I examine the implications of the WTP results for gender gaps in the labor market. First, I document that observed gender differences in sorting on the job tasks that I investigate are substantial in both the experiment sample and the ACS. Sorting on these five tasks as measured by a segregation index can account for more than three quarters of occupational segregation by gender in the ACS.

Next, I calculate the gender differences in sorting that are predicted by my model given the preference estimates and wage differentials associated with the tasks. I find that these predicted differences in sorting can explain nearly 70 percent of the segregation attributable to the gender-typical tasks in the experiment sample, and more than 40 percent in the ACS. These results suggest that gender differences in preferences for the five tasks that I examine can account for approximately a third of occupational segregation in the U.S. labor market. This pattern of findings is robust to a variety of alternative specifications, including controlling for other job amenities and repeating the analysis in other data sources with better measures of human capital.

By contrast, the preference estimates appear to explain little of the gender wage gap, but results are not conclusive. This finding is largely driven by the fact that some of the tasks that I examine widen the wage gap while others narrow it, and these effects offset each other.[5] Therefore, estimates of the contribution of preferences for tasks to the gender wage gap are inherently sensitive to the choice of which activities to examine, as well as the magnitude of the wage differentials associated with tasks.

1.1.1 Related Literature

This paper contributes to a large literature on the determinants of occupational segregation and other gender gaps in the labor market. Several studies examine gender differences in preferences for job amenities and other workplace characteristics. Experimental evidence from laboratory and field settings suggests that women are less likely than men to choose competitive compensation schemes (Niederle and Vesterlund 2007; Flory *et al.* 2015), and more likely to select team-based pay (Kuhn and Villeval 2015).[6] Other research argues that women place greater value on flexible scheduling arrangements and working fewer hours due to household constraints (Goldin 2014; Wiswall and Zafar 2018; Mas and Pallais 2017;

[5]In particular, the female-typical task of *helping and caring for others* is associated with a wage penalty that increases the magnitude of the gender wage gap, while the female-typical task of *documenting and recording information* offers a wage premium and thus decreases the magnitude of the gap.

[6]See Bertrand (2011) and Croson and Gneezy (2009) for a review of the literature on gender differences in preferences related to attitudes and personality traits.

Goldin and Katz 2016; Wasserman 2019; Denning *et al.* 2019; Cortes and Pan 2019).[7]

Most relevant to this paper are a handful of studies in economics (Lordan and Pischke 2018; Cortes and Pan 2018; Fortin 2008; Grove *et al.* 2011) and a larger body of research in psychology (e.g. Su *et al.* 2009; Pinker 2008) suggesting that women have a greater preference than men for jobs that involve helping others or working with people rather than things. In particular, Lordan and Pischke (2018) summarize a large number of O*NET variables as three latent factors that they label *people*, *brains* and *brawn*, and show suggestive evidence that women have a relative preference for people compared with brawn jobs.[8]

The current project contributes to this prior literature by providing evidence that women have a higher WTP than men for helping tasks and a lower WTP for activities related to equipment and machinery, consistent with the notion of a gender difference in preferences for people versus things. In contrast to previous studies that have used observational methods or descriptive surveys, however, this paper offers the first set of experimental evidence that preferences for tasks as measured by WTP differ by gender.

This project also relates to research contending that women have a comparative advantage in performing interpersonal tasks relative to certain physical activities (Cortes *et al.* 2018; Baker and Cornelson 2018; Ngai and Petrongolo 2017; Borghans *et al.* 2014; Weinberg 2000; Beaudry and Lewis 2014; Bacolod and Blum 2010; Black and Spitz-Oener 2010; Welch 2000).[9] Much of this literature emphasizes the same stylized fact that motivates this project - that women and men work in jobs that involve different activities - but proposes the alternative interpretation that skills rather than preferences may account for these differences. Task-

[7]Studies also report evidence that workers prefer colleagues of the same gender (Pan 2015) and that women (men) are more likely to participate in a group activity that requires stereotypically female (male) topical knowledge (Coffman 2014), consistent with a role for norms and identity in explaining gender gaps (Akerlof and Kranton 2000).

[8]Lordan and Pischke (2018) find that women tend to have higher reported job satisfaction if they work in an occupation with higher people and lower brawn content; men exhibit a similar qualitative pattern of reported job satisfaction, but the magnitudes are smaller. The authors also ask a sample of secondary students to choose between pairs of occupations, and find that female students in particular are more likely to choose jobs with higher people content.

[9]In particular, Cortes *et al.* (2018) contend that a female advantage in social skills may explain women's increased representation in cognitive occupations in recent decades. Baker and Cornelson (2018) document evidence of gender differences in sorting on the spatial, motor and sensory skill requirements of jobs.

specific preferences and skills are likely to be correlated, and I cannot shed light on the process of preference and skill formation. However, this project assesses the extent to which gender differences in sorting on tasks can be explained by women and men responding differently to the same wage offer because of their task valuations, rather than women and men receiving different wage offers due to skill differences.

Finally, this project builds on recent studies that use hypothetical choice data to estimate preferences for job amenities (Mas and Pallais 2017; Wiswall and Zafar 2018; Maestas *et al.* 2018; Datta 2019). In particular, Mas and Pallais (2017) estimate WTP for flexible work arrangements and report similar results from field and hypothetical choice experiments, suggesting that a purely hypothetical approach can generate amenity valuations that are relevant for real-world decisions. Similarly, Wiswall and Zafar (2018) use a hypothetical choice experiment to assess WTP for several workplace attributes, and find that estimated preferences predict subsequent college major and job choices and can explain a meaningful share of the gender wage gap in their sample.

The remainder of the paper proceeds as follows. Section 1.2 discusses the gender-typical tasks that I examine. Section 1.3 describes the design of the hypothetical choice experiment. Section 1.4 lays out the model and econometric strategy, and Section 1.5 reports the results of the experiment. Section 1.6 discusses implications of the experimental results for gender gaps in the labor market. Finally, Section 1.7 concludes.

1.2 Gender-Typical Tasks

The goal of the hypothetical choice experiment is to elicit preferences for gender-typical tasks that women are more likely to perform than men, or vice versa. However, representative data on the frequency of task performance among U.S. workers are not available. Therefore, as a proxy for the concept of frequency I use information on the importance of tasks from the O*NET, a U.S. Department of Labor database of occupational characteristics.[10]

[10]The O*NET data are based primarily on surveys of workers in each occupation. The survey questions related to the task importance variables ask respondents, "How important is X to the performance of your

I focus on O*NET measures in the *generalized work activities* and *work styles* domains, which I interpret as providing information about conceptual categories of tasks that are performed in a broad range of jobs. Importantly, these task categories are not explicitly linked to formal educational credentials. To examine gender differences in these measures, which are reported at the occupation level, I link the O*NET variables to information on the share of workers in an occupation who are female. I use data on currently employed workers aged 18 and older from the 2012-2016 American Community Survey (ACS) to construct the occupational female share.[11]

I consider O*NET work activities and work styles that are positively (negatively) correlated with the occupational female share to be measures of female-typical (male-typical) tasks. However, a large number of the work activities and work styles have statistically significant bivariate relationships with the female share, and many of the O*NET variables are correlated with each other. To select a set of these measures for inclusion in the experiment, I follow a hybrid quantitative and qualitative approach.

I estimate a series of ordinary least squares (OLS) regression models predicting the female share based on the O*NET variables. Specifically, within each O*NET domain I regress the female share on: 1) all tasks simultaneously, 2) a group of tasks with the most positive and most negative bivariate coefficients, and 3) a group of tasks that are rated as highly predictive using a random forest algorithm. I also repeat the regression analysis including controls for broad occupation cluster, mean log hourly wages and the share of workers with a college degree or more in each occupation.[12] I then search qualitatively for O*NET variables that are statistically significant and consistent in sign across specification, and combine some similar measures, yielding a final set of five measures. Appendix A.1 provides further details on the process of task selection.

Table 1.1 displays the names of the three female-typical and two male-typical tasks that I

---

current job?", where X is an occupational characteristic.

[11]The ACS collects demographic and socioeconomic information from a random sample of the U.S. population and is administered annually by the U.S. Census Bureau.

[12]Tables A.1, A.2, A.3 and A.4 shows the results of this analysis.

include in the experiment, along with a list of the O*NET variable(s) on which each task is based. The designation of *working and communicating with others and displaying a cooperative attitude*[13] as a female-typical task is consistent with the notion that women may prefer jobs that involve working with people (Lordan and Pischke 2018) or have a comparative advantage in performing interpersonal activities (Cortes *et al.* 2018). The female-typical task of *helping and caring for others* also involves interpersonal interaction; the selection of this task supports evidence that care work is overwhelmingly performed by women (England 2005; Folbre 2012). Similarly, labeling *operating, repairing and maintaining vehicles, devices or equipment*[14] as male-typical is consistent with the hypothesis that men prefer working with "things" (Su *et al.* 2009). The O*NET rates the female-typical task of *documenting and recording information* as highly important in many female-dominated occupations in healthcare, education and social services.[15] Finally, the male-typical task of *making decisions and solving problems* is ranked as important in a range of professional occupations that remain majority male, such as physician, lawyer and many STEM jobs.[16]

In addition to the name of each task, I provide survey participants with a definition and examples, also displayed in Table 1.1. I use the O*NET documentation as a guide, but modify the examples to ensure that they do not reference only female-dominated or only male-dominated occupations. While the O*NET is designed to measure task differences across occupations, the hypothetical scenarios ask participants to consider changing the amount of time spent on a task while keeping other aspects of the job the same. The examples in Table 1.1 are therefore designed to enable participants to envision performing each task in a wide range of jobs.

Figure 1.2 displays mean levels of the five gender-typical tasks for women and men,

[13] Hereafter, I refer to this task as *working and communicating with others*.

[14] Hereafter, I refer to this task as *operating and repairing equipment*.

[15] Figures A.3 and A.4 display the female share and gender-typical task levels by major occupation category.

[16] Tables A.5, A.6, A.7, A.8 and A.9 display the ten occupations with the highest and the ten occupations with the lowest levels of each task measure, along with the female share in each displayed occupation.

Table 1.1: *Gender-Typical Tasks*

Female-Typical Tasks

Helping and caring for others

Definition: Providing personal assistance, medical attention, emotional support, or other personal care to people such as co-workers, customers, or patients.

Examples: Helping a co-worker complete an assignment, assisting a customer in finding a product, or caring for injured people in a hospital.

*O*NET measure(s)*:

Assisting and Caring for Others (work activity)

Documenting and recording information

Definition: Entering, transcribing, recording, storing, or maintaining information in written or electronic form.

Examples: Recording the weights of trucks that use highways, documenting proceedings in a court room, or maintaining information about a patient's health.

*O*NET measure(s)*:

Documenting/Recording Information (work activity)

Working and communicating with others and displaying a cooperative attitude

Definition: Generally working with others rather than alone and being pleasant and good-natured with others on the job.

Examples: Meeting with co-workers to discuss a project, answering a client's questions over the phone, or facilitating a workshop.

*O*NET measure(s)*:

Social Orientation (work style)
Cooperation (work style)

Male-Typical Tasks

Operating, repairing and maintaining vehicles, devices or equipment

Definition: Running, navigating, servicing, repairing, adjusting, or testing vehicles, machines, devices, moving parts, or equipment.

Examples: Driving a car or truck, adjusting the settings on a medical device, or repairing a circuit board.

*O*NET measure(s)*:

Operating Vehicles, Mechanized Devices, or Equipment (work activity)
Repairing and Maintaining Mechanical Equipment (work activity)
Repairing and Maintaining Electronic Equipment (work activity)

Making decisions and solving problems

Definition: Analyzing information and evaluating results to choose the best solution and solve problems.

Examples: Selecting the menu options for a cafeteria, choosing a location for a retail store, or finalizing the budget for a school.

*O*NET measure(s)*:

Making Decisions and Solving Problems (work activity) using the O*NET measures on which the tasks are based.[17] The tasks are rescaled to reflect percentiles weighted by employment in the ACS, such that a value of 50 indicates the median task level. As expected, task levels differ dramatically by gender. Women work in jobs in which the mean percentiles of the female-typical tasks of *helping and caring for others*, *documenting and recording information* and *working and communicating with others* are 17, 11, and 22 percentage points higher, respectively, than the jobs in which men work. By contrast, women work in jobs in which the mean percentile of the male-typical task of *operating and repairing equipment* is 20 percentage points lower, compared with the jobs held by men.[18]

Table 1.2 displays the R^2 and adjusted R^2 statistics from OLS regressions of the female share on all 57 work activity and work style variables in Column 2, and the selected task measures in Column 3.[19] The adjusted R^2 based on all the O*NET variables in Column 2 is 0.77, indicating that these measures can explain a very large fraction of the variation in the female share. Furthermore, the adjusted R^2 of 0.67 in Column 3 suggests that the five selected tasks capture the majority of the female share variation, and more than 85 percent of the variation accounted for by the full set of work activities and work styles.

1.3 Experiment Design

1.3.1 Survey Recruitment and Preliminary Questions

I recruit and compensate participants in the experimental survey using Amazon Mechanical Turk (MTurk). MTurk is a platform that enables researchers and others to pay individuals to perform online activities such as completing surveys.[20] Research suggests that samples

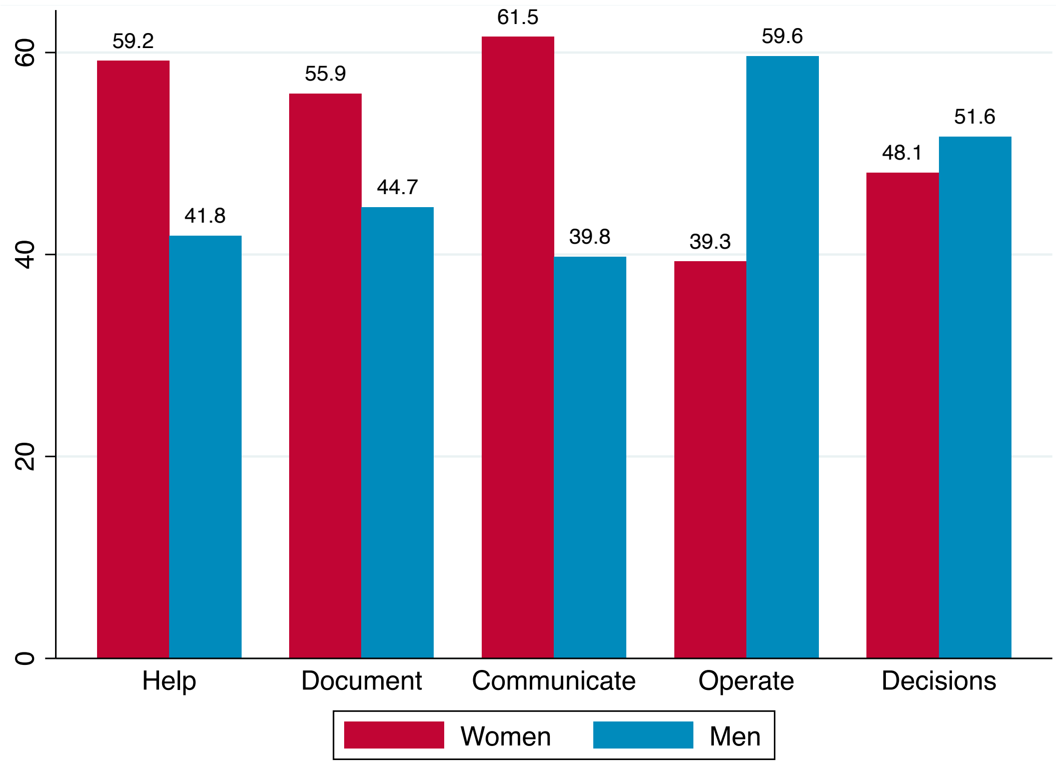
---

[17]For the selected tasks that combine multiple O*NET variables, I average the component variables to create a single measure.

[18]The mean percentile of *making decisions and solving problems* is 3.5 percentage points lower among women than men. However, Table A.1 shows that the relationship between the female share and this task is larger in magnitude when controlling for other work activities.

[19]The O*NET variables are standardized to have a mean of zero and standard deviation of one.

[20]An MTurk requester posts a description of the activity, including the number of participants required, the pay, and typically the amount of time required. Participants complete the activity on a first come, first served



Notes: This figure shows mean levels of the five gender-typical tasks included in the experiment, separately by gender, using the O*NET work activities and work styles listed in Table 1.1. Tasks based on multiple O*NET variables are constructed by averaging the component variables. Each task is measured in percentiles weighted by currently employed individuals aged 18 and older in the 2012-2016 ACS.

Figure 1.2: *Task Percentiles by Gender*

recruited through MTurk have adequate psychometric properties such as internal consistency and test-retest reliability (Buhrmester *et al.* 2011), and that levels of measurement error are similar compared with representative survey samples (Snowberg and Yariv 2018).[21] I administered the survey to 1,931 participants over two rounds in June 2018,[22] restricting the sample to MTurk participants who are U.S. residents and who have an approval rating of at

basis. Requesters then review the work and approve or deny each submission.

[21]In addition, recent studies use MTurk to recruit samples for descriptive surveys of labor market activity (Abraham and Amaya 2018; Katz and Krueger 2019).

[22]Participants were paid $1.80 to take the survey, which required an average of 10 minutes to complete. Compensation was therefore approximately equal to the Massachusetts minimum wage in 2018. Participants completed the survey on the Qualtrics platform, and then submitted a unique completion code on MTurk to receive compensation.

Table 1.2: *Predicting the Occupational Female Share*

	Controls	All Tasks	Selected Tasks
R^2	0.405	0.802	0.672
Adjusted R^2	0.401	0.774	0.668
N	464	464	464

Notes: This table shows the R^2 and adjusted R^2 statistics and the number of observations from a series of OLS regressions in which the outcome is the share of currently employed individuals aged 18 and older who are female in each occupation in the 2012-2016 ACS. In Column 1, the predictors are the mean log hourly wage, the mean log usual hours of work per week, and the share of workers with at least a college degree in each occupation, also from the ACS. In Column 2, the predictors are all 57 variables in the work activities or work styles domains of the O*NET. In Column 3, the predictors are the O*NET work activities and work styles listed in Table 1.1 that are selected for inclusion in the experiment. The O*NET variables are standardized to have a mean of zero and standard deviation of one. For the selected tasks that combine multiple O*NET variables, I average the component variables to create a single measure and re-standardize. The regressions include the 464 ACS occupations that can be matched to the O*NET.

least 95 percent on the platform.[23]

The survey begins by asking about the participant's current employment status, hours of work per week, pay rate, industry and occupation.[24] If participants report that they are not currently employed, then all questions about the current job are modified to refer to the most recent job.[25] I also specify that these questions refer to work other than completing activities on the MTurk platform. Next, the survey asks participants to report the number of hours per week they spend in their current or most recent job performing each of the five gender-typical tasks.[26] After the hypothetical scenarios, the survey gathers information on

[23]MTurk requesters can specify that participants meet certain requirements in order to be eligible to complete an activity. Participant location is self-reported on MTurk, but the Qualtrics platform collects data on the latitude and longitude of the respondent's device or IP address. I exclude from the analysis participants who were physically present outside of the U.S. while completing the survey.

[24]The response options for occupation and industry correspond to major occupation and industry categories using the ACS occupation and the North American Industry Classification System (NAICS) codes.

[25]In addition, I ask non-employed participants about the number of months since they last worked, and if they report that they have never worked I omit questions about the current or most recent job.

[26]The answer choices correspond to intervals that each represent 10 percent of the participant's total weekly

gender, age, race, ethnicity and educational attainment.

1.3.2 Hypothetical Scenarios

The survey asks participants to consider a series of hypothetical scenarios, each associated with one of the five gender-typical tasks. In the scenarios, participants are asked to envision that they have been given a choice between two jobs that differ in terms of pay and the amount of time spent on the focal task. Participants are then asked to indicate which job they would prefer. I randomize the order in which the scenarios appear and the display order of the answer choices within each scenario.

In each hypothetical scenario, one job is randomly selected to offer the participant's wage in the current or most recent job, expressed using the pay period that the participant reports.[27] The other job offers a wage that is higher than the participant's current or most recent wage by a randomly selected percentage from the following set: 1) 0 percent, 2) 5 percent, 3) 10 percent, 4) 15 percent, or 5) 20 percent.[28]

I hypothesize that worker preferences over the amount of time spent on tasks are not finely tuned, such that it is easiest for participants to choose between two mutually exclusive bins, one of which includes their optimal allocation of hours to the focal task, and one of which does not. Therefore, participants are given a choice between spending less than a cutoff of C hours per week on the focal task in one job, and C or more hours per week on the focal task in the other job (termed the *high-task* job).

For a randomly selected half of the sample, C is equal to 10 percent of total hours worked per week in the participant's current or most recent job. This cutoff can be interpreted as

hours of work. There is no restriction that the number of hours reported on the five tasks sum to the total number of hours worked per week, as participants likely spend a portion of their time on tasks that the survey does not ask about, and may also perform some of the gender-typical tasks concurrently.

[27]Participants can report their pay on an hourly, weekly, bi-weekly, twice monthly, monthly or yearly basis. For participants who have never worked, I use a value of $20 per hour for the wage rate in this job. Note that $20/hour falls between the mean hourly wage of $23.86 and median of $17.81 for all U.S. workers, based on the 2016 Occupational Employment Statistics (OES) data.

[28]For *operating and repairing equipment*, the set of percentages is {0%, 5%, 10%, 15%, 20%, 25%}, based on pilot data suggesting that WTP for this task is larger in magnitude, compared with the other tasks.

defining the extensive margin of task performance, as it gives participants a choice between spending little or no time on an activity compared with at least some time. This design has the advantage of involving the same comparison for all tasks.

However, gender differences in preferences for tasks may be largest at cutoffs along the intensive margin of task performance. Therefore, in the other half of the sample I choose a cutoff percentage of time, P, separately for each task, to maximize the difference between the share of women and the share of men who report spending at least P percent of their on the focal task, in a pilot version of the survey.[29] The cutoff number of hours, C, is then equal to P percent of the hours worked per week in the participant's current or most recent job.[30]

Participants are told to assume that other than the wage and the amount of time spent on the focal task, the two jobs in each scenario are exactly the same in all other ways, including total hours worked per week, schedule, co-workers, benefits, and the set of activities they do when not performing the focal task. Emphasizing that other aspects of the job do not vary decreases the probability that participants will view time spent on the focal task as related to factors such as discrimination and the gender composition of co-workers that may be correlated with task performance in real-world settings. In order to clarify that the high-task job does not require working more hours in total, I specify that in both positions, participants would work the same number of hours per week as in their current or most recent job.[31]

The experiment aims to measure preferences over conceptual categories of tasks that may contribute to gender gaps in the labor market. Therefore, the hypothetical scenarios do not provide additional information about job context, such as occupation or industry, that would tie the WTP estimates to a specific set of detailed work activities that constitute

[29]Specifically, the cutoff percentages are 20 percent of hours worked for *helping and caring for others* and *documenting and recording information*, 40 percent of hours worked for *working and communicating with others*, 0 percent of hours worked (which I operationalize as less than 1 hour versus at least one hour) for *operating and repairing equipment*, and 30 percent of hours worked for *making decisions and solving problems*.

[30]For participants who report that they have never worked, the extensive margin cutoff C is equal to 4 hours, and the intensive margin cutoff is equal to P percent of 40 hours.

[31]For participants who have never worked, I specify that both jobs involve 40 hours of work per week.

the focal task in a given context, as well as to a specific set of counterfactual activities that cannot be performed at the same time as the focal task.

As a result, however, the WTP estimates may reflect the distribution of contexts and, in particular, counterfactual activities that participants envision. I assume that participants are most likely to imagine their own current or most recent job. Therefore, the counterfactuals imagined may differ by gender, given the well-documented gender differences in job choices. While I cannot rule out bias from this issue, I discuss evidence in Section 1.5 that suggests that it is unlikely to meaningfully affect results.

Figure 1.3 displays an example of the hypothetical scenarios shown to participants where the focal task is *working and communicating with others*, the cutoff number of hours is on the extensive margin, and the participant reports working 40 hours per week and pay of $20 per hour in the current job.

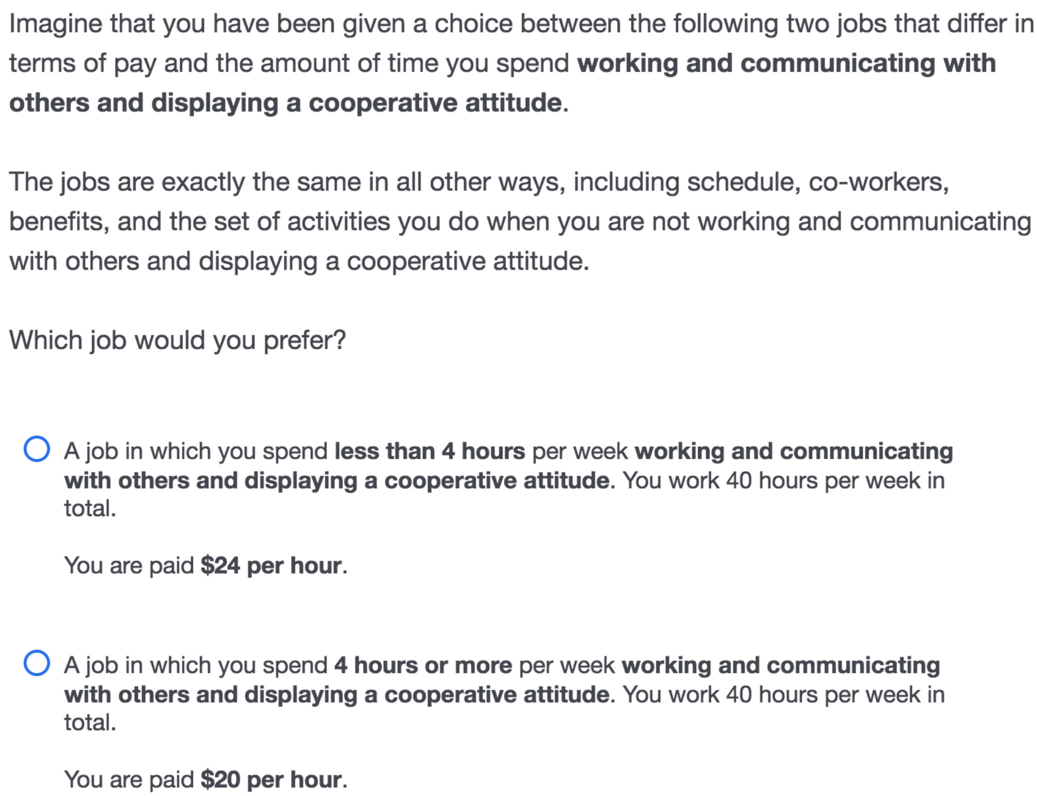
1.4 Model and Econometric Strategy

I use data from participant choices in the hypothetical scenarios to estimate WTP for gender-typical tasks, separately by gender, guided by a simple discrete choice model. As in Rosen (1986), I assume that workers derive utility from the wage and the amenities offered by a job.

In the hypothetical scenarios, participants are instructed to choose between two jobs that are the same except for pay and the amount of time spent on the focal gender-typical task. Thus participant i's utility from job $j \in \{1, 2\}$ related to task $k \in \{1, \ldots, 5\}$ can be expressed as

$$U_{ijk} = \alpha_i + \theta_k T_{jk} + \delta_k \ln w_{ijk} + \varepsilon_{ijk} \tag{1.1}$$

where α_i reflects participant-specific factors affecting utility from work, T_{jk} is an indicator for job j being the high-task job for gender-typical task k (described hereafter as the *high-task k* job or the job with a *high level of task k*), w_{ijk} is the wage offer for participant i in job j


Imagine that you have been given a choice between the following two jobs that differ in terms of pay and the amount of time you spend **working and communicating with others and displaying a cooperative attitude**.

The jobs are exactly the same in all other ways, including schedule, co-workers, benefits, and the set of activities you do when you are not working and communicating with others and displaying a cooperative attitude.

Which job would you prefer?

- A job in which you spend **less than 4 hours** per week **working and communicating with others and displaying a cooperative attitude**. You work 40 hours per week in total.

  You are paid **$24 per hour**.

- A job in which you spend **4 hours or more** per week **working and communicating with others and displaying a cooperative attitude**. You work 40 hours per week in total.

  You are paid **$20 per hour**.

Notes: This figure shows an example of a scenario from the hypothetical choice experiment related to the task *working and communicating with others* for a participant who reports working 40 hours per week and being paid $20/hour in the current or most recent job.

Figure 1.3: *Example of Hypothetical Scenario*

related to task k, θ_k and δ_k are preference parameters indexed by task, and ε_{ijk} is a worker-, job- and task-specific preference parameter. For convenience, let $j = 1$ index the high-task k job, such that $T_{1k} = 1$ and $T_{2k} = 0$, $\forall k$.

The parameter θ_k can be interpreted as reflecting preferences for the task k amenity, while δ_k reflects preferences over wages. The magnitude of θ_k relative to δ_k determines mean WTP for task k. The ε_{ijk} parameter shifts WTP for individual i. I assume that ε_{ijk} has a standard Extreme Value (EV) Type I distribution.[32]

[32]The EV Type I distribution describes the behavior of a random variable that is the maximum of some

An important benefit of the experimental setting is that w_{ijk} is observed for both jobs in the participant i's choice set; by contrast, standard survey and administrative data sources contain information only on realized wages. In addition, amenities or other characteristics of the hypothetical job that participant i envisions that are not directly affected by the amount of time spent on task k will not vary between jobs, and will be absorbed by the α_i term.

Participant i chooses the job with the high level of task k if

$$
\begin{aligned}
U_{i1k} &> U_{i2k} \\
\kappa_{ik} &< \theta_k + \delta_k \omega_{ik}
\end{aligned} \tag{1.2}
$$

where $\kappa_{ik} \equiv \varepsilon_{i2k} - \varepsilon_{i1k}$ and $\omega_{ik} \equiv \ln(w_{i1k}/w_{i2k})$ is the log difference in wage offers between the high-task and low-task jobs for individual i. The EV Type I assumption implies that κ_{ik} has a standard logistic distribution. Therefore, the parameters δ_k and θ_k can be estimated by logistic regression, where the outcome, y_{ik}, is an indicator for participant i choosing the high-task k job, and the predictors are an intercept term and ω_{ik}.[33]

My primary hypothesis is that δ_k and θ_k differ by gender. I therefore estimate the model for the entire experiment sample and separately by gender, yielding coefficients $\hat{\delta}_{gk}$ and $\hat{\theta}_{gk}$, where $g \in \{a, f, m\}$ indexes all participants, women and men, respectively.

To derive an expression for WTP, note that the logistic distribution for κ_{ik} implies that for an individual of gender g, the probability of choosing the high-task k job is given by

$$
\begin{aligned}
\Pr(y_{ik} = 1) &= \mathrm{F}_\kappa(\theta_{gk} + \delta_{gk}\omega_{ik}) \\
&= \frac{1}{1 + \exp(-(\theta_{gk} + \delta_{gk}\omega_{ik}))}
\end{aligned} \tag{1.3}
$$

where $\mathrm{F}_\kappa(\cdot)$ is the CDF of κ_{ik}. Conditional on the parameters for gender g, this probability

underlying sample. This distributional assumption can be motivated by the notion that preferences are shaped by repeated exposure to a task or bundle of tasks.

[33]This strategy follows the approach used in recent literature estimating preferences for job amenities based on discrete choice experiments (Mas and Pallais 2017; Maestas *et al.* 2018).

can be expressed as a function of ω_k:

$$\mathrm{H}_{gk}(\omega_k) \equiv \mathrm{F}_\kappa(\theta_{gk} + \delta_{gk}\omega_k). \tag{1.4}$$

For each value of ω_k, the share choosing the high-task job is the proportion of participants willing to pay $100 * (1 - \exp(\omega_k))$ percent of their wage to spend more time on task k.

At the mean and median of the κ_{ik} distribution for gender g, $\Pr(y_{ik} = 1) = 0.5$ and thus $\omega_k = -\theta_{gk}/\delta_{gk}$. Therefore, the mean WTP for more time spent on task k as a proportion of the wage among individuals of gender g is given by

$$\lambda_{gk} \equiv 1 - \exp\left(-\frac{\theta_{gk}}{\delta_{gk}}\right). \tag{1.5}$$

I estimate λ_{gk} for $g \in \{a, f, m\}$ and the gender difference in λ_{gk}:

$$\beta_k \equiv \lambda_{fk} - \lambda_{mk}, \tag{1.6}$$

using the Delta method to calculate robust standard errors.

I hypothesize that $\beta_k > 0$ if task k is female-typical and $\beta_k < 0$ if task k is male-typical. Furthermore, if $\delta_{fk} \approx \delta_{mk}$, then $\mathrm{H}_{fk}(\omega_k) > \mathrm{H}_{mk}(\omega_k)$ when $\beta_k > 0$ and $\mathrm{H}_{fk}(\omega_k) < \mathrm{H}_{mk}(\omega_k)$ when $\beta_k < 0$, for almost all values of ω_k. Therefore, the model predicts that if women (men) have a higher WTP for task k at the mean, then more women (men) will generally sort into high-task k jobs.

1.5 WTP Results

1.5.1 Descriptive Statistics

This section presents descriptive statistics for the experiment sample recruited using the MTurk platform. Table 1.3 displays summary statistics for currently employed individuals in the experiment sample, overall and by gender, in Columns 1-3, and comparable statistics for employed individuals aged 18 and older in the 2012-2016 ACS in Columns 4-6.[34] I focus

[34] Table A.10 displays comparable statistics for employed and non-employed individuals in both samples.

Table 1.3: *Summary Statistics - Employed Only*

	Experiment			ACS		
	All	Women	Men	All	Women	Men
Female	0.526	1.000	0.000	0.472	1.000	0.000
Age	34.4	35.1	33.5	42.0	42.0	42.1
White	0.719	0.769	0.663	0.648	0.645	0.651
Black	0.080	0.073	0.087	0.110	0.126	0.096
Hispanic	0.111	0.079	0.147	0.161	0.147	0.174
Other Race	0.090	0.079	0.103	0.080	0.082	0.079
HS or less	0.084	0.086	0.082	0.337	0.293	0.376
Some college	0.229	0.220	0.239	0.238	0.247	0.230
Associate's degree	0.124	0.144	0.101	0.090	0.104	0.077
Bachelor's degree	0.412	0.385	0.443	0.212	0.224	0.201
Graduate degree	0.151	0.164	0.135	0.123	0.131	0.116
Hours per week	37.6	36.4	39.0	39.3	36.6	41.7
Wage (hourly)	21.01	18.68	23.67	25.05	22.17	27.67
N	1,742	917	825	7,031,598	3,367,987	3,663,611

Notes: This table shows summary statistics in the experiment sample compared with the 2012-2016 ACS, restricting to currently employed participants in the experiment sample and currently employed individuals aged 18 and older in the ACS. Statistics in the ACS are weighted by the ACS person weight.

on the statistics for employed individuals because 90 percent of the experiment sample reports being currently employed, in contrast to 60 percent of individuals in the ACS.

The experiment sample is substantially younger (34 versus 42 years in Table 1.3), more likely to be White (72 versus 65 percent of the sample), and more educated, compared to the overall U.S. population as captured by the ACS. Specifically, 56 percent of the experiment sample reports having at least a bachelor's degree and 8 percent has a high school diploma or less, while 34 percent of the employed ACS sample falls in each of these categories. Experiment participants also report modestly lower hourly wages ($21/hour versus $25/hour in the ACS).[35]

The experiment sample is 53 percent female. Women in the experiment sample are substantially more likely than men to be White (77 percent versus 66 percent). In addition, women are less likely than men to be employed (89 versus 92 percent in Table A.10), work fewer hours per week (36 versus 39 hours), and have lower hourly wages ($19/hour versus $24/hour).[36]

Figures 1.4 and 1.5 show the distribution of employment by major occupation and industry categories in the experiment sample compared with the ACS.[37] Computer and math, education, arts, sports and media, and business operations and finance occupations are substantially over-represented in the experiment sample, while managerial, production, maintenance and transportation, health practitioner and technician, and personal care and cleaning occupations are under-represented. Among industry categories, information, educational services, professional, scientific and technical services, finance and insurance, administrative services, and arts, entertainment and recreation are over-represented in the experiment sample, while jobs in health and social assistance, manufacturing, construction and extraction, public administration, and wholesale trade are under-represented.

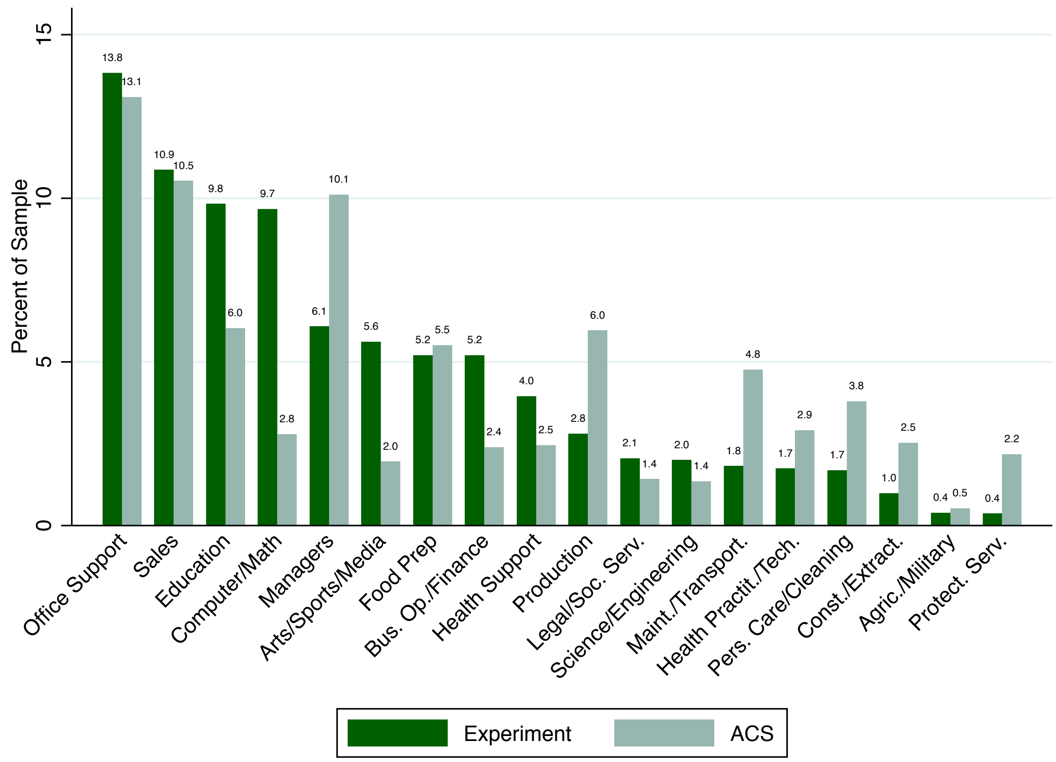
Figure 1.6 displays the percentage of weekly hours worked that participants report spending on the five gender-typical tasks, in 10 percentage point intervals. The distributions reveal substantial heterogeneity across task; for example, only 1 percent of the sample spends no time *working and communicating with others*, while 54 percent spends no time *operating and repairing equipment*.

---

[35]In both samples, I exclude hourly wage observations that are less than $3 or greater than $200. Hourly wages in the ACS are calculated as annual earnings divided by annual hours of work, and are inflated to 2018 dollars.

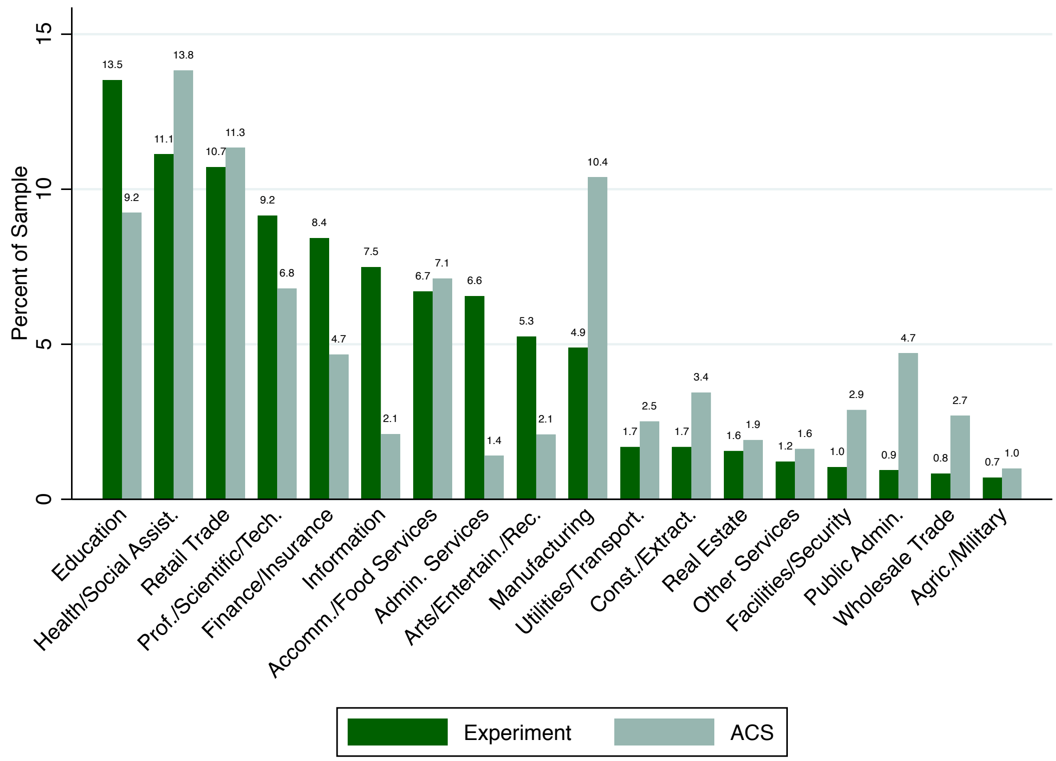
[36]The comparable gender gaps in employment, hours of work and hourly wages have the same sign in the ACS. However, women are slightly less likely to have a college degree in the experiment sample (55 versus 58 percent in Table 1.3), but are more likely to have a college degree in the ACS (36 versus 32 percent). I estimate a version of the WTP analysis in which I weight the experiment sample to match the ACS by gender, race, educational attainment and major occupation, to ensure that factors such as the education distribution and gender differences in the racial and ethnic composition of the experiment sample are not driving results.

[37]The figures use data on the most recent job for participants in the experiment sample who are not currently employed, but restrict to employed workers in the ACS.



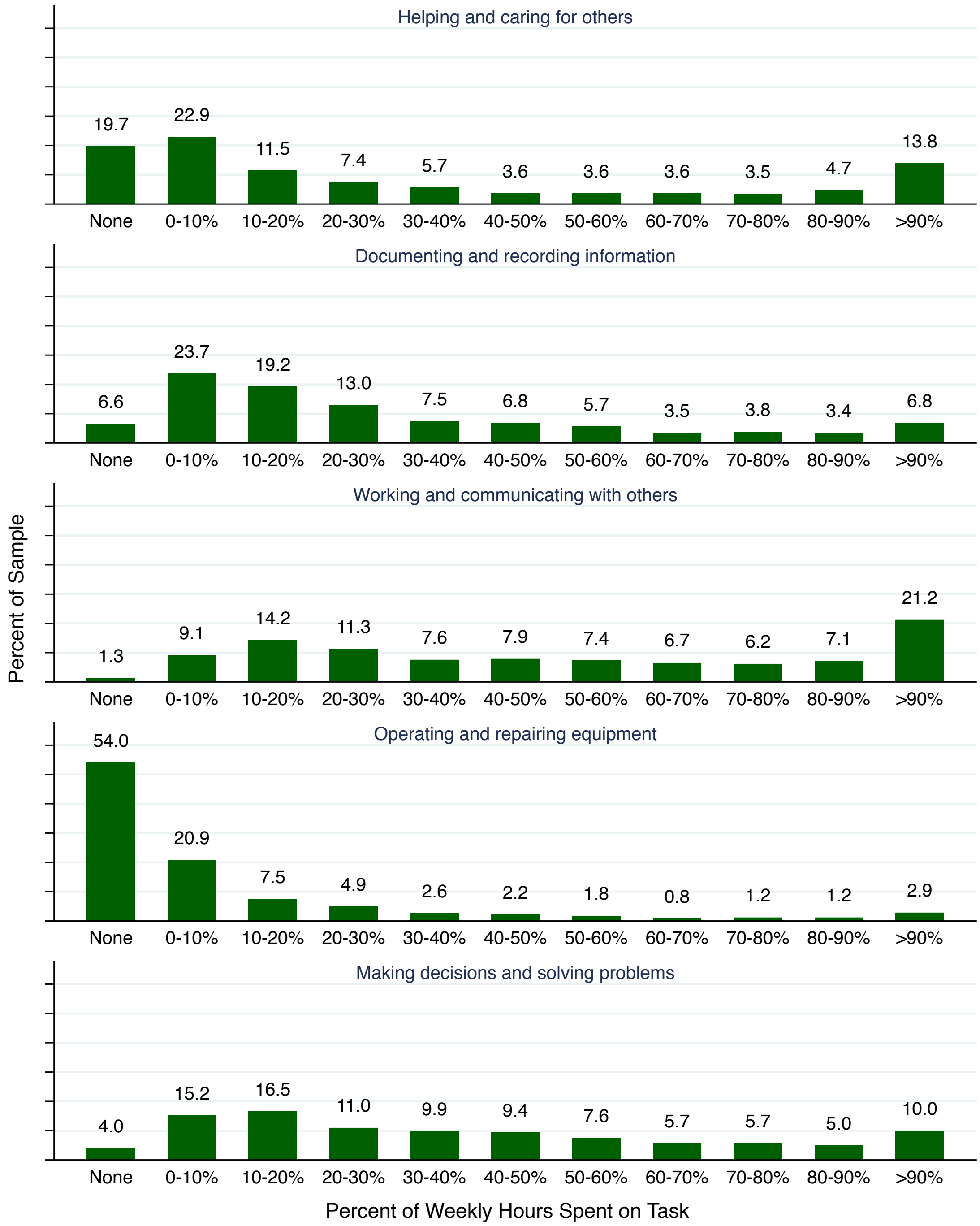
Notes: This figure shows the share of the experiment sample compared with the 2012-2016 ACS in each occupation category. The data from the experiment sample include occupation in the most recent job for participants who are not currently employed. The data from the ACS include only currently employed individuals aged 18 and older. Occupation categories are based on the two-digit occupation codes in the ACS, with some additional aggregation.

Figure 1.4: *Occupation Categories in Experiment Sample vs. ACS*



Notes: This figure shows the share of the experiment sample compared with the 2012-2016 ACS in each industry category. The data from the experiment sample include industry in the most recent job for participants who are not currently employed. The data from the ACS include only currently employed individuals aged 18 and older. Industry categories are based on the two-digit North American Industry Classification System (NAICS) codes, with some modifications.

Figure 1.5: *Industry Categories in Experiment Sample vs. ACS*



Notes: This figure shows the distribution of time spent on each gender-typical task in the current or most recent job among experiment participants.

Figure 1.6: *Task Distributions - Entire Sample*

Figure 1.7 shows the distribution of time spent on tasks separately for women and men. It is clear that women spend substantially more time than men on the female-typical tasks of *helping and caring for others*, *working and communicating with others*, and *documenting and recording information*, and less time on the male-typical task of *operating and repairing equipment*. For example, women are 17 percentage points more likely than men to spend at least 50 percent of their time *helping and caring for others*, and 18 percentage points more likely than men to spend no time *operating and repairing equipment*.[38] These results suggest that the O*NET variables measuring task importance do capture information about the frequency of task performance, as hypothesized.

1.5.2 Baseline WTP

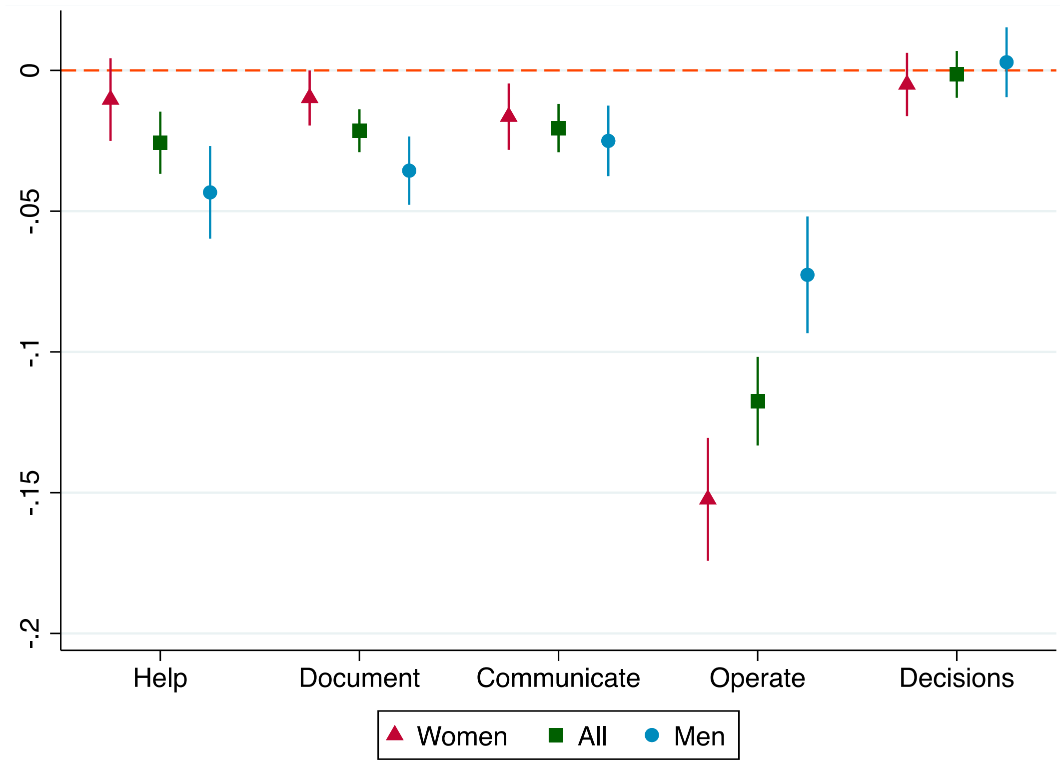
Table 1.4 reports mean WTP as a proportion of the wage for the jobs that involve spending more time on the gender-typical tasks (i.e. the high-task jobs), using data from the hypothetical choice experiment. This table is based on job choices from all participants, regardless of whether they face the extensive or intensive margin cutoffs in the hypothetical scenarios. Each cell in the first three rows of the table gives an estimate of λ_{gk} from a gender g- and task k-specific regression. The first, second and third rows display results for all participants, women and men, respectively, while the columns indicate task. The final row presents estimates of β_k, the female-to-male difference in WTP. Figures 1.8 and 1.9 plot the estimates of λ_{gk} and β_k, respectively.

In the sample as a whole, WTP is negative or close to zero for all gender-typical tasks. Specifically, the estimate in Column 1 of Table 1.4 indicates that participants must be compensated an additional 2.6 percent in order to be willing to work in a job with a high level of *helping and caring for others*. Similarly, participants must be paid approximately 2 percent more to spend more time *documenting and recording information* or *working and communicating with others*. The WTP estimate for *operating and repairing equipment* is largest in

[38] The distribution of time spent on the male-typical task of *making decisions and solving problems* is fairly similar across genders, which is consistent with the finding in Figure 1.2 that the overall gender difference in the mean level of this task in the O*NET is relatively small.

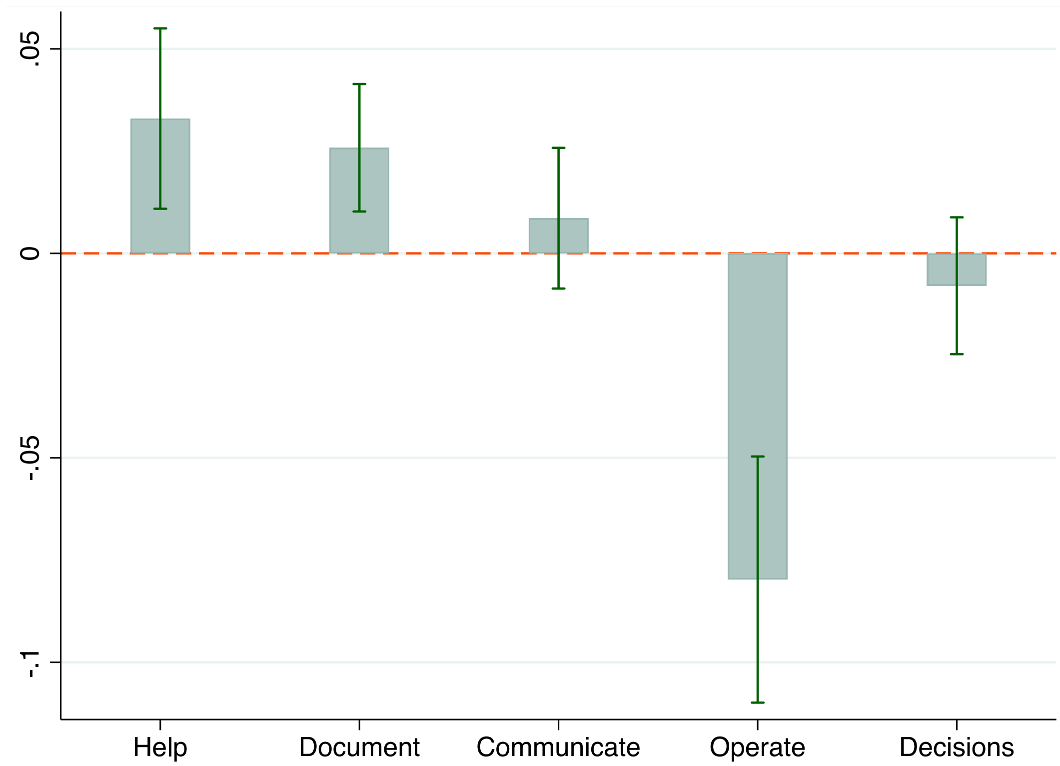
Notes: This figure shows the distribution by gender of time spent on each gender-typical task in the current or most recent job among experiment participants.

Figure 1.7: *Task Distributions - By Gender*



Notes: This figure plots the estimates from Table 1.4 of WTP for spending more time on the gender-typical tasks as a proportion of the wage for the entire sample (λ_{ak}), women (λ_{fk}) and men (λ_{mk}), using choice data from the experiment.

Figure 1.8: *WTP for Tasks as Share of Wage*



Notes: This figure plots the estimates from Table 1.4 of the female-to-male difference in WTP (β_k), using choice data from the experiment.

Figure 1.9: *WTP for Tasks: Gender Difference (W−M)*

Table 1.4: *WTP for Tasks as Share of Wage*

	Help	Document	Communic.	Operate	Decisions
All	-0.026**	-0.021**	-0.020**	-0.118**	-0.001
	(0.006)	(0.004)	(0.004)	(0.008)	(0.004)
Women	-0.010	-0.010*	-0.016**	-0.152**	-0.005
	(0.007)	(0.005)	(0.006)	(0.011)	(0.006)
Men	-0.043**	-0.036**	-0.025**	-0.073**	0.003
	(0.008)	(0.006)	(0.006)	(0.011)	(0.006)
Diff (W-M)	0.033**	0.026**	0.009	-0.080**	-0.008
	(0.011)	(0.008)	(0.009)	(0.015)	(0.009)
N	1,931	1,931	1,931	1,931	1,931

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows estimates of WTP for spending more time on the gender-typical tasks as a proportion of the wage for the entire sample (λ_{ak}), women (λ_{fk}) and men (λ_{mk}), and the female-to-male difference in WTP (β_k), using choice data from the experiment.

absolute value; participants require an additional 12 percent in pay to be willing to work in a job with a high level of this activity. Finally, WTP for *making decisions and solving problems* is very close to zero and insignificant.

It is striking that there is no task for which mean WTP is positive in the overall sample, despite the fact that for all tasks except *operating and repairing equipment*, over half the sample reports working in a high-task job.[39] Indeed, WTP among women in the second row is negative for all three female-typical tasks and statistically significant for two of the three (*documenting and recording information* and *working and communicating with others*), while WTP among men in the third row is negative and statistically significant for the male-typical task of *operating and repairing equipment*. It may be that many workers have substantial "white space" or downtime in their jobs during which they are essentially idle. Therefore, participants may interpret spending more time on any one activity as requiring greater effort because it reduces the downtime available to them.

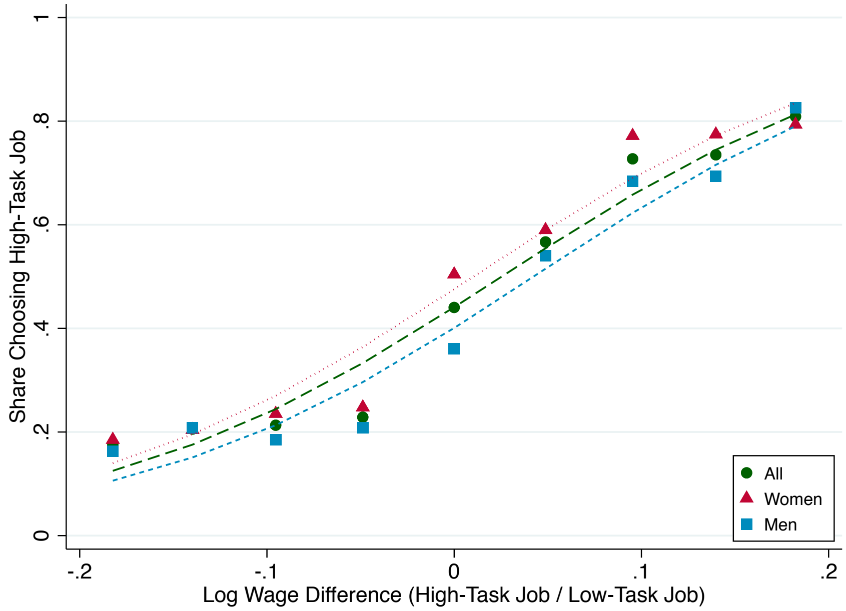
[39] I define participants as working in a high-task k job if they report spending more time on task k in their current or most recent job than the cutoff number of hours in the scenario related to task k that they are shown.

The final row of Table 1.4 shows evidence of significant gender differences in WTP for three of the five tasks examined. Women are willing to pay 3.3 and 2.6 percentage points more than men for the female-typical tasks of *helping and caring for others* and *documenting and recording information*, respectively. Similarly, men's WTP for the male-typical task of *operating and repairing equipment* is 8.0 percentage points higher than the estimate for women. These significant gender differences are consistent with the hypothesis that $\beta_k > 0$ ($\beta_k < 0$) for female-typical (male-typical) tasks.

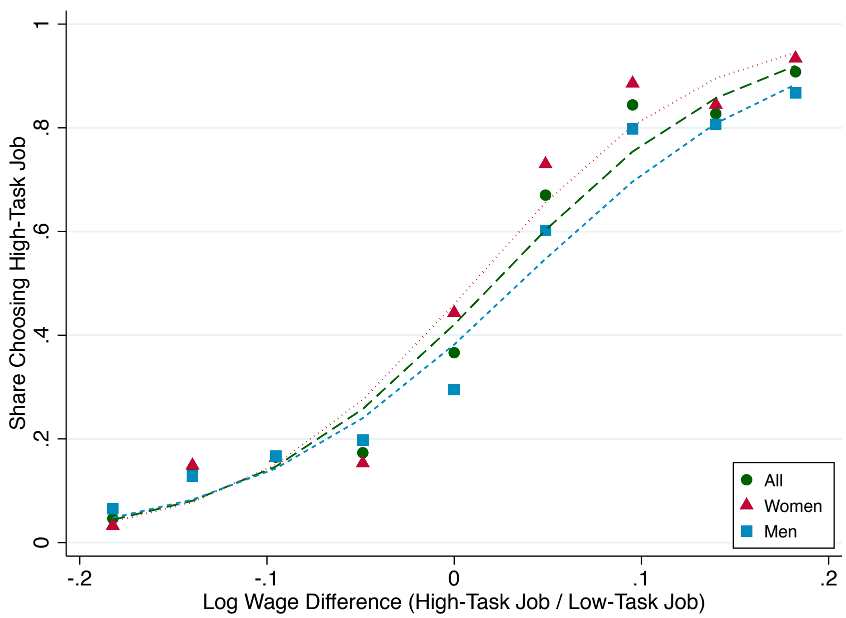
I find no significant gender differences in WTP for the female-typical task of *working and communicating with others* or the male-typical task of *making decisions and solving problems*. The finding for *making decisions and solving problems* suggests that the observed correlation between this task and the female share may be due to factors such as discrimination that constrain women from entering jobs that offer decision-making authority. The result for *working and communicating with others* is surprising given the large observed gender difference in this task in the O*NET in Figure 1.2.

Figures 1.10 and 1.11 plot the share of participants choosing the high-task job for each gender-typical task against the log difference in wage offers, for the entire sample and for women and men separately, along with the predicted probabilities from the logistic specification. It is clear that women are more likely to choose a job with a high level of *helping and caring for others*, and men are more likely to choose a job with a high level of *operating and repairing equipment*, at nearly all wage differentials. By contrast, it is evident from the figures that there is little gender difference in job choices at any wage differential for *working and communicating with others* or *making decisions and solving problems*. Gender differences in choices related to *documenting and recording information* occur primarily with a wage differential of zero or when the high-task job offers a higher wage.

When the wage offer is the same in both jobs, women are 14 percentage points more likely than men to choose a job with a high level of *helping and caring for others*, 15 percentage points more likely than men to choose a job with a high level of *documenting and recording information*, and 24 percentage points less likely than men to choose a job with a high level



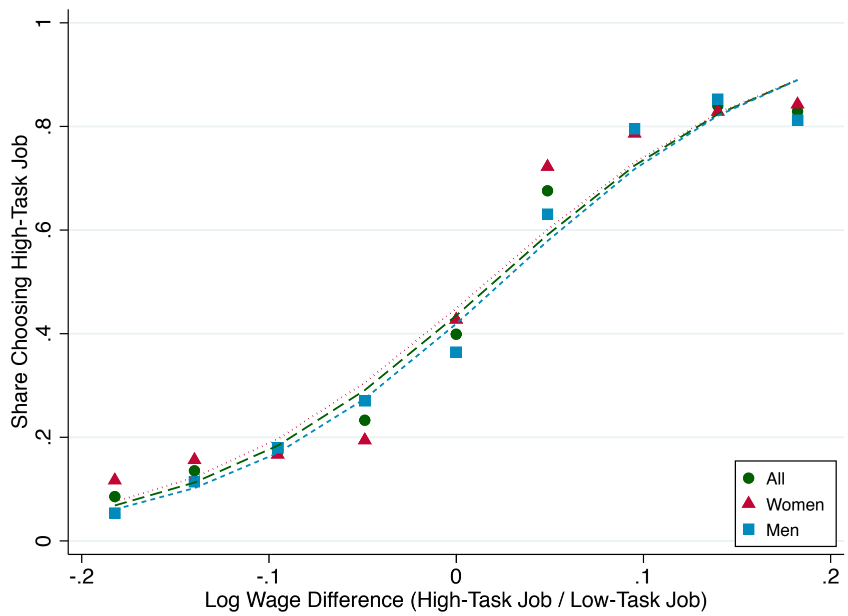
(a) Helping and caring for others



(b) Documenting and recording information

Figure 1.10: *Job Choices - Female-Typical Tasks*

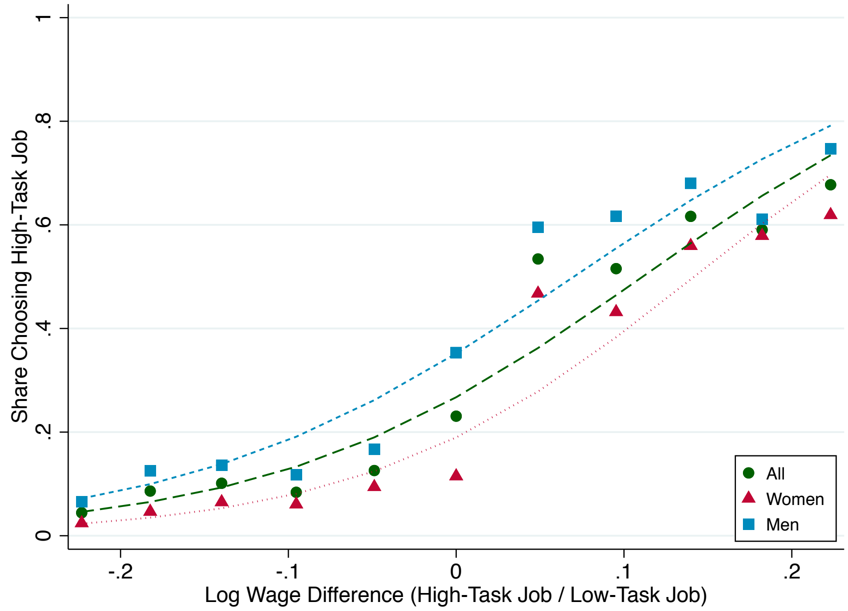
Continued on next page



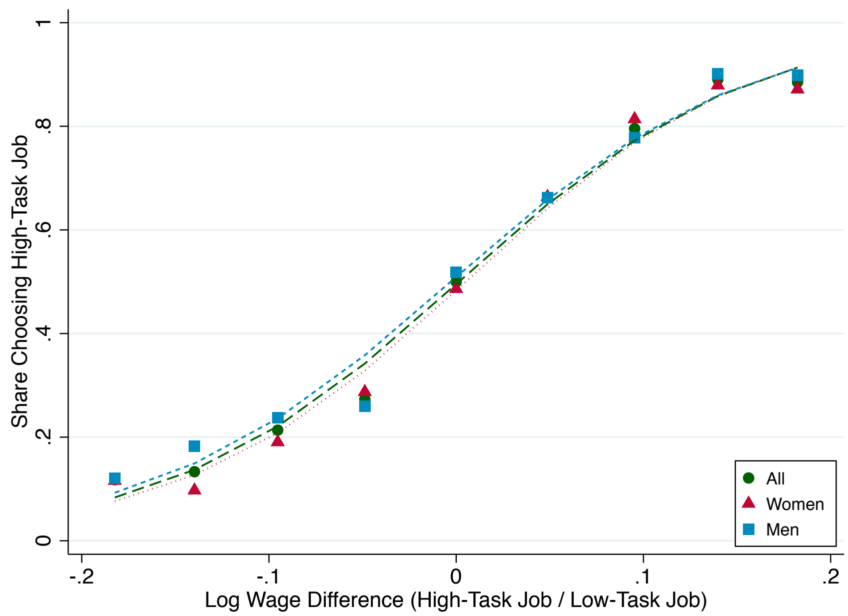
(c) Working and communicating with others

Notes: This figure shows the share of all experiment participants, women and men choosing the high-task job in the hypothetical scenarios as a function of the log difference in wage offers, for female-typical tasks. The high-task job is the job that involves spending more than the cutoff amount of time on the focal task.

Figure 1.10: *(Continued) Job Choices - Female-Typical Tasks*



(a) Operating and repairing equipment



(b) Making decisions and solving problems

Notes: This figure shows the share of all experiment participants, women and men choosing the high-task job in the hypothetical scenarios as a function of the log difference in wage offers, for male-typical tasks. The high-task job is the job that involves spending more than the cutoff amount of time on the focal task.

Figure 1.11: *Choice of High-Task Job - Male-Typical Tasks*

of *operating and repairing equipment*. These differences are statistically significant.[40]

1.5.3 Heterogeneity and Robustness Checks

Table 1.5 reports WTP estimates separately for participants shown hypothetical scenarios with extensive margin cutoffs (Panel A) and intensive margin cutoffs (Panel B), as well as the differences between the estimates (Panel C).[41] Results for both sub-samples are qualitatively similar to those in Table 1.4, although WTP estimates are generally more negative on the intensive margin, consistent with participants having a distaste for spending a greater amount of time on the gender-typical tasks. The gender difference in WTP for *helping and caring for others* is also larger in magnitude on the intensive margin (5.9 percentage points versus an insignificant 1.4 percentage points on the extensive margin).

Table A.11 displays WTP estimates separately for participants with at least a college degree and those without a college degree. Results are similar across groups and compared with baseline estimates in Table 1.4. The exception is that WTP for *making decisions and solving problems* is positive and significant among college workers and negative and significant among non-college workers, and the difference across groups is also significant (4.5 percentage points when pooling women and men). This finding is consistent with college-educated workers being more likely to hold jobs involving problem-solving and high-stakes decision-making.

Tables A.12 and A.13 show WTP estimates excluding participants who are inattentive and who are not currently employed, respectively. I measure inattention by asking participants at the end of the survey to indicate the decisions they made in the hypothetical choice experiment for a randomly selected two of the five gender-typical tasks. I consider a

[40]Women are also are 6 percentage points more likely to choose a job with a high level of *working and communicating with others* and 3 percentage points less likely to choose a job with a high level of *making decisions and solving problems*, but these differences are not significant.

[41]As described in Section 1.3.2, participants are asked to choose between a job in which they spend less than a cutoff number of hours, C, on the focal task, and a job in which they spend C or more hours on the focal task; the value of C depends on whether participant is assigned to the extensive or the intensive margin cutoffs.

Table 1.5: *WTP for Tasks - Extensive vs. Intensive Margin*

(a) Extensive Margin

	Help	Document	Communic.	Operate	Decisions
All	-0.015^{+}	-0.016**	0.003	-0.127**	0.006
	(0.007)	(0.005)	(0.006)	(0.012)	(0.006)
Women	-0.008	-0.004	0.006	-0.155**	0.001
	(0.011)	(0.006)	(0.009)	(0.016)	(0.009)
Men	-0.022*	-0.030**	-0.000	-0.089**	0.011
	(0.010)	(0.008)	(0.009)	(0.017)	(0.009)
Diff (W-M)	0.014	0.026*	0.006	-0.065**	-0.009
	(0.015)	(0.010)	(0.012)	(0.024)	(0.012)
N	953	953	953	953	953

(b) Intensive Margin

	Help	Document	Communic.	Operate	Decisions
All	-0.038**	-0.028**	-0.042**	-0.109**	-0.008
	(0.009)	(0.006)	(0.006)	(0.011)	(0.006)
Women	-0.013	-0.016*	-0.037**	-0.150**	-0.010
	(0.011)	(0.008)	(0.009)	(0.015)	(0.008)
Men	-0.071**	-0.043**	-0.048**	-0.058**	-0.005
	(0.015)	(0.010)	(0.009)	(0.014)	(0.009)
Diff (W-M)	0.059**	0.027*	0.011	-0.092**	-0.005
	(0.018)	(0.013)	(0.012)	(0.020)	(0.012)
N	978	978	978	978	978

Continued on next page

Table 1.5: *(Continued) WTP - Extensive vs. Intensive Margin*

(c) Difference (Extensive − Intensive)

	Help	Document	Communic.	Operate	Decisions
All	0.023*	0.012	0.045**	-0.018	0.014+
	(0.011)	(0.008)	(0.009)	(0.016)	(0.009)
Women	0.004	0.013	0.042**	-0.004	0.011
	(0.015)	(0.010)	(0.012)	(0.022)	(0.011)
Men	0.049**	0.013	0.047**	-0.031	0.016
	(0.018)	(0.013)	(0.013)	(0.022)	(0.013)
Diff (W-M)	-0.045+	-0.000	-0.005	0.027	-0.004
	(0.024)	(0.016)	(0.018)	(0.031)	(0.017)
N	1,931	1,931	1,931	1,931	1,931

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates for experiment participants who face extensive margin cutoffs in the hypothetical scenarios in Panel A, WTP estimates for participants who face intensive margin cutoffs in Panel B, and the difference between the estimates in Panel C.

participant to be inattentive if they answer either question incorrectly.[42] Results among attentive and employed participants are similar to baseline estimates in Table 1.4.

1.5.4 External Validity

If workers sort according to preferences for tasks, then individuals working in jobs that involve more time spent on a task are predicted to have a higher WTP for that activity. To test this hypothesis, Table 1.6 shows WTP for participants currently working in a high-task job (Panel A) versus those currently working in a low-task job (Panel B), with the difference between the estimates in Panel C. I designate participants as working in a high-task k job if the number of hours they report spending on task k in their current or most recent job is greater than the cutoff number of hours in the hypothetical scenario they face for that task.

Consistent with the prediction, WTP estimates for those currently in high-task jobs

[42] I find that 18 percent of participants are inattentive using this definition.

Table 1.6: *WTP for Tasks - Currently in High-Task vs. Low-Task Job*

(a) Currently in High-Task Job

	Help	Document	Communic.	Operate	Decisions
All	0.015*	-0.010*	-0.001	-0.047**	0.014**
	(0.007)	(0.005)	(0.005)	(0.011)	(0.005)
Women	0.020*	0.001	0.003	-0.085**	0.013+
	(0.008)	(0.006)	(0.006)	(0.016)	(0.007)
Men	0.007	-0.026**	-0.005	-0.019	0.014+
	(0.011)	(0.008)	(0.007)	(0.014)	(0.007)
Diff (W-M)	0.013	0.027**	0.008	-0.067**	-0.001
	(0.014)	(0.010)	(0.010)	(0.021)	(0.010)
N	988	1,156	1,408	694	1,272

(b) Currently in Low-Task Job

	Help	Document	Communic.	Operate	Decisions
All	-0.075**	-0.042**	-0.077**	-0.157**	-0.030**
	(0.010)	(0.007)	(0.010)	(0.011)	(0.007)
Women	-0.063**	-0.034**	-0.086**	-0.181**	-0.039**
	(0.015)	(0.011)	(0.016)	(0.015)	(0.010)
Men	-0.084**	-0.049**	-0.068**	-0.116**	-0.019
	(0.012)	(0.010)	(0.013)	(0.016)	(0.012)
Diff (W-M)	0.021	0.014	-0.018	-0.065**	-0.020
	(0.019)	(0.015)	(0.020)	(0.022)	(0.015)
N	935	767	515	1,229	651

Continued on next page

Table 1.6: *(Continued) WTP - Currently in High-Task vs. Low-Task Job*

(c) Difference (High-Task − Low-Task)

	Help	Document	Communic.	Operate	Decisions
All	0.090**	0.033**	0.076**	0.109**	0.044**
	(0.012)	(0.009)	(0.011)	(0.015)	(0.009)
Women	0.083**	0.036**	0.089**	0.095**	0.052**
	(0.017)	(0.012)	(0.017)	(0.022)	(0.012)
Men	0.091**	0.023+	0.063**	0.097**	0.033*
	(0.017)	(0.013)	(0.015)	(0.021)	(0.014)
Diff (W-M)	-0.008	0.013	0.026	-0.002	0.019
	(0.024)	(0.018)	(0.023)	(0.031)	(0.018)
N	1,923	1,923	1,923	1,923	1,923

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates for experiment participants who report currently working in a high-task job in Panel A, WTP estimates for participants who report currently working in a low-task job in Panel B, and the difference between the estimates in Panel C. I designate participants as working in a high-task k job if the number of hours they report spending on task k in their current or most recent job is greater than the cutoff number of hours in the hypothetical scenario they face for that task.

are higher than estimates for those in low-task jobs, for all tasks and both genders.[43] For example, workers who currently spend more *working and communicating with others* are willing to pay 7.6 percentage points more for this task than those who currently spend less time on this activity. In addition, overall WTP for *helping and caring for others* and *making decisions and solving problems* is positive (approximately 1.5 percent) and statistically significant among those currently spending more time on these tasks. These results indicate that WTP estimates are correlated with equilibrium labor market outcomes, suggesting that the hypothetical choice experiment measures task valuations with real-world relevance.

The qualitative pattern of gender differences in WTP in Table 1.6 matches the baseline results in Table 1.4. However, in some cases the differences are smaller in magnitude and lose significance.[44] This finding is not surprising, as one might expect gender differences to shrink or even disappear among workers who make the same sorting decisions and therefore are likely to have more homogeneous preferences, compared with the overall population.

The experiment sample is not statistically representative of the broader U.S. population, and Section 1.5.1 documents that the distributions of race, education, occupation and industry differ meaningfully between the experiment sample and the ACS. To assess how this selection may affect results, in Table 1.7 I repeat the WTP analysis, weighting the sample to match currently employed workers in the ACS by gender, race (White versus non-White), college degree receipt and major occupation. Results are similar to the unweighted estimates, although the gender difference in WTP for *helping and caring for others* is somewhat larger in magnitude (7.0 percentage points compared with 3.3 percentage points).

[43] The differences in WTP between those in high-task and low-task jobs are statistically significant for all tasks when pooling women and men, and for nearly all tasks when considering women and men separately.

[44] Specifically, the gender difference in WTP for *helping and caring for others* is insignificant in both the high-task and low-task sub-samples, and the gender difference in *documenting and recording information* is insignificant in the low-task sub-sample.

Table 1.7: *WTP for Tasks - Weighted to Match ACS*

	Help	Document	Communic.	Operate	Decisions
All	-0.039**	-0.029**	-0.028**	-0.099**	-0.008
	(0.010)	(0.007)	(0.008)	(0.013)	(0.006)
Women	-0.004	-0.015+	-0.025**	-0.138**	-0.013+
	(0.011)	(0.008)	(0.009)	(0.017)	(0.007)
Men	-0.074**	-0.043**	-0.030*	-0.055**	-0.003
	(0.018)	(0.011)	(0.012)	(0.018)	(0.010)
Diff (W-M)	0.070**	0.027*	0.005	-0.083**	-0.010
	(0.021)	(0.014)	(0.015)	(0.024)	(0.012)
N	1,728	1,728	1,728	1,728	1,728

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates for currently employed individuals in the experiment sample that have been weighted to match currently employed workers aged 18 and older in the 2012-2016 ACS by gender, race (White versus non-White), college degree receipt, and major occupation category.

1.5.5 Interpretation

The goal of the experiment is to ensure that participants make choices only on the basis of preferences over pay and gender-typical tasks. Therefore, wages are randomly assigned and participants are instructed to assume that jobs are the same except for explicitly stated differences. However, it is possible that WTP estimates may still reflect task-specific skills, concerns about discrimination or preferences for job characteristics that participants view as correlated with tasks.

To investigate this possibility, at the end of the survey I ask participants about the motivations for their choices in a randomly selected two out of the five hypothetical scenarios. Table 1.8 displays responses by all participants (Column 1) and participants choosing the high- and low-task jobs (Columns 2 and 3, respectively), pooling across tasks.[45]

The two most common responses aside from the job offering better pay (cited by nearly 60 percent of participants) are that the chosen job sounds "more enjoyable/interesting," and

[45] Figure A.5 shows the wording of the response options available to participants. I randomize the order of responses and allow multiple entries.

Table 1.8: *Reasons for Choices*

	All	High-Task	Low-Task
Offers better pay	0.594	0.632	0.564
More enjoyable/interesting	0.338	0.433	0.265
Better fit for skills	0.338	0.381	0.305
Develop new skills	0.177	0.281	0.097
Require less effort	0.234	0.075	0.356
More people like me	0.076	0.107	0.052
Would be treated better	0.066	0.071	0.061
More prestigious	0.054	0.074	0.039
N	3,862	1,683	2,179

Notes: This table shows the share of the experiment sample citing each of the reasons listed as a motivation for choices made in the hypothetical scenarios. Each participant was asked to indicate reasons for the choices made in scenarios relating to a randomly selected two out of the five gender-typical tasks. Column 1 shows reasons cited by all participants, Column 2 shows reasons cited bt participants who chose the high-task job in that scenario, and Column 3 shows reasons cited by participats who chose the low-task job.

that the chosen job would be a "better fit for my existing skills and abilities," each cited by 34 percent of participants. I interpret these responses to reflect choice motivations related to current preferences and skills, respectively. In addition, 18 percent of participants indicate that the chosen job would "allow me to strengthen or develop new skills," which suggests an investment motivation. The preference response has a correlation coefficient of 0.31 and 0.27 with the responses related to existing and new skills, respectively.

These results are consistent with the hypothesis that task-specific preferences and skills are correlated or even jointly determined. Workers may find it more interesting or enjoyable to perform tasks in which they have a productivity advantage, and may also invest in developing skills relevant to tasks they enjoy performing. I cannot provide insight into the process of task-specific preference and skill formation. To the extent that task-specific skills have a causal impact on preferences, I interpret this effect as a component of the preference parameter that I estimate.

Another possibility is that WTP estimates reflect career concerns related to task-specific skills. Specifically, participants may be willing to accept lower wages in a job in which they have a comparative advantage because they believe they are more likely to be promoted or less likely to be terminated in that position. In addition, participants may be willing to pay to develop competencies that they believe will lead to higher wages in the future. While career concerns are likely to be small in a hypothetical setting, I cannot rule them out. However, Tables A.14 and A.15 show that WTP results are similar when restricting the sample to participants who do not cite a better fit for existing skills or developing new skills, respectively, as a motivation for their choice.

Table 1.8 also indicates that 23 percent of participants say that the chosen job would "require less effort." This answer is much more common among those choosing a low-task job (36 percent) than among those choosing a high-task job (8 percent),[46] suggesting that participants do indeed view jobs involving more time spent on a gender-typical task as requiring greater effort. This finding provides evidence that participants are focused on the

[46] By contrast, all other response options are selected more often by participants choosing the high-task job.

difference in the amount of time spent on the focal task between the two scenarios, rather than the time spent on other activities. Therefore, it seems unlikely that gender differences in the counterfactual activities that participants envision are driving results.[47]

Finally, only between 5 and 10 percent of participants cite reasons for their choices related to identity (the chosen job would have "more people like me"), discrimination (participants would be "treated better" in the chosen job), and prestige. Therefore, it does not seem that beliefs about discrimination or gender identity are major factors affecting choices.

1.6 Implications for Gender Gaps

In this section, I examine the implications of the WTP estimates for gender gaps in the labor market. I focus on gender differences in job sorting on the five gender-typical tasks, occupational segregation, and the gender wage gap, using data from the experiment sample and the ACS.

1.6.1 Observed Sorting and Segregation

I begin by documenting observed gender differences in sorting on the five gender-typical tasks in the experiment sample and the ACS. These observed differences provide a baseline against which I compare the gender differences in sorted that are predicted by the preference estimates from the experiment.

I define the gender difference in sorting on task k to be

$$Q_k \equiv p_{fk} - p_{mk} \tag{1.7}$$

[47] As an additional robustness check, in Table A.16 I report estimates of the difference in differences in WTP across gender and tasks. If participants envision the same job involving the same set of activities for all scenarios, then the counterfactual tasks imagined will difference out in the comparison between tasks. The gender differences in WTP for all female-typical tasks compared with *operating and repairing equipment* are statistically significant, as is the difference between *helping and caring for others* and *making decisions and solving problems* and between *documenting and recording information* and *making decisions and solving problems*.

where p_{gk} is the share of workers of gender $g \in \{f, m\}$ employed in a job that involves a high level of task k.

In the experiment sample, I consider workers to have a job with a high level of task k if they report that in their current or most recent job, they spend more time on task k than the cutoff number of hours in the hypothetical scenario for task k that they are shown. In the ACS, I use the occupation-level O*NET measures to classify jobs as high task or low task, as I have no data on the frequency of task performance. Specifically, I select a cutoff percentile for each task such that the share of workers in the ACS above the cutoff matches the share of workers in a high-task job in the experiment sample, and consider occupations to be high task if they fall above this percentile.[48]

The first column of Panel A in Table 1.9 reveals substantial gender differences in sorting (Q_k) in the experiment sample, as suggested by the gender-specific distributions of time spent on tasks in Figure 1.7. Women are 15.4, 10.5 and 8.0 percentage points more likely than men to work in jobs that involve a high level of the female-typical tasks of *helping and caring for others*, *documenting and recording information*, and *working and communicating with others*, respectively. By contrast women are 18 percentage points less likely than men to spend more time *operating and repairing equipment*. The gender gap in sorting on *making decisions and solving problems* is close to zero.[49]

The first column of Panel B in Table 1.9 shows a similar pattern of results in the ACS. Women are approximately 27.0, 14.1 and 25.9 percentage points more likely than men to work in jobs that involve high levels of *helping and caring for others*, *documenting and recording information*, and *working and communicating with others*, respectively, and 30.5 percentage points less likely than men to hold a job with a high level of *operating and*

[48]The analyses in this section restrict the experiment sample to all participants with valid wages and information on race, ethnicity and education. The ACS sample used in this section includes all individuals in the 2012-2016 data who are aged 18 and older, currently employed, have valid non-zero earnings, and work in an occupation that can be matched to the O*NET. Appendix A.1 provides further detail on the O*NET measures.

[49]In fact, women are approximately 1 percentage point more likely to be in a job with a high level of *making decisions and solving problems*, contrary to the hypothesis that this activity is male-typical.

Table 1.9: *Sorting and Segregation Explained by Tasks*

(a) Experiment

	Q	I	b	$\hat{Q}$	$\hat{Q}/Q$	$\hat{I}$	$\hat{I}/I$
Help	0.154		-0.089	0.058	0.375		
Document	0.105		0.115	0.060	0.566		
Communic.	0.080		-0.014	0.027	0.332		
Operate	-0.179		0.008	-0.167	0.938		
Decisions	0.007		0.168	-0.011	-1.608		
Index		0.262				0.179	0.684
N			1,785				

(b) ACS

	Q	I	b	$\hat{Q}$	$\hat{Q}/Q$	$\hat{I}$	$\hat{I}/I$
Help	0.270		-0.191	0.024	0.089		
Document	0.141		0.133	0.054	0.380		
Communic.	0.259		0.050	0.027	0.104		
Operate	-0.305		-0.017	-0.143	0.470		
Decisions	-0.011		0.222	-0.007	0.595		
Index		0.389				0.158	0.405
N			6,419,869				

Notes: This table shows observed and predicted gender differences in sorting on the gender-typical tasks and task-based segregation in the experiment sample (Panel A) and the 2012-2016 ACS (Panel B). Q ($\hat{Q}$) is the observed (predicted) gender difference in the share of workers sorting into the high-task job. b is the coefficient on the task from the wage regression. I ($\hat{I}$) is the value of the segregation index based on observed (predicted) sorting.

repairing equipment.[50] These differences are somewhat larger in magnitude than those in the experiment sample, which is not surprising given that the O*NET measures used to construct the gender-typical tasks are selected on the basis of their robust correlation with the occupational female share.

I also calculate a task-based segregation index as in Duncan and Duncan (1955) that summarizes the gender differences in sorting on all tasks with a single statistic. Consider a set of job categories categories, indexed by $j \in \{1, \dots, J\}$, where each category is defined by having a high or low level of each gender-typical task, such that $J = 2^5$. The index is then given by

$$I \equiv \frac{1}{2} \sum_{j=1}^{J} |p_{fj} - p_{mj}| \tag{1.8}$$

where p_{gj} is the proportion of workers of gender g sorting into job category j. The index is scaled such that a value of 0 indicates gender equality, while a value of 1 reflects total segregation.

The second column of Table 1.9 reports a value of 0.262 for the task-based segregation index in the experiment sample in Panel A, and a value of 0.389 in the ACS in Panel B. By comparison, Blau *et al.* (2013) calculate a gender segregation index in the 2009 ACS using the 2000 Census occupation codes, which include approximately 500 categories, and report a value of 0.510.[51] Thus more than three quarters (76 percent) of occupational segregation by gender can be accounted for by gender differences in sorting on 32 occupation categories defined by gender-typical tasks. This comparison provides a baseline estimate of how much preferences over the tasks I examine contribute to segregation under the assumption that observed gender differences in sorting on these tasks are entirely due to preferences.

[50] Again, the gender difference in sorting on *making decisions and solving problems* is close to zero; women are 1.1 percentage points less likely to work in jobs with a high level of this task.

[51] I calculate an occupational segregation index using currently employed individuals aged 18 and older in the 2012-2016 ACS data, and find a value of 0.497 using the full set of 478 occupation codes that appear in the data and a value of 0.498 using the 464 occupation codes that can be matched to the O*NET.

1.6.2 Predicted Sorting and Segregation

Next, I calculate predicted gender differences in sorting on the gender-typical tasks, using the model and estimated preference parameters from the hypothetical choice experiment and wage differentials associated with the tasks in the experiment sample and ACS.

For the estimated preference parameters, I use coefficients from a pooled specification in which I stack the choice data from the experiment for all tasks and both genders and run a single logistic regression, clustering standard errors by participant. An observation in this regression is at the participant-by-scenario level, the outcome is an indicator for choosing the high-task job in that scenario, and the predictors are the log difference in wage offers between the high-task and low-task jobs in the scenario (ω_{ik}) and a set of indicators for the scenario relating to task k (T_{jk} for $k \in \{1, \ldots 5\}$). I allow the $\hat{\theta}_{gk}$ task coefficients to differ by gender, but estimate a single $\hat{\delta}$ coefficient for the entire sample.

Estimating a $\hat{\delta}$ coefficient that does not vary by task k allows me to calculate predicted probabilities of sorting into the job categories $j \in \{1, \ldots, 32\}$, which are necessary for calculating a predicted segregation index. Furthermore, I restrict $\hat{\delta}$ to be the same for women and men to ensure that predicted gender differences in sorting are driven only by estimated differences in preferences for tasks, rather than any gender difference in the coefficient on the wage differential.[52]

Table 1.10 shows WTP estimates based on the pooled specification. Results are very similar to the WTP estimates based on task-specific regressions and gender- and task-specific $\hat{\delta}_{gk}$ coefficients in Table 1.4.[53]

---

[52]This specification can be motivated by a modification of (1.1) in which the utility of person i of gender $g(i) \in \{f, m\}$ from job $j \in \{1, \ldots, 32\}$ related to scenario c is given by

$$U_{ijc} = \alpha_i + \sum_{k=1}^{5} \theta_{g(i)k} T_{jk} + \delta \ln w_{ijc} + \varepsilon_{ijc} \tag{1.9}$$

where w_{ijc} is the wage offer for participant i in job j related to scenario c, $\theta_{g(i)k}$ is a task- and gender-specific preference parameter, δ is a preference parameter not indexed by task or gender, ε_{ijc} is a worker- , job- and scenario-specific preference parameter with a standard EV Type I distribution, and the other terms are as defined in (1.1). Note that in contrast to (1.1), in which $j \in \{1, 2\}$, here j indexes the job categories defined in Section 1.6.1.

[53]I also find similar results when I estimate WTP based on task-specific regressions in which the $\hat{\delta}_k$ coefficient

Table 1.10: *WTP for Tasks - All Tasks Pooled, Single δ Parameter*

	Help	Document	Communic.	Operate	Decisions
Women	-0.009	-0.012*	-0.017**	-0.143**	-0.006
	(0.007)	(0.006)	(0.006)	(0.009)	(0.006)
Men	-0.039**	-0.039**	-0.027**	-0.061**	0.003
	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Diff (W-M)	0.030**	0.027**	0.010	-0.082**	-0.009
	(0.010)	(0.009)	(0.009)	(0.012)	(0.009)
N	9,655	9,655	9,655	9,655	9,655

Notes: Standard errors clustered by participant. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates from a pooled specification in which choices are stacked and all tasks are included in a single regression. The $\hat{\delta}$ coefficient on the log difference in wage offers is also restricted to be the same for women and men in this specification.

Using the $\hat{\delta}$ and $\hat{\theta}_{gk}$ coefficients from the pooled specification reported in Table 1.10 and the logistic parameterization, I calculate the share of workers of gender $g \in \{f, m\}$ who are predicted to sort into a high-task k job as a function of ω_k, the log difference in wage offers between high-task k and low-task k jobs, as in (1.4):

$$\hat{\mathrm{H}}_{gk}(\omega_k) \equiv \frac{1}{1 + \exp(-(\hat{\theta}_{gk} + \hat{\delta}\omega_k))}. \tag{1.10}$$

The predicted gender difference in sorting on task k is then

$$\hat{\mathrm{Q}}_k \equiv \hat{\mathrm{H}}_{fk}(\hat{\omega}_k) - \hat{\mathrm{H}}_{mk}(\hat{\omega}_k) \tag{1.11}$$

where $\hat{\omega}_k$ is an estimate of ω_k.

To generate $\hat{\omega}_k$, I estimate earnings equations of the form

$$\ln w_i = X_i'\beta + \xi_i \tag{1.12}$$

in the experiment sample and the ACS, where $\ln w_i$ is the log hourly wage of individual i, X_i is a vector of individual i's observable characteristics, and ξ_i is an individual-specific

differs by task but is restricted to be the same for women and men, and a specification in which all tasks are pooled but the $\hat{\delta}_g$ coefficient is allowed to differ by gender.

residual with expectation zero. The X_i vector includes a set of indicators for i's current job involving high levels of the gender-typical tasks, as well as controls for gender, race and ethnicity, education, potential experience and, in the ACS only, geography and year.[54] My estimate of ω_k is then

$$\hat{\omega}_k \equiv b_k \tag{1.13}$$

where b_k is the coefficient on the task k indicator from the OLS estimation of (1.12).

I interpret the b_k coefficients (termed *task wage differentials*) as equilibrium wage differentials associated with the gender-typical tasks. These differentials may reflect preferences for tasks as well as preferences for other job amenities correlated with tasks that affect labor supply decisions, producing compensating differences.[55] Search frictions and departures from perfect competition such as monopsony and bargaining may also contribute to the differentials.

The third column of Table 1.9 reports the task wage differentials (b_k) in the experiment sample (Panel A) and the ACS (Panel B). In both data sources, the female-typical task of *documenting and recording information* and the male-typical task of *making decisions and solving problems* are associated with substantial wage premiums (0.133 and 0.222, respectively, in the ACS), while the female-typical *helping and caring for others* offers a large wage penalty (-0.191 in the ACS). The wage differentials associated with *working and communicating with others* and *operating and repairing equipment*, by contrast, are quite small in magnitude in both samples.

---

[54] I exclude hourly wage observations that are less than $3 or greater than $200 in 2018 dollars. In the ACS, I measure hourly wages as annual earnings divided by annual hours worked, inflate wages to 2018 dollars, and weight the regression by the ACS person weight. The race and ethnicity variables consist of mutually exclusive indicators for Black, Hispanic, and other race, with White non-Hispanic as the omitted category. The education variables consist of indicators for high school diploma or less in the experiment sample (less than a high school diploma and high school diploma in the ACS), some college, associate degree and graduate degree, with bachelor's degree as the omitted category. Potential experience is measured as age minus years of education minus six, restricted to be greater than or equal to zero, and the regression also includes the square of potential experience. In the ACS only, the geography variables consist of indicators for region and for metropolitan area status.

[55] In Rosen (1986), firms have a distribution of costs associated with offering an amenity, and the equilibrium wage differential equates the share of workers willing to pay that differential with the share of firms with a cost of offering the amenity that is equal to or less than the differential. By contrast, I view tasks as inherent to production in certain firms, but an amenity from the perspective of workers.

Column 4 reveals non-trivial predicted gender differences in sorting ($\hat{Q}_k$) based on the task wage differentials and preference estimates. In the experiment sample (Panel A), women are predicted to be 5.8, 6.0 and 2.7 percentage points more likely than men to work in jobs with high levels of the female-typical tasks of *helping and caring for others*, *documenting and recording information* and *working and communicating with others*, respectively. Column 5 shows that these predicted differences can account for 38, 57 and 33 percent, respectively, of observed sorting differences. In addition, women are predicted to be 16.7 percentage points less likely than men to work in jobs with a high level of the male-typical task of *operating and repairing equipment*, accounting for 94 percent of the observed sorting difference.[56]

In the ACS (Panel B), women are predicted to be 2.4, 5.4 and 2.7 percentage points more likely than men to work in jobs with high levels of *helping and caring for others*, *documenting and recording information*, and *working and communicating with others*, respectively, accounting for 9, 38 and 10 percent of observed differences in sorting. Women are also predicted to be 14.3 percentage points less likely than men to work in jobs with a high level of *operating and repairing equipment*, accounting for 47 percent of the observed difference.[57]

The predicted gender differences in sorting on *helping and caring for others* and *documenting and recording information* are somewhat smaller in magnitude in the ACS compared with the experiment sample because the task wage differentials in the ACS are farther from the overall mean WTP, where the gender difference in sorting is approximately maximized. Combined with the fact that observed sorting differences are larger in the ACS, these smaller predicted sorting differences result in preferences accounting for a lower share of actual sorting in the ACS compared with the experiment sample. It is not surprising that the WTP estimates are more predictive of sorting behavior in the sample used to elicit preferences than in the overall population.

[56]Women are also predicted to be 1.1 percentage points less likely to work in jobs with a high level of *making decisions and solving problems*, although they are slightly more likely to report actually working in jobs involving high levels of this activity.

[57]Finally, women are predicted to be 0.7 percentage points less likely to work in jobs with a high level of *making decisions and solving problems*, explaining 60 percent of the small observed gender difference.

I also calculate a value for the task-based segregation index, using the estimated preference parameters to predict gender differences in sorting into the task-based job categories. Specifically, I assume that worker i of gender g's choice of job category has a multinomial logistic distribution, and that the probability of sorting into category j can be expressed as a function of ω_j, the log wage in job category j, as follows:

$$\hat{h}_{gj}(\omega_j) \equiv \frac{\exp(\hat{V}_{gj})}{\sum_{j'=1}^{J} \exp(\hat{V}_{gj'})} \tag{1.14}$$

where $\hat{V}_{gj} = \sum_{k=1}^{5} \hat{\theta}_{gk} T_{jk} + \hat{\delta}\omega_j$.[58] I estimate ω_j as the sum of the wage differentials for tasks that have a high level in job category j:

$$\hat{\omega}_j = \sum_{k=1}^{5} b_k T_{jk}. \tag{1.16}$$

Thus the value of the task-based segregation index based on predicted sorting is given by

$$\hat{I} = \frac{1}{2} \sum_{j=1}^{J} |\hat{h}_{fj}(\hat{\omega}_j) - \hat{h}_{mj}(\hat{\omega}_j)|. \tag{1.17}$$

The predicted segregation index has a value of 0.179 using wage differentials from the experiment sample, which is 68 percent of the magnitude of the observed segregation index calculated in that sample. In the ACS, the predicted segregation index has a value of 0.158, which is approximately 41 percent of the index value of 0.389 based on observed sorting and about 31 percent of the index value reported by Blau *et al.* (2013). This result suggests that preferences for the gender-typical tasks I examine can account for a degree of segregation that is nearly a third of overall occupational segregation by gender in the U.S. labor market.

[58] The multinomial logistic parameterization can be motivated by a model in which the utility of person i of gender $g(i) \in \{f, m\}$ from job category $j \in \{1, \ldots, 32\}$ is given by

$$U_{ij} = \alpha_i + \sum_{k=1}^{5} \theta_{g(i)k} T_{jk} + \delta\omega_j + \varepsilon_{ij} \tag{1.15}$$

where ε_{ij} is a worker- and job-specific preference parameter with a standard EV Type I distribution. See Cameron and Trivedi (2005) for a derivation. This model differs from (1.9) in that ε_{ij} is indexed only by i and job category j, and not by scenario c. However, note that the predicted probability of sorting into a high-task k job ($\hat{Q}_k$) given by (1.10) and (1.11) is equal to the probability calculated by summing $\hat{h}_{gj}(\hat{\omega}_j)$ over all job categories j that have a high level of task k.

1.6.3 Robustness Checks for Sorting and Segregation

The results in Table 1.9 suggest that gender differences in sorting on the gender-typical tasks are substantial, both in the experiment sample and the ACS, and that preferences for these tasks can account for a meaningful share of the observed sorting differences and overall occupational segregation by gender.

This qualitative pattern of results is robust to a variety of alternative specifications, including restricting the sample by age range, restricting the sample to full-time workers, repeating the analysis separately for individuals with and without a college degree, and adding controls for major industry category into the wage equation in (1.12) for the ACS sample.[59] Findings are also comparable when I define the indicators for participants being in a high-task job using the extensive or intensive margin cutoffs, rather than the cutoff that each participant faces.[60]

A potential limitation of the ACS is that hourly wages must be constructed by dividing annual earnings by annual hours worked, increasing the likelihood of measurement error.[61] Therefore, I repeat the analysis using the Current Population Survey Merged Outgoing Rotation Group (CPS MORG) data, which includes information on hourly wages for workers paid by the hour, and weekly earnings for others. Panel A of Table A.17 documents patterns of gender differences in observed and predicted sorting in the CPS MORG data that are similar to results in the ACS.[62]

[59]Specifically, results in the ACS are very similar when restricting to workers aged 18 to 64 and excluding those who are self-employed, in institutional group quarters, or in the military, and when restricting to workers aged 25-64 who are employed full-time, defined as working at least 35 hours per week and 27 weeks per year. Results are qualitatively similar for college and non-college workers, but the magnitude of the observed and predicted gender differences in sorting are much larger for the less educated group, especially in the ACS, consistent with the fact that occupational segregation is greater among non-college workers (Blau *et al.* 2013).

[60]In these versions of the analysis, I also use preference parameters estimated based on the extensive or intensive margin sub-samples, respectively.

[61]Furthermore, annual weeks worked is constructed by multiplying usual hours of work per week by weeks worked in the previous year. Beginning in 2008, ACS respondents report weeks worked within intervals; therefore, the precise number of weeks worked is imputed using the 2005-2007 ACS, introducing additional measurement error and potential for bias.

[62]See Appendix A.1 for further details on sample and variable construction in the CPS MORG data.

One concern is that the task wage differentials may reflect preferences for other job amenities that are systematically bundled with the gender-typical tasks.[63] Panel B of Table A.17 repeats the analysis in the ACS including a set of measures for amenities that the literature suggests may be differentially valued by women compared with men.[64] The b_k coefficients do indeed decrease in magnitude, but results on predicted sorting and segregation are similar to the baseline ACS specification in Table 1.9.[65]

Another concern is that the wage differentials may be correlated with unobserved human capital. A large literature on compensating differentials has noted that estimates tend to be "wrong-signed," i.e. amenities that are expected to be considered desirable by most or all workers are associated with wage premiums, suggesting that individuals in jobs with these attributes have unobservably higher skills (e.g. Brown 1980). Tasks are not a clear amenity or disamenity, so ability bias may be less of a concern in this context. Nevertheless, to investigate this possibility I repeat the analysis using data from the Panel Study of Income Dynamics (PSID), which includes information on actual work experience, and the National Longitudinal Surveys of Youth 1979 and 1997 (NLSY79 and NLSY97), which can be used to construct measures of cognitive, non-cognitive and social skills.[66] Panels C, D and E of

[63]The fact that the tasks may be correlated with other job amenities does not necessarily imply that the task wage differentials are biased, but it does suggest that they cannot be interpreted as compensating differentials for the tasks alone, and may also explain the disparity between predicted and observed sorting levels for the overall sample. For example, 52 percent of workers in the experiment sample report working in jobs with a high level of *helping and caring for others*, although this task is associated with a wage penalty of -0.089 and the WTP estimates imply that workers must be compensated an additional 2.6 percent, on average, to be willing to spend more time on this activity. One potential explanation for this finding is that jobs with a high level of *helping and caring for others* have other amenities that workers value.

[64]Specifically, the specification includes mean log hours worked per week by occupation, which Denning *et al.* (2019) find can account for a large portion of the gender wage gap, a set of five O*NET measures that Goldin (2014) uses as a proxy for time flexibility (the work activity *establishing and maintaining interpersonal relationships* and the work context variables *time pressure*, *contact with others*, *structured versus unstructured work*, and *freedom to make decisions*), and the O*NET work context variable *level of competition*. The O*NET measures are standardized to have a mean of zero and standard deviation of one.

[65]Indeed, the predicted gender differences in sorting on *helping and caring for others* and *documenting and recording information* are larger compared with the baseline specification.

[66]I follow Blau and Kahn (2017) in constructing a measure of actual experience in the PSID. In the NLSY79 and NLSY97, I use scores on the Armed Forces Qualifying Test (AFQT) to capture cognitive skill, and follow the methodology of Deming (2017) to create measures for non-cognitive and social skill. See Appendix A.1 for further details on sample and variable construction in the PSID, NLSY79 and NLSY97.

Table A.17 show that patterns of observed and predicted gender differences in sorting and segregation in the PSID, NLSY79 and NLSY97, respectively, are qualitatively similar to those in the ACS and experiment sample.

Thus incorporating additional measures of general human capital into the analysis does not appear to meaningfully change the pattern of findings. I cannot rule out bias from task-specific human capital.[67] However, it is important to note that the model predicts some gender difference in sorting regardless of the exact value of the estimate for ω_k. Figure A.6 plots predicted gender differences in sorting by task as a function of ω_k, using the $\hat{\delta}$ and $\hat{\theta}_{gk}$ coefficients.[68] It is clear that the estimated preference parameters imply meaningful sorting differences for the three tasks with significant gender differences in WTP for a wide range of values of ω_k.

1.6.4 Gender Wage Gap

Lastly, I assess the contribution of observed and predicted gender differences in sorting to the gender wage gap. The female-to-male log hourly wage gap explained by observed gender differences in sorting on task k is given by

$$\mathbb{WG}_{expl,k} \equiv Q_k b_k. \tag{1.18}$$

The total wage gap explained by the five gender-typical tasks is the sum of the task-specific gaps:

$$\mathbb{WG}_{expl,tot} \equiv \sum_{k=1}^{5} Q_k b_k. \tag{1.19}$$

[67]Workers can generally be expected to sort into jobs based on comparative advantage in performing the activities required, as in Roy (1951). As discussed above, task-specific preferences and skills are likely to be correlated, and workers sorting into a high-task k job may therefore have both a high WTP and a comparative aptitude for task k. Without further assumptions, however, it is not clear how task-based sorting on comparative advantage is likely to affect the task wage differentials.

[68]The vertical lines show the estimated task wage differentials ($\hat{\omega}_k = b_k$) in the ACS from Table 1.9.

The task-specific and total gender wage gaps that can be explained by predicted gender differences in sorting simply replace Q_k in (1.18) and (1.19) with $\hat{Q}_k$, as follows:

$$\hat{\mathbb{WG}}_{expl,k} \equiv \hat{Q}_k b_k \tag{1.20}$$

$$\hat{\mathbb{WG}}_{expl,tot} \equiv \sum_{k=1}^{5} \hat{Q}_k b_k. \tag{1.21}$$

For female-typical tasks, I expect that $Q_k > 0$, so that a wage premium associated with task k ($b_k > 0$) explains a positive differential that serves to narrow the female-to-male wage gap, while a wage penalty ($b_k < 0$) explains a negative differential that serves to widen the wage gap. Similarly, for male-typical tasks, I expect that $Q_k < 0$, so that $b_k < 0$ ($b_k > 0$) explains a positive (negative) wage gap. Note that if the sign of $Q_k b_k$ differs across tasks, these gaps may partially or fully offset each other.

Table 1.11 displays results on the gender wage gap in the experiment sample (Panel A) and the ACS (Panel B). The first three columns repeat the estimates of observed gender differences in sorting (Q_k), the task wage differentials (b_k) and predicted gender differences in sorting ($\hat{Q}_k$), for reference. Column 4 shows that some of the task-specific wage gaps based on observed sorting ($Q_k b_k$) are non-trivial in magnitude. In particular, the substantial wage penalty associated with *helping and caring for others* implies that this task can explain a negative wage gap of −0.014 in the experiment sample and −0.052 in the ACS. By contrast, the wage premium associated with *documenting and recording information* implies that this task explains a positive, and thus "wrong-signed," wage gap of 0.012 in the experiment sample and 0.019 in the ACS.[69] However, this positive and negative gap largely cancel out, leading to a total explained wage gap that is close to zero (−0.003) in the experiment sample and negative but modest in size (−0.017) in the ACS. By contrast, I document sizable raw wage gaps of −0.199 and −0.183 in the experiment sample and the ACS, respectively. Column 5 reveals a similar pattern based on predicted gender differences in sorting. The

[69]In both samples, the fact that the task wage differentials for *working and communicating with others* and *operating and repairing equipment* are quite small in magnitude means that these tasks contribute relatively little to the wage gap despite substantial gender differences in sorting. Conversely, *making decisions and solving problems* does not contribute meaningfully to the gender wage gap because sorting differences are small, although this task provides a large wage premium.

Table 1.11: *Wage Gap Explained by Tasks*

(a) Experiment

	Q	b	$\hat{Q}$	Qb	$\hat{Q}b$
Help	0.154	-0.089	0.058	-0.014	-0.005
Document	0.105	0.115	0.060	0.012	0.007
Communic.	0.080	-0.014	0.027	-0.001	-0.000
Operate	-0.179	0.008	-0.167	-0.001	-0.001
Decisions	0.007	0.168	-0.011	0.001	-0.002
Total Explained by Tasks				-0.003	-0.002
Total Wage Gap		-0.199			
N		1,785			

(b) ACS

	Q	b	$\hat{Q}$	Qb	$\hat{Q}b$
Help	0.270	-0.191	0.024	-0.052	-0.005
Document	0.141	0.133	0.054	0.019	0.007
Communic.	0.259	0.050	0.027	0.013	0.001
Operate	-0.305	-0.017	-0.143	0.005	0.002
Decisions	-0.011	0.222	-0.007	-0.003	-0.001
Total Explained by Tasks				-0.017	0.005
Total Wage Gap		-0.183			
N		6,419,869			

Notes: This table shows the gender wage gaps explained by observed and predicted gender differences in sorting on the gender-typical tasks in the experiment sample (Panel A) and the 2012-2016 ACS (Panel B). Qb ($\hat{Q}b$) is the female-to-male log hourly wage gap explained by the observed (predicted) gender difference in sorting on the task.

estimated preference differences over the tasks I examine can explain gender wage gaps of close to zero (-0.002 in the experiment sample and 0.005 in the ACS).

These findings suggest that preferences for job tasks do not contribute meaningfully to the gender wage gap. However, I only examine five task categories, and given that certain tasks tend to widen the gap while others tend to narrow it, the sum of the task-specific gaps is inherently sensitive to the choices of which activities to include in the analysis. Results are also likely to depend on the precise values of the estimated task wage differentials and how tasks are measured. In particular, I elicit preferences over the binary options of spending more versus less time on a task, and it may be that valuations of time spent on tasks measured more finely may contribute more to the gender wage gap.

To illustrate this possibility, I estimate a series of decompositions of the gender wage gap, as follows:

$$\bar{w}_f - \bar{w}_m = \underbrace{(\bar{X}_f - \bar{X}_m)'b}_{Explained} + \underbrace{b_{fem}}_{Residual} \tag{1.22}$$

where $\bar{w}_g$ is the mean log hourly wage for workers of gender $g \in \{f, m\}$, $\bar{X}_g$ is a vector of the mean levels for gender g of the observable characteristics in X_i from (1.12), with the exception of the control for gender, b is a vector of coefficients on X_i from the OLS estimation of (1.12), and b_{fem} is the coefficient on the female indicator from this regression. The first component of the gap is the portion *explained* by gender differences in mean observable characteristics, sometimes called the endowments effect, while the second component is a *residual* portion.[70]

Table 1.12 displays the results of this decomposition in the ACS. Column 1 reports the total, residual and explained wage gap based on a vector of observable characteristics that includes race and ethnicity, education, potential experience, geography, and year. These controls explain a positive female-to-male wage gap of 0.033, despite the total raw wage gap of -0.183, implying a residual gap of -0.216. In other words, women are predicted to earn about 3 percent more than men on the basis of their level of education and other basic observable characteristics. The specification in Column 2 adds the full set of occupation codes available in the ACS into the X_i vector.[71] The occupation indicators explain a gender wage gap of -0.056, about 31 percent of the total gap, and reduce the residual gap by 34 percent, to -0.142, compared with Column 1. It is clear that choice of occupation plays a meaningful role in the gender wage gap.

[70]Note that (1.22) is a version of a Oaxaca-Blinder decomposition (Blinder 1973; Oaxaca 1973) where the coefficients used to weight the difference in observable characteristics are from a pooled regression, rather than a specification including only women or only men, as is commonly the case.

[71]There are 464 codes in the 2012-2016 ACS that match to the O*NET and appear in the analysis sample for the decomposition.

Table 1.12: *Wage Gap Decomposition in ACS*

	(1)	(2)	(3)	(4)
Summary:				
Total Wage Gap	-0.183	-0.183	-0.183	-0.183
Residual Gap	-0.216	-0.142	-0.170	-0.190
Explained Gap	0.033	-0.041	-0.014	0.007
Explained By:				
Occupation		-0.056		
Continuous Tasks			-0.034	
Binary Tasks				-0.017
N	6,419,869	6,419,869	6,419,869	6,419,869

Notes: This table shows the results of a decomposition of the female-to-male log hourly wage gap in the 2012-2016 ACS. All columns include controls for race and ethnicity, education, potential experience, geography, and year. Column 2 also includes occupation indicators. Column 3 includes the gender-typical tasks measured as continuous variables standardized to have a mean of zero and standard deviation of one. Column 4 includes the binary measures of the gender-typical tasks that I use in Tables 1.9 and 1.11.

In Column 3, I replace the occupation indicators with the five gender-typical tasks, measured as continuous variables standardized to have a mean of zero and standard deviation of one. These continuous task variables can explain a gender wage gap of -0.034, about 19 percent of the raw gap, and reduce the residual gap by about 21 percent compared with Column 1. Finally, Column 4 includes the binary task measures that I use in the sorting and wage gap analyses in Tables 1.9 and 1.11. As already noted, these indicators can explain a gender wage gap of -0.017, approximately 9 percent of the total gap, and they reduce the residual portion of the gap by approximately 12 percent, from -0.216 to -0.190.

Thus while the continuous task measures capture more than 60 percent of the explanatory power of the full set of occupation controls, the binary variables capture only 30 percent. These decomposition results suggest that the contribution of preferences for job tasks to the gender wage gap remains an open question.

1.7 Conclusion

This paper examines gender differences in preferences for job tasks and the implications for gender gaps in the labor market. I estimate WTP for a set of gender-typical tasks performed more frequently by one gender than the other, using a hypothetical choice experiment embedded in a survey. I find that women have a significantly higher WTP than men for the female-typical tasks of *helping and caring for others* and *documenting and recording information*, and a significantly lower WTP for the male-typical task of *operating and repairing equipment*. WTP is higher for participants who currently spend more time on the focal tasks, suggesting that the estimates are correlated with labor market outcomes.

I find that the estimated preference parameters can explain a meaningful portion of observed gender differences in sorting on the gender-typical tasks, in both the experiment sample and the ACS. Indeed, gender differences in WTP for tasks can account for approximately a third of overall occupational segregation by gender in the ACS. By contrast, preferences for the gender-typical tasks I examine appear to explain little of the gender wage gap, but this finding may be sensitive to which tasks are selected and how they are measured.

These results suggest that gender differences in valuations of work activities have a role to play in explaining sorting patterns in the labor market, as hypothesized. Task-specific skills and discrimination may also be important, however, and I am not able to examine how preference formation may be related to these other factors. Further research is needed to understand the extent to which current task valuations are caused by inherent gender differences in preferences rather than past resource constraints and gender norms.

Chapter 2

Why Aren't Workers Paid Their Marginal Product? Unobserved Skills vs. Overconfidence[1]

Chapter Abstract

How do firms interpret a worker's wage bid when making hiring decisions? When worker skills are not perfectly observed, employers must form expectations about productivity based on the information available, including the worker's proposed wage. If these beliefs are accurate, then firms will pay workers their marginal product in efficiency units and a higher-wage worker will complete a discrete project more quickly, leaving the total wage bill unchanged. We simply test this prediction by exploiting an institutional feature of an online labor market that creates quasi-random variation in which workers are recommended to employers. We show that the recommendation system arbitrarily induces firms to hire workers who make different wages bids but score similarly on a measure of the likelihood that they will be hired, based on historical data from the platform. Higher-wage employees do work fewer hours, as expected, but increase the total wage bill, suggesting that these

[1]Co-authored with John Horton.

individuals may be systematically overvalued.

2.1 Introduction

What happens if a firm is induced to hire a higher-wage worker? A core feature of the competitive labor market model is that workers are paid their marginal product. The model predicts that a more productive worker will be compensated precisely for their additional productivity, such that a higher-wage worker completes a fixed amount of work in fewer hours, leaving the total wage bill for that work unchanged. This framework implies that firms should be indifferent over applicants that differ in their productivity—or equivalently, that differ in terms of their wage. In most labor market settings, however, firms cannot perfectly observe worker skills, and must make inferences about productivity based on the information available, including the worker's proposed wage. Are these these assessments about productivity are accurate, and if not, are higher-wage workers systematically undervalue or overvalued? The answer to this question has important implications for theories of wage determination.

This paper explores the prediction that the hourly wage is a sufficient statistic for a worker's marginal productivity. We investigate what happens when firms are quasi-randomly induced to hire a worker with a higher or lower hourly wage. Our setting is an online labor market in which employment relationships are short-term and each job can be interpreted as a discrete project. In this context, we observe whether higher-wage workers complete a fixed amount of work more quickly. By contrast, hours of work do not provide direct evidence on productivity in typical open-ended employment relationships, in which workers are assigned a continuous flow of projects. Furthermore, several factors that theory suggests may lead workers to be paid more or less than their marginal product are unlikely to be relevant in our setting. The short-term nature of employment implies that relational contracting (Macneil, 1974) and on-the-job training (Mincer, 1974) are rare or absent, as are institutional constraints such as unions (Freeman and Medoff, 1984). Because work is performed remotely, wages are unlikely to reflect compensating differentials for pleasant or unpleasant working conditions (Rosen, 1986).

In this labor market, employers post a job opening and workers submit applications for the position, including a wage bid. Once a worker is hired, the platform monitors the number of hours that the individual works and processes wage payments. Thus hours of work and wages are measured virtually without error. In addition, employers can view a potential worker's full employment history on the platform, including past wage rates and feedback from former employers. Firms therefore have comparatively rich data that they can use to estimate the productivity of job applicants.

As in other settings, naive estimates of the elasticity of hours and earnings with respect to the wage may be biased in our context due to unobserved heterogeneity across firms and workers. In particular, employers posting openings for larger projects may attract more highly skilled applicants, leading to a positive correlation between hours of work and hourly wages that may attenuate or even reverse the sign of the negative relationship predicted by theory.

Our empirical strategy overcomes this challenge by exploiting the fact that the platform makes recommendations to employers about which workers in their applicant pool they should hire. These recommendations depend on a continuous score that the platform assigns to each applicant. The score represents the platform's prediction of the probability that the employer will hire the worker based on all the information about an applicant that is observable to the platform, including past employer decisions and the wage bid for the current position. The score is calculated using a machine learning model trained on historical data from the platform. Critically for our purposes, the platform recommends a worker if their score is above a threshold and does not if they are below. This institutional feature creates the conditions for a regression discontinuity design.

We show that these platform recommendations matter. In a tight bandwidth around the score cutoff, receiving a recommendation increases the probability that a worker is hired by more than 40 percent. However, average worker attributes, including wage bids and rank in the applicant pool, are continuous through the recommendation cutoff. As workers are not aware of their score or their recommendation status, they cannot adjust their score or condition their wage bid on that score.

We use the recommendation discontinuity to examine the impact on hours worked and the wage bill when employers are encouraged to hire a worker with a higher wage bid, but a similar score. We interpret the score as the platform's prediction of the applicant's value to the employer, which we define as the log ratio of estimated productivity to the worker's wage bid. We assume that conditional on the score, any remaining variation in individual worker characteristics that are inputs into the score—including the wage bid—is not informative about the probability that a worker will be hired. In other words, if two applicants for a job both have scores that are close to the discontinuity, then in expectation employers believe these applicants to be equivalent in terms of their value.

This assumption implies that the distribution of wage bids among workers in the bandwidth varies quasi-randomly across openings. Because employers are more likely to hire an applicant who is recommended, higher wages among recommended applicants will lead to higher worker wages, holding constant expected worker value. Therefore, we use the difference between mean wage bids above versus below the cutoff as an instrument for the hired worker's realized wage.[2]

The competitive market model predicts that employers who are induced to hire a higher-wage but similar-score worker will find that the worker completes the project more quickly, leaving the total wage bill unchanged. Our results are not consistent with these sharp predictions. Higher-wage workers do indeed work fewer hours, but the decrease is much smaller in magnitude than the elasticity of negative one predicted by the model. Our estimates suggest that a one percent increase in the hired worker's hourly wage is associated with a decrease of less than half a percentage point in hours worked. In addition, the wage bill increases substantially when firms hire more expensive workers, contrary to the prediction that overall costs should stay the same.

[2]Note that our approach differs from the literature on gift exchange, as our strategy induces employers to hire workers who make different wage bids but have similar expected productivity relative to the wage, rather than randomly assigning wages irrespective of productivity (Gneezy and List, 2006).

We explore several potential explanations for these results. First, higher-wage workers may be more productive in a way that justifies their higher cost on dimensions other than hours worked, such as the quality of output or the likelihood of completing a project. Similarly, it may be that employers initially undervalue higher-wage workers, and decide to expand the scope of the project to involve more hours of work when they are pleasantly surprised. Both of these hypotheses imply that employers should be more satisfied when they are induced to hire higher-wage workers.

By contrast, employers may overvalue higher-wage workers, perhaps because they believe applicant wage bids to provide additional information about productivity even after conditioning on other observable characteristics. Similarly, the score may be a poor approximation of firm valuations, such that employers are being induced to hire workers that they would otherwise judge to be overpriced. In either of these cases, employers are likely to be less satisfied when the platform recommendation convinces them to hire a higher-wage applicant.

2.1.1 Related Literature

This paper contributes to a large literature on wage determination and the relationship between pay and productivity. Especially relevant to this project is research that examines the variance of pay relative to skills. Romer (1992) argues that firms compress wages due to worker concerns about fairness and relative pay, implying that higher ability workers are a better bargain and paying lower ability workers less would harm productivity.[3] This hypothesis cannot explain our results, which suggest that wages may be more dispersed than ability.[4] Our findings are more consistent with Terviö (2008), who contends that it may be optimal for pay to have a higher variance than underlying ability if project or firm size is

[3]Other studies that propose that wages may be less dispersed than marginal productivity because of fairness and relative status considerations include Frank (1984), Lazear (1989), and Akerlof and Yellen (1990).

[4]Indeed, fairness concerns are more likely to be relevant within firms among workers with traditional in-person employment relationships who have opportunities to interact and share information about pay.

highly variable.[5] As discussed in Section 2.7, higher wage workers may be more likely to complete a project, and a high probability of completion or high quality of output may be more important to some firms or projects than to others.

An extensive body of work suggests other reasons that workers may not be paid their marginal product at a point in time, including monopsony (Manning, 2003), search frictions and bargaining (Mortensen, 2011), on-the-job-training (Mincer, 1974), unions (Freeman and Medoff, 1984), employer learning (Altonji and Pierret, 2001; Lange, 2007), compensating differentials (Rosen, 1986), relational contracting (Macneil, 1974), and efficiency wages (Krueger and Summers, 1988).[6] However, while these theories predict that workers may be paid more or less than their marginal product, they generally do not explain why wages would be more dispersed than ability within an employer's applicant pool, as we find.[7]

Lastly, this project builds on research that examines how information affects employer decisions and worker outcomes in an online labor market. Pallais (2014) shows that providing detailed employer feedback increases the probability that inexperienced workers in an online labor market will be subsequently hired by other employers, consistent with Terviö (2009). Barach *et al.* (forthcoming) document that employers are more likely to hire a worker who is recommended by a labor market platform, even if no guarantee of quality is offered, while Horton (2017) demonstrates that platform recommendations can increase the share of vacancies that are filled. This previous work motivates our strategy of using platform recommendations as an instrument for the probability that a worker is hired and the wage of the hired worker.

The remainder of the paper proceeds as follows. Section 2.2 briefly describes the online labor market that is our empirical context. Sections 2.3 and 2.4 lay out the model and

[5] See Rosen (1981) for a similar argument and assignment model approach.

[6] Models that incorporate search frictions, in particular, do not necessarily imply a departure from a competitive equilibrium, but rather suggest that wages reflect the distribution of worker and firm costs and outside options, not merely productivity conditional on employment.

[7] Indeed, Zoega and Booth (2005) find that monopsony lowers wages for high-ability workers more than for low-ability workers, leading to wage compression. In addition, Hoxby and Leigh (2004) report evidence that unionization compresses teacher pay relative to ability.

empirical strategy, respectively. Section 2.5 describes the data, while Section 2.6 reports our results. Finally, Section 2.7 discusses potential explanations for our findings and concludes.

2.2 Empirical Context

The setting for our analysis is a large online labor market. In online labor markets, employers hire workers to perform tasks that can be done remotely. Services provided by the platform include soliciting and disseminating job openings, hosting user profile pages, processing payments, arbitrating disputes, certifying worker skills, and maintaining a reputation system (Horton, 2010; Filippas *et al.*, 2020).

On the platform, employers write job descriptions, categorize the nature of the work and required skills and then post the opening to the platform website. The labor market is designed to connect employers with freelancers or independent contractors, and thus jobs are generally short-term and project-based. In theory, employers posting job openings on the platform we study may be simultaneously advertising these openings in other online labor markets and the traditional offline market. However, survey evidence suggests that online and offline hiring are weak substitutes and that multi-homing of job openings is rare.[8]

Workers construct profiles similar to resumes. These profiles contain a rich set of information about the worker, including details of past jobs, education history, and skills. Importantly, the platform verifies much of the data on worker profiles, including hours of work, hourly wage rates, total earnings and employer ratings from past contracts.

Workers learn about job openings through electronic searches or email notifications. In addition, employers can search worker profiles and invite workers to apply (Horton, 2017). Workers submit job applications, which generally include a cover letter and a wage bid

[8]When asked what they would have done with their most recent project if the platform were not available, only 15 percent of employers respond that they would have made a local hire. Online employers report that they are generally deciding among (a) getting the work done online, (b) doing the work themselves and (c) not having the work done at all. The survey also finds that 83 percent of employers listed their last job opening only on the platform in question.

for hourly jobs or total project bid for fixed-price jobs. While in other settings it is not the norm for workers to propose a wage when applying for a job, this process is typical for independent contractors.

After a worker submits an application, the employer can interview the applicant. Employers can then hire an applicant at the terms proposed in their application, or make a counteroffer, which the worker can then counter. Back-and-forth bargaining between worker and employer is fairly rare, as approximately 90 percent of hired workers receive the wage they initially proposed (Barach and Horton, 2020). The process is not an auction and neither the employer nor the worker is bound to accept any offer.

The platform facilitates the matching process between employers and workers by recommending some subset of the applicants to a job opening Barach *et al.* (forthcoming). These recommendations are based on a score assigned to each applicant using a machine learning model trained on historical data from the platform. Specifically, the platform recommends a worker to the employer if their score is above a threshold and does not if they are below.

Once a worker is hired, the platform gives employers the ability to precisely monitor the number of hours that a worker is working. In order to work on hourly contracts, workers must install custom tracking software on their computers that serves as a digital punch clock. When the worker is working, the software records the count of keystrokes and mouse movements, and captures an image of the worker's computer screen at random intervals. These data are sent to the platform's servers and made available to the employer for inspection in real time. In addition, the platform processes worker payments, meaning that hours of work and hourly wages are measured virtually without error.

2.3 Model

2.3.1 Observable Productivity

Consider a labor market with many workers i and employers j. Each employer has a job opening for a project of an exogenously-determined size Y_j that it can sell in the product market for price p, with labor as the sole factor of production.[9] Each worker has a publicly observable, exogenously-determined productivity level of y_i.[10] Firms post openings for their job, and worker i applies to employer j with an hourly wage bid of w_{ij}.

Firms choose a worker to maximize profits:

$$\arg\max_i \; pY_j - \frac{Y_j}{y_i} w_{ij}. \tag{2.1}$$

Denote $R_{ij} \equiv \frac{w_{ij}}{y_i}$ and $r_{ij} \equiv \ln R_{ij} = \ln w_{ij} - \ln y_i$. Then firms select the worker in their applicant pool with the minimal r_{ij}. Equivalently, we can define the value of worker i to the firm to be

$$u(y_i, w_{ij}) \equiv \ln y_i - \ln w_{ij}, \tag{2.2}$$

such that firms select a worker to maximize $u(\cdot,\cdot)$.

In equilibrium in the absence of search frictions, all firms choose a worker with a wage bid that satisfies the first order condition of profits with respect to Y_j,

$$p = \frac{w_{ij}}{y_i}. \tag{2.3}$$

Normalize the output price to $p = 1$. Then worker i who is hired by employer j is paid

$$w_{ij}^h = y_i, \tag{2.4}$$

their marginal product.

This simple model generates two empirical predictions:

Prediction 1. The number of hours worked by worker i hired by firm j, $h_{ij} \equiv \frac{Y_j}{y_i}$, will be decreasing in the wage w_{ij}^h, with an elasticity of negative one.

[9] In our data, employers may post multiple jobs; we assume here for simplicity that each employer has a single job opening.

[10] For simplicity, we assume here that worker productivity is constant across employers and jobs. The model generates identical predictions if we allow productivity to vary by employer, although in that case the worker's equilibrium wage will also differ by employer.

Proof. The elasticity of hours with respect to wages, ϵ_{ij}, is

$$\epsilon_{ij} = \frac{\partial \ln h_{ij}}{\partial \ln w^h_{ij}} = \frac{\partial [\ln Y_j - \ln w^h_{ij}]}{\partial \ln w^h_{ij}} = -1. \tag{2.5}$$

Prediction 2. The wage bill for worker i hired by firm j, C_{ij}, is constant, with an elasticity of zero.

Proof. Equation 2.4 implies that the ratio R_{ij} is a constant: $R_{ij} = \frac{w^h_{ij}}{y_i} = 1$. Therefore, the wage bill of firm j is

$$C_{ij} = Y_j R_{ij} = Y_j, \tag{2.6}$$

regardless of which worker i is hired.

2.3.2 Unobservable Productivity

Now consider the case where worker productivity, y_i, is unobservable, as in most real-world labor markets. However, firm j can observe an n-dimensional vector of worker attributes, $X_{ij} \in R^n$, that includes the worker's wage bid. We can express X_{ij} as $X_{ij} = [Z_i', w_{ij}]'$, where Z_i is an $(n-1)$-dimensional vector of worker attributes that excludes the worker's wage bid.

Firms now select workers by forming an estimate of the productivity of each worker in their applicant pool. We assume that the firm's estimate of the worker's productivity depends only on Z_i, and is not affected by the worker's wage bid.

Assumption 1. Firm j's estimate of worker i's log productivity, $\ln y_i$, can be expressed as $\hat{y}_j(X_{ij}) = \hat{y}_j(Z_i, w_{ij}) = \hat{y}_j(Z_i)$.

In other words, firms do not believe that the wage bid provides additional information about worker skills, after conditioning on other observable characteristics. This assumption seems plausible in the online labor market setting that we study, where firms have access to a rich set of information about potential workers including past wage rates.

Define firm j's estimate of the value it will gain from hiring worker i to be $\hat{u}_j(X_{ij})$. Using the definition of worker value in Equation 2.2, Assumption 1 implies that

$$\hat{u}_j(X_{ij}) = \hat{y}_j(Z_i) - \ln w_{ij}. \tag{2.7}$$

Therefore, if firm j is indifferent between two workers i and i' such that $\hat{u}_j(X_{ij}) = \hat{u}_j(X_{i'j})$, and worker i makes a higher wage bid ($w_{ij} > w_{i'j}$), then worker i must also have higher expected productivity ($\hat{y}_j(Z_i) > \hat{y}_j(Z_{i'})$).

The firm hires the worker in its applicant pool that maximizes $\hat{u}_j(\cdot)$, or may choose not to hire at all. We further assume that $\hat{y}_j(Z_i)$ is an unbiased estimate of the true expectation of productivity conditional on Z_i.

Assumption 2. $\mathrm{E}[\hat{y}_j(Z_i)] = \mathrm{E}[\ln y_i | Z_i]$.

Then in the absence of search and informational frictions, workers are paid their expected marginal product conditional on Z_i, such that $\ln w^h_{ij} = \mathrm{E}[\ln y_i | Z_i]$, and Predictions 1 and 2 hold.[11]

2.4 Empirical Strategy

The goal of the empirical strategy is to test the predictions of the model in Section 2.3 about the elasticity of hours worked and the wage bill with respect to wages, using data from the online labor market. Ordinary least squares (OLS) estimates of the elasticities, however, may be biased due to unobservable characteristics of job openings and applicants. Therefore, we use an institutional feature of the labor market that introduces quasi-random variation into which workers are hired, among those with similar predicted value to the employer but different wage bids.

[11]To see that this is true, define $\ln y_i = \mathrm{E}[\ln y_i | Z_i] + v_i$, and note that $\ln w^h_{ij} = \mathrm{E}[\ln y_i | Z_i]$ is orthogonal to v_i by construction.

2.4.1 Platform Score

Let $s(X_{ij})$ be a score that constitutes a continuous mapping from X_{ij} to $[0,1]$ created by a platform that acts as a labor market intermediary.[12] We assume that if the platform were trying to predict the firm's estimate of a worker's value, $\hat{u}_j(X_{ij})$, then conditional on the score, no additional information is provided by any component of the observed vector of worker characteristics.

Assumption 3. $\mathrm{E}[\hat{u}_j(X_{ij})|s(X_{ij}), x_{ijk}] = \mathrm{E}[\hat{u}_j(X_{ij})|s(X_{ij})],\ \forall x_{ijk} \in X_{ij}$.

The platform cannot perfectly predict the firm's estimate of a worker's value. However, Assumption 3 says that conditional on $s(X_{ij})$, the residual variation in employer beliefs around the platform's prediction is orthogonal to the wage bid and other characteristics of the worker that are used to compute the score. This assumption seems reasonable given that $s(\cdot)$ is an algorithm that the platform has developed to predict hiring on the basis of X_{ij}.

Assumption 3 is crucial for our identification strategy, because it implies that in a narrow bandwidth around some value of the score, two workers with different wage bids have the same expected value to the firm. Therefore, Predictions 1 and 2 hold in expectation within the bandwidth.[13]

2.4.2 Regression Discontinuity Design

Suppose there is a threshold score s^* such that the platform recommends a worker i who has applied to employer j if $s(X_{ij}) \geq s^*$, and otherwise does not.

Assumption 4. The distribution of $\hat{u}_j(X_{ij})$ and the components of X_{ij} conditional on $s(X_{ij})$, $F_{\hat{u}_j(X_{ij})|s(X_{ij})}(u|s(X_{ij}))$ and $F_{x_{ik}|s(X_{ij})}(x|s(X_{ij}))$ for $x_{ijk} \in X_{ij}$, are continuous in $s(X_{ij})$.

[12]For simplicity, we assume here that each worker applies exactly one job opening. On the platform, workers may apply to multiple openings, and may receive a different score for each application.

[13]The model predictions hold even when the expected worker value in the bandwidth is not equal to zero and expected worker productivity does not equal the worker's wage bid, i.e. $\mathrm{E}[\hat{u}_j(X_{ij})|s(X_{ij})] = \mathrm{E}[\hat{y}_j(X_{ij}) - \ln w_{ij}|s(X_{ij})] = c \neq 0$.

Assumption 3 implies that the expectation of $\hat{u}(X_{ij})$, the worker's estimated value, does not jump discontinuously at the threshold.[14] This is similar to the typical regression discontinuity assumption that the potential outcomes are continuous in the running variable.

However, if the recommendation has teeth, then employers will estimate a recommended worker's productivity to be higher than that of a non-recommended worker with identical observable characteristics. In other words, in the presence of a recommendation system, firm j's estimate of worker i's value will be $\tilde{u}_j(X_{ij}, REC_{ij}) \equiv \hat{u}_j(X_{ij}) + \tau REC_{ij}$, where $REC_{ij} = \mathbb{1}(s(X_{ij}) \geq s^*)$ is an indicator for a worker being recommended and $\tau > 0$. This discontinuity in firm's beliefs about worker productivity created by the platform recommendations motivates a fuzzy regression discontinuity (RD) design.

The first stage of our RD captures the effect of a recommendation on the probability that a worker will be hired. We estimate the following linear probability model in a small bandwidth of size δ around the threshold:

$$HIRE_{ij} = \beta_0 + \beta_1 REC_{ij} + \beta_2 \tilde{s}(X_{ij}) + \beta_3 REC_{ij} * \tilde{s}(X_{ij}) + \phi_{jt} + \varepsilon_{ij}, \tag{2.8}$$

where $HIRE_{ij}$ is an indicator for firm j hiring worker i, $\tilde{s}(X_{ij}) \equiv s(X_{ij}) - s^*$ is the score centered around the cutoff, ϕ_{jt} are job opening-by-month fixed effects, and ε_{ij} is a residual with mean zero. The parameter β_1 gives the increase in the probability of being hired associated with a recommendation, controlling for linear terms in the score and allowing different slopes above versus below the threshold.

As with all RD designs, β_1 is a local estimate. Our model suggests that the firm hires the worker in its applicant pool with the highest estimated value. Therefore, the recommendation affects a firm's hiring decision if it changes employer beliefs about which applicant is the highest-value worker. For a specific employer, the impact of the recommendation discontinuity depends on where the threshold falls in the distribution of estimated worker value in the applicant pool. Complier firms are likely to be those who are deciding between

[14] In other words: $\mathrm{E}[\hat{u}_j(X_{ij})|s(X_{ij}) = s^*] = \lim_{s(X_{ij})\uparrow s^*} \mathrm{E}[\hat{u}_j(X_{ij})|s(X_{ij})] = \lim_{s(X_{ij})\downarrow s^*} \mathrm{E}[\hat{u}_j(X_{ij})|s(X_{ij})]$.

applicants with scores just below and just above the cutoff. Some employers may be always takers, in the sense they are considering a set of workers with score values far above the cutoff, all of whom are recommended. Other employers may be never takers who only receive applications from non-recommended workers substantially below the cutoff. Indeed, some employers may choose not to hire at all if they believe all their workers to be low-value.[15]

The second stage of our RD uses the discontinuity in hiring at the threshold to examine the impact of worker wage bids on hours of work and the wage bill at the job level. This strategy differs from the more typical RD approach of analyzing the effect of the hiring discontinuity on worker-level outcomes. We rely on Assumption 3, which implies that the distribution of the wage bid close to the cutoff fluctuates quasi-randomly across openings for reasons that are unrelated to $\hat{u}_j(X_{ij})$, the employer's beliefs about the worker's value. We can therefore exploit this variation to look at what happens when employer are induced to hire workers who make higher wage bids, holding constant expected worker value.

Specifically, because employers are more likely to hire workers who are recommended, the wages of workers above the cutoff will be positively correlated with the wage of the hired worker. Thus we can use the difference in the mean wage bid of workers above versus below the cutoff as an instrument for the hired worker's wage. This additional first-stage specification is given by

$$\ln w_{ij}^h = \pi_0 + \pi_1 \Delta_j + \xi_j, \tag{2.9}$$

where $\Delta_j \equiv \ln\overline{w}_j^{abv} - \ln\overline{w}_j^{bel}$ gives the mean log wage bid above the threshold minus the mean log wage bid below the threshold among workers who apply to employer j, restricting to a narrow bandwidth around the cutoff, w_{ij}^h is again the hourly wage paid to worker i hired by employer j, and ξ_j is a mean-zero residual.

Finally, our second stage equations use the Δ_j instrument to estimate the elasticity of

[15]Firms may have different reservation value levels below which they choose not to hire. In addition, the score is not a perfect prediction of firm estimates of worker value, $\hat{u}_j(\cdot)$. Therefore, some employers may choose not to hire when they have applicants in the bandwidth, or with even higher score values.

hours worked and the wage bill with respect to the hourly wage. We estimate

$$\ln h_{ij} = \theta_0 + \theta_1 \ln\hat{w}^h_{ij} + \nu_j \tag{2.10}$$

and

$$\ln C_{ij} = \gamma_0 + \gamma_1 \ln\hat{w}^h_{ij} + \mu_j, \tag{2.11}$$

where h_{ij} and C_{ij} are again the hours worked and wage bill of worker i hired by firm j, $\ln\hat{w}^h_{ij}$ is the fitted value of $\ln w^h_{ij}$ from Equation 2.9, and ν_j and μ_j are residuals with mean zero.[16]

In this context, the exclusion restriction is satisfied if the Δ_j instrument is not correlated with opening-level characteristics that affect hours worked and the wage bill. The assumption would be violated if, for example, employers with higher values of Δ_j attracted higher-wage applicants with specialized skills. To address this concern, we estimate versions of Equations 2.9, 2.10 and 2.11 that control for the mean wage bid in the applicant pool for each job opening.

2.5 Data

We use data from the online labor market on job posts created between January 2014 and May 2016 that were advertised publicly on the platform. As our model generates predictions about the relationship between hours worked and hourly wages, we limit our sample to openings for jobs that paid workers hourly. In addition, we restrict the sample to openings for which the employer hired exactly one worker, and to openings that received greater than 1 and less than 200 applications.

We use data on job applications made to these openings. In general, the platform recommends an applicant to an employer if the applicant is assigned a score of at least 0.5. In other words, the threshold is $s^* = 0.5$, and accordingly we define $REC_{ij} = \mathbb{1}(s(X_{ij}) \geq 0.5)$.[17]

[16]Note that although w^h_{ij}, h_{ij} and C_{ij} are indexed by both worker and firm, only one worker is hired for each job opening, and therefore these second-stage equations are estimated at the job opening level.

[17]In Sections 2.3 and 2.4 above, j indexes employer. In our data, employers may post multiple openings.

However, we find that during various periods in the timeframe under study, the platform appears to have experimented with alternative cutoffs, such that some applicants with scores below 0.5 are recommended and some with scores above 0.5 are not.

Define the false negative rate to be the ratio of the number of applicants with scores above 0.5 who are not recommended over the number who are recommended, and the false positive rate to be the ratio of the number of applicants with scores below 0.5 who are recommended over the number who are not recommended. We restrict the sample to job applications submitted on days when the false negative rate was less than 0.1 and the false positive rate was less than 0.2. Furthermore, we exclude applications that include an hourly wage bid of \$0.25 or less or of \$100 or more, applications made by workers who list an hourly wage on their profile of \$0.25 or less or of \$100 or more, and applications that receive a score of 0.05 or less.

In the analyses performed at the application level, such as Equation 2.8, we restrict the data to applications that receive a score within a bandwidth of $\delta = 0.1$ around score cutoff, $s(X_{ij}) \in (s^* - \delta, s^* + \delta) = (0.4, 0.6)$. We also restrict to applications to openings that receive at least one application with a score below the threshold but within the bandwidth and one application with a score above the threshold but within the bandwidth.

For the analyses performed at the level of the job opening, such as Equations 2.9, 2.10 and 2.11, we merge the applications data with information on job outcomes, specifically the wage of the hired worker, hours of work and the total wage bill or cost to the employer. We restrict this sample to openings where the hired worker's wage is greater than \$0.25 and less than \$100, and hours of work and the wage bill are greater than zero.

In all analyses, we use a version of the score that has been centered around the cutoff, $\tilde{s}(X_{ij}) = s(X_{ij}) - s^*$, such that a value of zero indicates the recommendation threshold.

Therefore, in remainder of the paper, j indexes opening.

2.6 Results

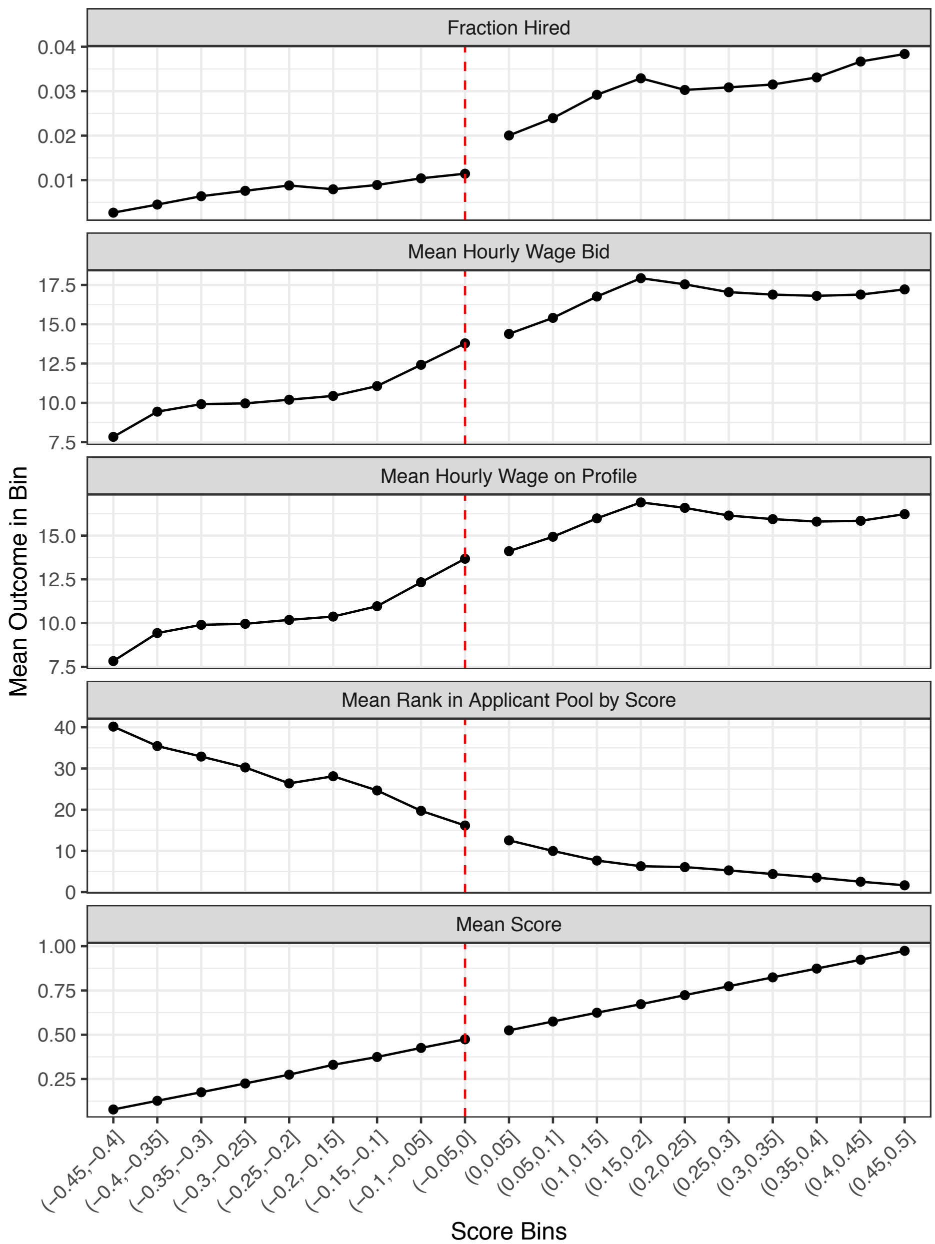
2.6.1 Effect of Recommendation on Hiring

This section presents results on the first-stage relationship for job applicants between receiving a platform recommendation and being hired. Figure 2.1 displays a series of binned scatter plots of application-level outcomes versus the score. Using our full sample of job applications not restricted to a narrow bandwidth around the threshold, we create bins of the score and calculate bin-level means. The first subplot is a visual representation of the first-stage specification in Equation 2.8 that shows the fraction of applicants hired in each score bin. The score has a positive relationship with the fraction hired even below the cutoff, as we would expect given our interpretation of the score as a prediction of the probability that an applicant will be hired. However, it is clear that the probability of being hired increases discontinuously at the recommendation threshold. By contrast, the hourly wage bid, hourly wage listed on the applicant's profile, and mean worker rank in the applicant pool are all smooth at the cutoff.[18]

Figure B.1 uses the applications data to fit local polynomial regressions that predict hiring as a function of the score, separately on each side of the threshold. We shows results for a grid of smoothing parameters (by column) and bandwidths (by row). All subplots show clear evidence of a discontinuity in the fitted lines at the cutoff.

To quantify this effect, Table 2.1 documents the relationship between the recommendation discontinuity and hiring in regression form, using the applications data. Column (1) shows the coefficient on the recommendation indicator from a linear probability model in which the outcome is an indicator for being hired. Column (2) controls for the score, centered at the cutoff. Column (3) reports results for our preferred specification from Equation 2.8, which allows the coefficient on the score to differ above versus below the cutoff. Lastly, Column (4) controls for a second-degree polynomial in the score. All specifications restrict the data to a bandwidth of 0.1 around the score cutoff, include opening-by-month fixed

[18]Section B.1 discusses the validity of the RD design in greater depth.



Notes: This figure shows binned scatter plots of mean application-level outcomes against score bins. The recommendation threshold is demarcated by a vertical dashed line. The rank variable indicates the worker's rank in the applicant pool by score, so that a rank of one indicates that the applicant has the highest score among all applicants to the opening.

Figure 2.1: *Binned Scatter Plot of Application-Level Outcomes and the Score*

effects, and cluster standard errors by worker.[19]

Table 2.1: *Effect of Recommendation on Hiring*

	Outcome:			
	Hired			
	(1)	(2)	(3)	(4)
Recommended	0.011**	0.006**	0.006**	0.006**
	(0.0001)	(0.0003)	(0.0003)	(0.0003)
Score		0.049**	0.026**	
		(0.002)	(0.003)	
Recommended*Score			0.050**	
			(0.005)	
Score Polynom.	N	N	N	2nd deg
Observations	3,648,710	3,648,710	3,648,710	3,648,710
R^2	0.181	0.181	0.182	0.182
Adjusted R^2	0.056	0.056	0.057	0.057

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports linear probability estimates of Equation 2.8, using the applications data and restricting to a bandwidth of 0.1 around the recommendation threshold. All columns include opening-by-month fixed effects and cluster standard errors by worker.

All columns in Table 2.1 show a highly significant and economically meaningful relationship between recommendation and hiring. Specifically, our preferred specification in Column (3) indicates that recommended workers are 0.6 percentage points more likely to be hired. Recommendation thus increases the probability of hiring by more than 40 percent percent, from a base of 0.014, for applicants near the threshold.

2.6.2 Effect of Delta Instrument on Hired Worker Wage

Table 2.2 reports estimates of our additional first stage in Equation 2.9. In Column (1), we regress the log hourly wage of the hired worker on the Δ_j instrument, using all openings that appear in the data for applications within the bandwidth. Column (2) controls for the mean log wage bid among applicants within the bandwidth, as it is possible that the distribution of wages may be correlated with the mean level of wage bids across openings. Similarly,

[19]Table B.1 shows summary statistics for the variables in Table 2.1, restricting to a bandwidth of 0.1 around the recommendation threshold.

Column (3) includes controls for the mean score above and below the threshold but within the bandwidth, as Δ_j may be associated with the score distribution. Columns (4)-(6) repeat the specifications in Columns (1)-(3), restricting the sample to job openings where the hired worker is within the bandwidth.[20]

In all models, we find a highly significant relationship between Δ_j and the hired worker wage. The coefficient in Column (1) suggests that an increase of 10 logs points in the mean log difference in wage bids above versus below the cutoff raises the hired worker's wage by 0.6 percent. The relationship is approximately twice as large in magnitude when restricting to openings where the hired worker is in the bandwidth in Column (4). This stronger association is not surprising given that employers who select a worker within the bandwidth are more likely to be compliers who change their hiring decisions in response to the platform recommendation. Adding controls for the the mean log wage bid in Columns (2) and (5) and the mean score above and below the cutoff in Columns (3) and (6) changes the coefficients very little.

2.6.3 Effect of the Wage on Hours Worked and the Wage Bill

Table 2.3 displays estimates of the effect of the hired worker's wage on hours worked. Column (1) shows results from the OLS regression of log hours worked on the log hourly wage, based on all openings in the applications data within the bandwidth. Column (2) reports the IV estimate of this relationship shown in Equation 2.10, using Δ_j as an instrument for the hired worker's wage. Column (3) controls for the mean log wage bid within the bandwidth in the first- and second-stage equations. Similarly, Column (4) controls for the mean score above and below the cutoff within the bandwidth. Columns (5) and (6) repeat Columns (1) and (2), restricting to openings where the hired worker is within the bandwidth.[21] In all specifications, the coefficient on the hired worker's wage can be

[20]Table B.2 shows summary statistics for the variables at the job opening level, using the full sample in Columns (1)-(3).

[21]We also estimate specifications comparable to Columns (3) and (4) but restricting to openings where the hired worker is within the bandwidth (not shown), and find results similar very similar to the estimate in

Table 2.2: *Effect of Delta Instrument on Hired Wage*

	Outcome:					
	Log Hourly Wage of Hired Worker					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ_j	0.063**	0.065**	0.062**	0.119**	0.129**	0.118**
	(0.005)	(0.003)	(0.005)	(0.009)	(0.005)	(0.009)
Mean Log Wage Bid		0.941**			1.041**	
		(0.002)			(0.004)	
Mean Score Below			4.965**			6.063**
			(0.213)			(0.386)
Mean Score Above			5.500**			5.490**
			(0.220)			(0.398)
Restrict to Openings w/ Hired Worker in BW	N	N	N	Y	Y	Y
Observations	117,624	117,624	117,624	36,036	36,036	36,036
R^2	0.002	0.562	0.020	0.005	0.702	0.028

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports OLS estimates of Equation 2.9. The data are aggregated to the job opening level. The Δ_j instrument is the difference between the mean log wage bid of applicants above versus below the recommendation threshold but within the score bandwidth of 0.1 around the cutoff. The other predictors are the mean log wage bid, the mean score below the threshold and the mean score above the threshold, among applicants within the bandwidth. Columns (1)-(3) include all job openings. Columns (4)-(6) restrict to openings where the hired worker is within the bandwidth.

interpreted as the elasticity of hours worked with respect to the hourly wage.

Table 2.3: *Effect of Hired Worker Wage on Hours Worked*

	Outcome:					
	Log of Hours Worked					
	OLS	IV	IV	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Log Hired Wage	-0.172^{**}				-0.190^{**}	
	(0.006)				(0.011)	
Fitted Hired Wage		-0.441^{**}	-0.433^{**}	-0.443^{**}		-0.238
		(0.153)	(0.148)	(0.155)		(0.154)
Mean Log Wage			0.233^{+}			
			(0.139)			
Mean Score Below				0.577		
				(0.875)		
Mean Score Above				0.757		
				(0.981)		
Restrict to Openings w/ Hired Worker in BW	N	N	N	N	Y	Y
Observations	117,624	117,624	117,624	117,624	36,036	36,036

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports OLS and IV estimates of Equation 2.10. The data are aggregated to the job opening level. In the IV estimates, the Δ_j instrument is used to instrument for the log hired wage. The mean log wage bid, the mean score below the threshold and the mean score above the threshold are included in the first- and second-stage equations as indicated. Columns (1)-(4) include all job openings. Columns (5)-(6) restrict to openings where the hired worker is within the bandwidth.

The elasticities estimated in Table 2.3 are much smaller in magnitude than the value of -1 implied by Prediction 1 in the model. The OLS estimates in Columns (1) and (4) are very similar to one another, and suggest that a 1 percent increase in the hourly wage decreases hours worked by 0.17 to 0.19 percent. The IV estimates using all openings in Columns (2) to (4) are somewhat larger, and imply a decrease in hours worked of 0.43 to 0.44 percent. Finally, the IV coefficient in Column (6), which restricts to openings where the hired worker is within the bandwidth, is just over half the magnitude of the coefficients in Columns (2) to (4) at -0.24, and is insignificant, possibly due to the relatively small sample size.[22]

Column (6).

[22] Table B.5 reports reduced form estimates of the impact of Δ_j on hours worked and the wage bill. The coefficient on the instrument is almost identical in Columns (1) and (2), which show results on log hours worked

Finally, Table 2.4 shows estimates of the effect of the hired worker's wage on the wage bill, or worker cost. The specifications in Table 2.4 are comparable to those in Table 2.3, but with the log total wage bill, rather than log hours worked, as the outcome. Again, the coefficients on the hired worker's wage can be interpreted as elasticities. Prediction 2 in the model implies that the wage bill for each opening should be constant with respect to the worker hired, and that the elasticity should therefore have a value of 0.

Table 2.4: *Effect of Hired Worker Wage on Wage Bill*

	Outcome:					
	Log of Total Wage Bill					
	OLS	IV	IV	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Log Hired Wage	0.830**				0.807**	
	(0.007)				(0.013)	
Fitted Hired Wage		0.550**	0.558**	0.545**		0.964**
		(0.174)	(0.168)	(0.176)		(0.184)
Mean Log Wage			0.238			
			(0.158)			
Mean Score Below				0.553		
				(0.994)		
Mean Score Above				1.003		
				(1.114)		
Restrict to Openings w/ Hired Worker in BW	N	N	N	N	Y	Y
Observations	117,624	117,624	117,624	117,624	36,036	36,036

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports OLS and IV estimates of Equation 2.11. The data are aggregated to the job opening level. In the IV estimates, the Δ_j instrument is used to instrument for the log hired wage. The mean log wage bid, the mean score below the threshold and the mean score above the threshold are included in the first- and second-stage equations as indicated. Columns (1)-(4) include all job openings. Columns (5)-(6) restrict to openings where the hired worker is within the bandwidth.

Contrary to Prediction 2, the estimates in Table 2.4 suggest that hiring a higher wage worker increases the wage bill meaningfully. The coefficients on the log hired worker's wage are positive and quite sizable in magnitude. Specifically, the IV estimate in Column (2)

for all openings and openings where the hired worker is in the bandwidth, respectively. However, the estimate in Column (2) has a much larger standard error and is insignificant. The combination of a stronger first stage and a similar but insignificant reduced form estimate when restricting to workers in the bandwidth leads to an IV estimate that is insignificant and smaller in magnitude, compared with the IV estimate using all openings.

implies that a 1 percent increase in the hourly wage increases the wage bill by 0.55 percent. Column (6), which restricts to openings where the hired worker is within the bandwidth, suggests an increase in the wage bill of nearly 1 percent is associated with a 1 percent rise in the hired worker's wage.[23]

2.7 Discussion and Conclusion

This paper examines the prediction that workers are paid their marginal productivity, based on a simple competitive model of the labor market. We use data from an online labor market in which each job can be interpreted as a discrete project, and hours of work and wages are measured with great accuracy. We exploit an institutional feature of the platform that quasi-randomly induces employers to hire workers who make different hourly wage bids, but are similar in terms of expected employer beliefs about their value.

The model implies that higher wage workers should complete a fixed amount of work more quickly, leaving the wage bill unchanged. Our results are not consistent with this prediction. While higher-wage workers do work fewer hours, our elasticity estimates are much small in magnitude than the value of -1 suggested by the model. In addition, we find that higher-wage workers cost employers substantially more. In essence, our results suggest that wages are more dispersed than productivity, such that lower-wage workers are better value for money, while higher-wage workers are worse.

Why do our findings contradict the competitive market model? There are several possible explanations:

Explanation 1. *Higher-wage workers are more productive on unobserved dimensions.*
It may be that higher-wage workers are more productive in a way that justifies their higher cost on a metric other than the amount of time required to complete a project. For example, a higher wage may translate into a higher quality of output or a higher likelihood of project completion. In addition, some employers may value quality or the probability of successful

[23] The OLS estimates in Columns (1) and (4) are also quite sizable, with values of 0.81 to 0.83.

completion more than others, perhaps because of the scale or importance of the project (Terviö, 2008). Equivalently, lower-wage workers may be less productive in a way we cannot observe. If this hypothesis is true, employers who are induced to hire a higher-wage worker may report being more satisfied, or at least will not report being less satisfied.

Explanation 2. *Employers expand the project scope.*

Perhaps higher-wage workers are undervalued by employers, in the sense that they are paid less than their marginal productivity and complete projects more quickly than anticipated. In other words, Assumption 2 may not hold because $\mathrm{E}[\hat{y}_j(Z_i)] < \mathrm{E}[\ln y_i | Z_i]$ among higher-wage workers, and the opposite is true among lower-wage workers. If this is the case, employers who are induced to hire a higher-wage worker may be pleasantly surprised and decide to expand the project scope, which would appear in our data as more hours worked and a higher total wage bill.[24] Again, this explanation suggests that employers who hire higher-wage workers will give more positive feedback.

Explanation 3. *Employers overestimate the productivity of higher-wage workers.*

Employers may believe higher-wage workers to be more productive than they actually are, such that Assumption 2 is violated because $\mathrm{E}[\hat{y}_j(Z_i)] > \mathrm{E}[\ln y_i | Z_i]$ among higher-wage workers, and the opposite among lower-wage workers. In particular, it may be the case the Assumption 1 does not hold, and employers take a worker's wage bid into consideration when estimating their productivity, even after conditioning on other observable characteristics. Perhaps employers believe that other firms know something they do not, and that a worker who makes a higher wage bid must be unobservably more productive because otherwise they would not be able to command this wage in the market. This hypothesis is consistent with Barach and Horton (2020), who find that employers hire workers with lower average past wages when an online labor market platform experimentally conceals information about compensation history, suggesting that employers view higher past wages

[24]Similarly, employers induced to hire a lower-wage applicant may be unpleasantly surprised by how long the worker is taking, and decide to cut the project short.

as a positive signal of worker ability. However, in reality it may be that workers who make higher wage bids conditional on their observed characteristics are simply overconfident. Equivalently, higher-wage workers may be productive but may unilaterally decide to expand the project scope and work additional hours on tasks that the employer does not need to be performed. This hypothesis suggests that employers induced to hire higher-wage workers will be less satisfied, on average.

Explanation 4. *Employers believe higher-wage workers to be worse value for money.*
Assumption 3 may be violated, such that employers believe higher-wage workers to be lower value in expectation after conditioning on the score. In other words, the score may be a bad approximation of employer beliefs. Thus the Δ_j instrument may induce employers to hire workers who make higher wage bids who they would otherwise judge to be less qualified relative to their cost, compared with other workers in their applicant pool. Table B.6 shows that worker wage bids negatively predict hiring within the bandwidth even after controlling flexibly for the score, providing some evidence in support of this hypothesis. In this case, our results suggest that worker wage bids do not reflect marginal productivity, and that wage bids are positively correlated with worker over-confidence. Again, we expect that employers induced to hire higher-wage workers will be less satisfied in this scenario.

Explanation 5. *Wages are affected by search frictions and/or monopsony.*
The literature on wage determination suggests many reasons that workers may not be paid their marginal product. Search costs may lead workers to accept wages below their marginal productivity. Similarly, vacancy costs may cause employers to be willing to pay workers more than their marginal productivity.[25] Our results may be generated by a situation in which higher-wage workers have relatively more bargaining power compared with low-wage workers, for example due to better outside options (Horton, 2020; Caldwell and Danieli,

[25]We might also expect that informational frictions contribute to workers' being paid more than their marginal product. If search costs for workers are low and employers form unbiased estimates of worker productivity but cannot observe other firms' estimates, then workers may be able to "shop around" to find an employer who is willing to pay more than their marginal productivity.

2018), leading them to have higher wages relative to their marginal product.[26] Alternatively, there may be monopsony in the labor market that affects low-wage workers more than higher-wage workers, such that low-wage workers are paid systematically less relative to their productivity. In both of these scenarios, wages are more dispersed than productivity, leading the elasticity of hours with respect to wages to be less than one in absolute value. We cannot rule out these hypotheses. However, they do not seem very realistic given that our identification strategy uses variation in wages *within* applicant pools, and our estimates are not sensitive to controlling for the mean wage within the applicant pool.[27] In addition, Zoega and Booth (2005) provide evidence that monopsony is more likely to cause wage compression relative to ability, rather than dispersion.

Further research is needed to disentangle these possible explanations. In particular, the finding that employers who are induced to hire higher-wage workers are more satisfied, on average, would be consistent with Explanations 1 and 2, while the finding that employers are less satisfied would be consistent with Explanations 3 and 4.

---

[26]This situation is theoretically possible in the baseline model with search frictions in Romer (1992), in which wages are an increasing function of a worker's outside option and outside options depend positively on ability.

[27]In other words, we would expect these differences in outside options and especially monopsony to occur primarily across job openings that attract different types of applicants.

Chapter 3

Algorithmic Technology and Skill Requirements in Cognitive Occupations

Chapter Abstract

Workers rely on data to make decisions in increasingly sophisticated ways. This paper examines how skill requirements in two cognitive occupations—marketing managers and financial analysts—are related to the use of technology that facilitates data-driven decision-making (*algorithmic technology*). Drawing on data from online job postings, I find that in both occupations, mention of algorithmic technology is positively associated with complementary technical skills, although the nature of those skills differs by field. By contrast, I find that algorithmic technology is negatively related to many frequently-listed non-routine cognitive skills in both occupations. These relationships are robust to controlling for employer, month and geographic area fixed effects. I link the skill measures to data on wages, and document that mention of algorithmic technology is positively correlated with realized hourly earnings across geographic area and time. These results suggest that data from online job postings can be valuable in understanding how technology use is related to skill requirements and

wages in cognitive occupations.

3.1 Introduction

Workers in many occupations use data to inform decision-making in increasingly sophisticated ways. Media reports stress the growing importance of big data in the workplace. In economics, one strand of research examines how algorithms can improve the quality of decisions made by policymakers and private sector firms (Kleinberg *et al.*, 2018). Another branch of literature assesses whether technological advances in algorithmic decision-making are likely to lower employment and wages for certain groups of workers, or eliminate some occupations entirely (Frey and Osborne, 2013). Previous waves of automation linked to information technology largely affected middle-skill workers engaged in routine tasks (Acemoglu and Autor, 2011). However, there is evidence that more recent innovations in machine learning and artificial intelligence in particular may have an impact on workers in high-skill cognitive occupations (Brynjolfsson and McAfee, 2014).

Technology that facilitates data-driven decision-making, hereafter *algorithmic technology*, may also transform the nature of the work performed in certain occupations. Theory suggests that technological change leads to the creation of new tasks as well as the automation of existing ones (Agrawal *et al.*, 2019; Acemoglu and Restrepo, 2018). In addition, firms may change the skill profile of workers they recruit as the tasks involved in an occupation shift. Relatively little attention has been paid to understanding how job tasks and skills evolve in occupations where workers begin using algorithmic technology.

This paper uses data from online job postings to examine how skill requirements change as firms adopt algorithmic technology in two cognitive occupations. The data include a list of skills mentioned in each post, providing insight into heterogeneity in skill demands within occupations. Technologies that enable data-driven decision-making take very different forms in different occupational contexts. I focus on specific occupations to avoid the challenge of constructing a measure of algorithmic technology that is relevant in all jobs. The case study approach also enables me to exploit the richness of the data by defining occupation-specific skills.

I focus on two professional occupations—marketing managers and financial analysts—in which algorithmic technology plays an important role, as measured by the share of job postings that mention machine learning or artificial intelligence. Over the past decade, employment in these occupations has increased while wages have held steady, suggesting rising or non-declining demand for these jobs. Marketing managers use marketing automation software to automate interactions with customers in a personalized manner and measure progress on important outcomes. Financial analysts utilize statistical software to develop financial models in order to predict business performance and make investment recommendations. Notably, neither marketing managers nor financial analysts are heavily regulated or require occupational licenses. Thus employers may have relative flexibility to change job descriptions and skills requirements in these occupations.

I use a difference-in-differences approach to assess which skill requirements predict the mention of algorithmic technology in job postings for marketing managers and financial analysts during a recent nine-year period. Specifically, the richness of the data allow me to control for time, employer and geographic area fixed effects at the job posting level. I construct measures of algorithmic technology separately for each occupation. For marketing managers, I measure algorithmic technology as the mention of marketing automation software in a job posting. For financial analysts, I focus on the mention of statistical software and general-purpose programming languages commonly used for statistical analysis. In addition, I define a set of skill composites for each job using the most commonly-mentioned skills in the first and last years of the period under study.

I find that in both occupations, mention of algorithmic technology is positively predicted by complementary software skills. However, the nature of these skills differs by field. For marketing managers, postings listing automation software are more likely to also mention software that requires limited or no coding ability, such as data visualization tools and especially customer relationship management (CRM) platforms, which are closely related to marketing automation. However, postings listing marketing automation software are less likely to mention software skills that involve coding.

These results suggest that marketing managers follow what I label a *data dashboard* model in their use of algorithmic technology. Marketing managers exploit rules-based automation to target and interact with customers, and draw on simple statistical analyses to measure outcomes and inform future actions. However, they do not need to understand the algorithms or statistical analyses at an advanced level to take advantage of this technology.

In the job postings for financial analysts, by contrast, the mention of statistical programming languages is positively predicted by other software skills that involve coding, as might be expected, as well as skills involving a more limited degree of data manipulation. However, postings mentioning statistical programming are less likely to list "point and click" software platforms such as CRM and even some financial software. In addition, postings listing statistical programming are more likely to require conceptual knowledge of highly quantitative domains such as economics, statistics and risk. These findings suggest an *analyst* model whereby financial analysts incorporate statistical prediction methods into their toolkit for building financial models, and become direct users or even producers of decision-making algorithms.

I also find that the mention of algorithmic technology is negatively predicted by the majority of the most frequently-listed non-routine cognitive skills in both occupations. These skills include interpersonal abilities and other general human capital, such as writing, as well as more occupation-specific competencies, such as product management and development for marketing managers and financial reporting for financial analysts. By contrast, a large literature on skill-biased technical change and the task content of work suggests that earlier rounds of information technology adoption led to increasing demand for non-routine cognitive skills, including broad competencies not directly related to technology such as social skills (Deming, 2017; Borghans *et al.*, 2014; Catherine, 2014).[1]

[1] Many of these studies examines changes in the employment distribution across occupations with different skill requirements, rather than changes in skill requirements within occupations (Acemoglu and Autor, 2011). However, there is some case study evidence that broad non-routine cognitive skills such as problem solving have become more important within certain occupations (Bartel *et al.*, 2007; Autor *et al.*, 2002).

These results suggest that the use of algorithmic technology by marketing managers and financial analysts is primarily accompanied by growing demand for complementary technical skills. However, the job postings data may reflect employer aspirations rather actual technology adoption or changes in worker skills. Therefore, in the last section of the paper, I examine how the skill measures from the job postings data are correlated with realized worker wages, which are likely to co-vary with worker skills as well as firm productivity. Specifically, for each occupation I aggregate the job postings by geographic area and year, and link to data on wages from the Occupational Employment Statistics (OES). I find that mention of algorithmic technology is positively related to hourly wages for both marketing managers and financial analysts.

3.1.1 Related Literature

This project contributes to an extensive literature on the impact of technology on workers and jobs over the last century (Goldin and Katz, 2007; Katz and Murphy, 1992; Katz and Autor, 1999; Acemoglu and Autor, 2012). A large body of evidence suggests that advances in computer and information technology, in particular, have automated routine tasks performed by middle-skill workers while increasing demand for high-skill workers engaged in non-routine cognitive work, leading to labor market polarization (Autor *et al.*, 2003, 2006; Acemoglu and Autor, 2011; Autor and Dorn, 2013a; Goos *et al.*, 2014; Deming, 2017). Recent studies, however, suggest that the employment and wage advantage enjoyed by highly educated workers may be eroding (Autor, 2014a; Beaudry *et al.*, 2014, 2016; Autor, 2014b). In particular, there is concern that innovations in machine learning and artificial intelligence are enabling the automation of complex cognitive tasks previously performed by highly skilled workers (Brynjolfsson and McAfee, 2011, 2014). Most of this work focuses on the reduced form impact of these technologies on wages and employment across the occupation distribution (Frey and Osborne, 2013; Felton *et al.*, 2019; Acemoglu and Restrepo, 2020, 2018; Webb, 2020).

This project contributes to the literature by providing insight into how the structure and task content of work—as measured by skill requirements—may change as the use of algorithmic technology becomes widespread within an occupation. In order to flexibly define technology use and skill categories to fit the occupational context, I focus on specific jobs. Thus I contribute to a strand of the literature that uses a quantitative case study approach to explore the effect of digital technology adoption on the workplace in production (Bartel *et al.*, 2007, 2004, 2003) and administrative jobs (Autor *et al.*, 2002; Dillender and Forsythe, 2019). In contrast to previous work, I examine professional occupations in which workers typically have a college degree or more.[2] In addition, while existing research suggests that technology adoption is often associated with an increase in overall education and skill requirements (Bartel *et al.*, 2007; Autor *et al.*, 2002), I find that the mention of algorithmic technology is correlated with complementary technical skills but not broad non-routine cognitive competencies.

Finally, this paper relates to recent work work that uses data from job postings to measure variation in skill requirements within occupations (Sasser Modestino *et al.*, forthcoming, 2016; Hershbein and Kahn, 2018; Dillender and Forsythe, 2019).[3] Most of the literature on skill-biased technical change, by contrast, examines changes in employment and wages across occupations with different task content.[4] In particular, I build on the work of Deming and Kahn (2018), who document that skill requirements in job postings are correlated with independent measures of earnings and firm performance, even controlling for detailed occupation. These results motivate my approach of measuring the use of algorithmic technology based on language in job postings.

The remainder of the paper proceeds as follows. Section 3.2 discusses the occupations analyzed, including how these jobs are selected. Sections 3.3 and 3.4 describe the data and

[2]A separate body of research examines the impact of technology on high-skill occupations such as lawyers (Remus and Levy, 2017) and radiologists (Levy, 2008) using a qualitative approach.

[3]For example, Hershbein and Kahn (2018) and Sasser Modestino *et al.* (2016) use data from job postings to investigate the impact of the Great Recession on education and experience requirements controlling for occupation and job title, respectively. Dillender and Forsythe (2019) find that firms that adopt technology for office and administrative jobs increase skill requirements for these roles.

[4]Many of these studies use data from the O*NET, a database of occupational characteristics, or its predecessor the Dictionary of Occupational Titles (DOT).

empirical strategy, respectively. Section 3.5 reports the results of the analysis, and Section 3.6 concludes.

3.2 Occupational Contexts

I select the cognitive occupations that I focus on in this project by ranking all occupations according to the share of job postings that mention keywords related to algorithmic technology in 2017, the last year of my data.[5] Specifically, I calculate the proportion of postings for each occupation that mention *machine learning* or *artificial intelligence* in the list of required skills. Although the specific technological tools used to facilitate data-driven decision-making are likely to vary widely across jobs, I interpret these keywords as a way for employers to signal that they value cutting-edge data analysis techniques.

I use a variant of the Standard Occupational Classification (SOC) codes developed by the O*NET, a Department of Labor database of occupational characteristics, that includes some additional detailed occupations. I conceptualize cognitive occupations as jobs that require high-level analytical and critical thinking skills and employ primarily college-educated workers. As a proxy for this type of job, I focus on professional occupations. However, I exclude computer and math, engineering and science occupations, as statistical analysis and the use of programming languages are an inherent component of many of these jobs. Therefore, methodological advances in computation or analytical techniques are unlikely to change the basic structure of these occupations. Lastly, I focus in particular on business occupations that are lightly regulated and do not require occupational licensing, as I hypothesize that firms may have more flexibility in adjusting job descriptions and skill requirements in these occupations in response to technological change.

Using this approach, I choose two occupations—*marketing managers* (O*NET SOC code 11-2021.00) and *financial quantitative analysts* (O*NET SOC code 13-2099.01). Financial quantitative analysts are, in fact, the top-ranked occupation in terms of the mention of

[5] Section 3.3 describes the job postings data in greater detail.

machine learning and artificial intelligence—more than 10 percent of job postings for this occupation mention these skills in 2017. However, relatively few job postings are categorized as falling into this occupation, as it is an additional detailed occupation defined by the O*NET that falls within the *all other financial specialists* SOC category.[6] Therefore, in the analysis I combine financial quantitative analysts with the much larger occupation of *financial analysts* (SOC code 13-2051), hereafter *traditional financial analysts*.

Figure 3.1 shows employment and mean hourly wages in real terms for marketing managers and traditional financial analysts by year for 2010-2017, using data from the OES. Employment has increased moderately over this period while wages have been largely flat, suggesting rising or steady demand for these jobs.[7] Figure 3.2 shows the number of job postings by month over a similar period for marketing managers, traditional financial analysts and financial quantitative analysts as a share of the total number of job postings in each month. The job postings are noisier than the employment data, but illustrate a similar pattern.

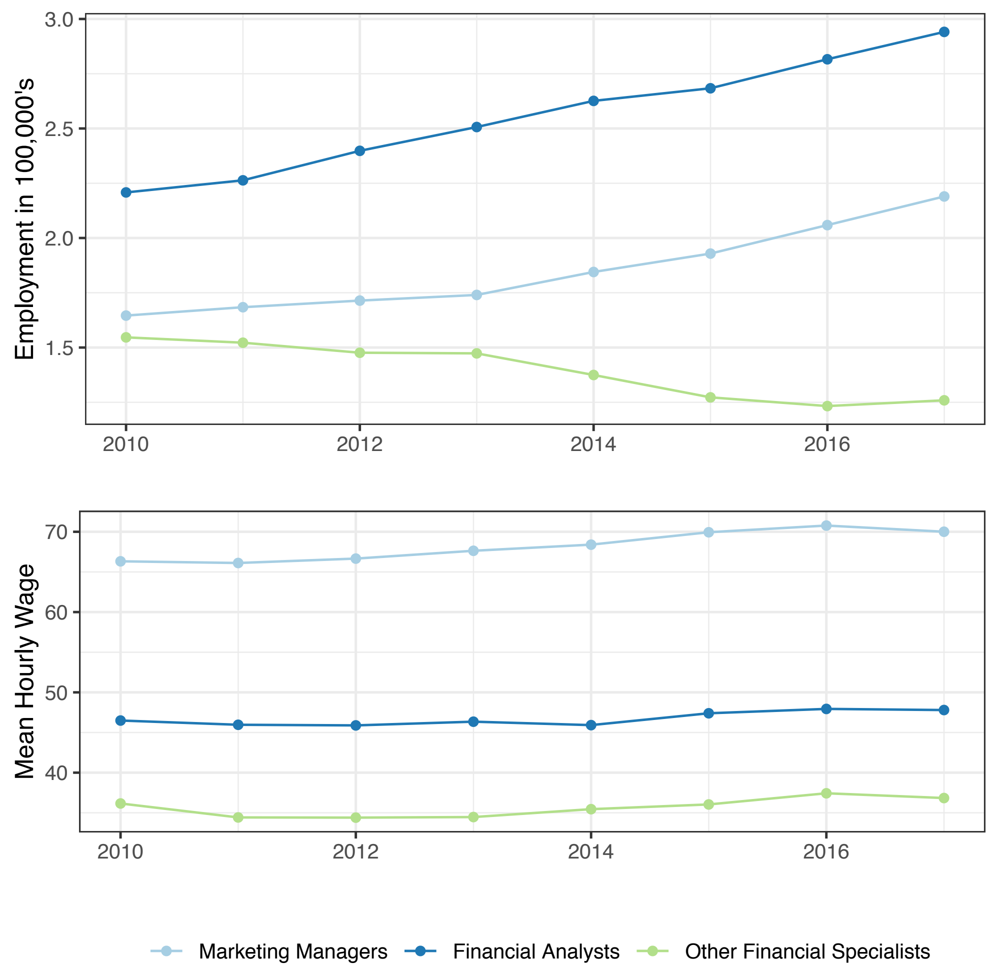
In the remainder of this section, I provide additional detail about the selected occupations.

3.2.1 Marketing Managers

Evidence from industry surveys suggests that the field of marketing has increasingly incorporated digital technology and data-driven decision-making into all aspects of its operations (Gartner, 2020). The rise of e-commerce and expansion of individuals' digital footprints has provided a wealth of data that marketers can use to improve targeting of potential customers, as well as new modes of interacting with current and future customers.

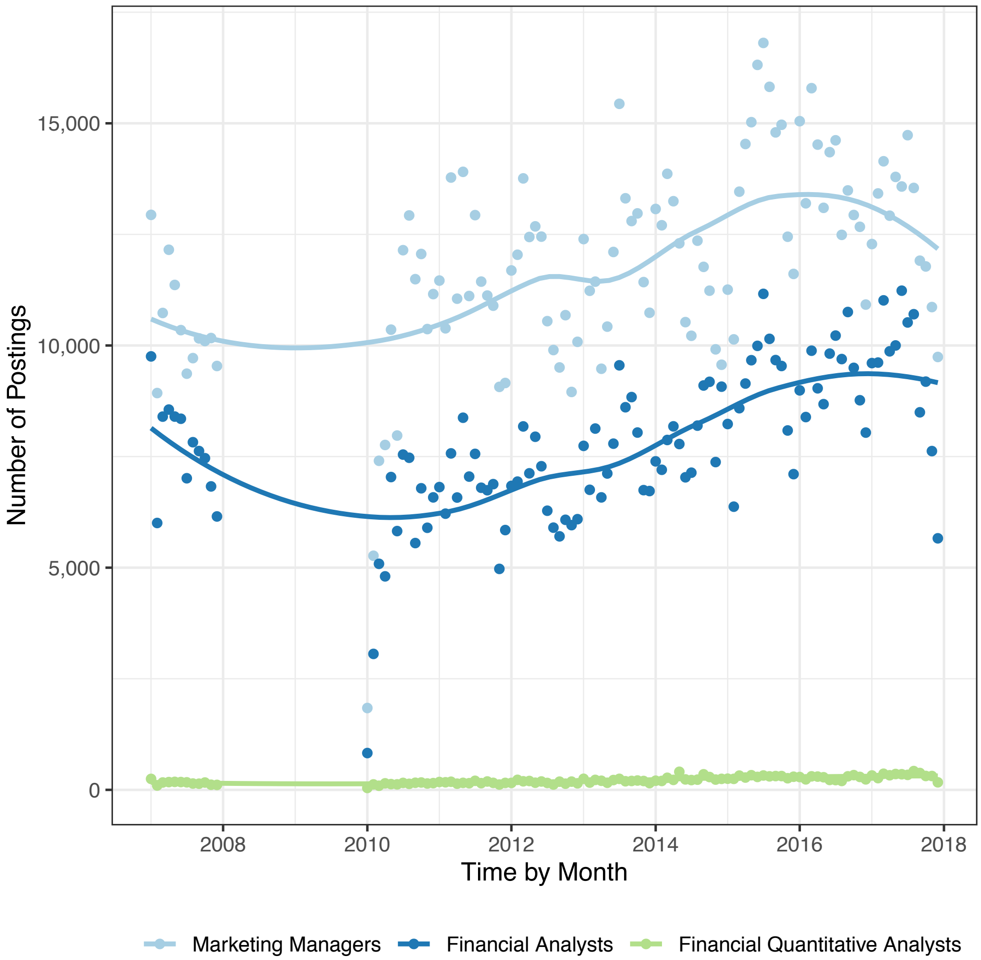
[6] In addition, there are no employment or wage data available on financial quantitative analysts, as the typical sources for this information use the SOC codes or an even more aggregated set of codes.

[7] The trends are similar if I calculate employment in each occupation as a share of total employed workers in each year, and calculate wages in each occupation as a share of the overall mean hourly wage. I also show employment and wages for the other financial specialists occupation, which includes financial quantitative analysts. Wages for other financial specialists are flat over the period under study, while employment appears to decrease. However, I cannot make inferences about trends for financial quantitative analysts without knowing what share of other financial specialists they represent.



Notes: This figure shows employment and the mean hourly wage in 2017 dollars by year for marketing managers, financial analysts and other financial specialists, using Occupational Employment Statistics (OES) data for 2010-2017.

Figure 3.1: *Employment and Wages in Selected Occupations*



Notes: This figure shows the number of job postings for marketing managers, financial analysts and financial quantitative analysts by month in 2007 and 2010-2017 (dots), along with fitted local polynomial regressions (lines).

Figure 3.2: *Job Postings in Selected Occupations*

I focus in particular on marketing automation software as the algorithmic technology that seems most relevant to the performance of marketing managers' jobs. Marketing automation systems perform tasks such as sending marketing emails, posting on social media and implementing advertising campaigns.[8] Typically, these systems track individuals' interactions with a company or brand, including website browsing behavior as well as purchase history, and enable marketers to customize interactions with customers on the basis of these data in a automized way. Marketing automation software also helps users to track their efforts and metrics of interest, often providing convenient data visualization and aggregation features. While many marketing automation platforms include recommendations that may be based on machine learning or other prediction algorithms, the core functionality of these systems involves the user defining simple rules that the software implements.

3.2.2 Financial Analysts

The job of traditional financial analysts is inherently quantitative and data-driven, as the core function of these individuals is to analyze financial information for the purpose of making investment decisions or recommendations.[9] However, industry publications suggest that asset management firms have increasingly incorporated sophisticated statistical and computational techniques into their analytical processes and even in some cases their core investment products (McKinsey, 2019). I interpret the O*NET occupation of *financial quantitative analysts* as a subset of a broader conceptual category of financial analysts in which workers rely heavily on these techniques.

In general, financial analysts must possess some knowledge of coding in order to implement a predictive statistical method. Therefore, I focus on programming languages as the measure of algorithmic technology that is most clearly related to advances in data-driven

---

[8]For example, see `https://www.hubspot.com/marketing-automation-information`.

[9]The O*NET summarizes the activities in this occupation as "conduct quantitative analyses of information affecting investment programs of public or private institutions."

decision-making in this job.

3.3 Data

I use data on employer vacancies from Burning Glass Technologies (BGT), a labor market analytics firm that aggregates, deduplicates and processes job advertisements posted online.[10] BGT reports information on each posting in a set of standardized fields that include the date posted, occupation, employer, geographic area and education and experience requirements. In addition, the data include a list of detailed skill requirements mentioned in each post, a subset of which are classified as software skills.[11]

My sample includes data on job postings for marketing managers (SOC code 11-2021), traditional financial analysts (SOC code 13-2051), and financial quantitative analysts (O*NET SOC code 13-2099.01) for the years 2007 and 2010 through 2017. The data from the two financial analyst codes are pooled in analyses for this occupation. I construct measures of skills related to algorithmic technology, separately for the two fields. For marketing managers, this variable takes the form of an indicator for the posting mentioning a marketing automation platform or company, such as Marketo or Hubspot (hereafter the indicator for *marketing automation* skills). For financial analysts, I measure algorithmic technology as an indicator for the posting listing at least one statistical software package such as SAS, MATLAB and SPSS or a general programming language commonly used for data analysis, including Python, R and SQL (hereafter the indicator for *programming* skills).

In addition, I define a set of skill composites based on frequently required skills. Specifically, I consider the top 50 most commonly listed software skills in 2017 and the top 25 mostly commonly listed software skills in 2007 in each occupation, and group these skills into composite measures by function. For example, the software skill composites include

---

[10] Several previous papers use these data, and have performed various validation checks, including comparing the occupation and industry distributions in the job postings to employment data (Sasser Modestino *et al.*, 2016, forthcoming; Hershbein and Kahn, 2018; Deming and Kahn, 2018; Rothwell, 2014).

[11] These skills are standardized to a certain extent by BGT's proprietary text processing algorithms.

Enterprise Resource Planning (ERP) software, data visualization tools, and skill keywords related to software development. I also consider the top 20 most commonly listed skills overall in 2017 and 2007 in each occupation.[12] Table C.1 lists the skill composites, along with the component skills and relevant occupation(s) for each composite. The composites are defined as as indicator for the posting mentioning at least one of the component skills as a requirement for the job. I construct the composites separately for each occupation, but in cases where a skill is top-listed for both occupations, I define the composite containing that skill identically for both occupations.[13]

As there are over 30 skill composites for each occupation, I also group the composites into more broadly defined skill categories. I divide the software skill composites into those that definitively require coding (e.g. software development),[14] those that involve data manipulation and may involve some limited coding (e.g. data visualization tools), and those related to software with a "point and click" user interface (e.g. Microsoft Office applications). I aggregate the composites composed of overall top-ranked skills into interpersonal competencies (e.g. communication and teamwork/collaboration), non-routine cognitive skills that represent general human capital (e.g. problem solving, writing and research), and abilities that might facilitate routine tasks (detail-oriented and organizational skills). For marketing managers only, I define the category of business-related non-routine cognitive skills (e.g. budgeting, business management, and project management). For financial analysts only, I group together composites related to finance and accounting, and also composites related to quantitative analysis (e.g. economics, risk and credit, and statistics and math). Similarly to the composites, the skill category variables are defined as indicators

[12]For financial analysts, I include the top software skills and top overall skills for each occupation code separately, as well as the top skills for two codes combined. As less than five precent of the combined postings are categorized as financial quantitative analysts, the combined rankings are very similar to the rankings for traditional financial analysts.

[13]In this case, a composite may include skills that are not top-ranked for one of the occupations. I also include some software skills that are not top-ranked but are of conceptual interest, such as machine learning and artificial intelligence, and exclude some skills that essentially synonymous with the name of the occupation, such as *marketing management* for marketing managers and *financial analysis* for financial analysts.

[14]The programming composite that I use as the measure of algorithmic technology for financial analysts is used as a predictor in this skill category for marketing managers.

for the job posting requiring at least one of the component skill composites.

Lastly, the analysis also controls for the total number of skills in a post, minimum education and experience requirements, and fixed effects for time measured in months, employer, and Metropolitan Statistical Area (MSA).[15]

In the final section of the paper, I aggregate the job postings data to the level of MSA by year for the period 2010-2017, separately for marketing managers and financial analysts. The skill variables thus indicate the share of job postings in a particular occupation, MSA and year that mention a specific skill category. I then link the aggregated job postings for each occupation to data on hourly wages by MSA and year from the OES, inflated to 2017 dollars. The data include approximately 300 MSA's for each occupation.

3.4 Empirical Strategy

Using the data on job postings, I estimate linear probability models of the form

$$TECH_{itjc} = \Theta' SKILL_{itjc} + \Pi' X_{itjc} + \tau_t + \nu_j + \psi_c + \varepsilon_{itjc}, \tag{3.1}$$

where $TECH_{itjc}$ is an indicator for job vacancy i posted in month t by employer j in MSA c listing algorithmic technology as a skill requirement, $SKILL_{ijtc}$ is a vector of indicators for skill composites referenced in the job posting, X_{ijtc} is a vector of education and experience requirements in the posting, τ_t, ν_j and ψ_c are fixed effects for time measured in months, employer and MSA, respectively, and ε_{itjc} is a mean-zero residual. I cluster standard errors by employer. I estimate the model separately for each of the two occupations, so all terms in Equation 3.1 can be interpreted as being indexed by occupation. I also estimate versions of the model in which I replace the skill composites with a vector of aggregated skill categories.

It is important to note that while I model algorithmic technology as an outcome, the

[15]The education requirement variables consist of indicators for not mentioning an education requirement or requiring less than a bachelor's degree (which is extremely rare for the occupations I study), requiring at minimum a bachelor's degree, and requiring a master's degree or more. The experience requirement variables consist of indicators for not mentioning an experience requirement, explicitly requiring 0 years of experience, requiring 1-5 years, requiring 5-10 years, and requiring 10-15 years.

coefficients I estimate do not have a causal interpretation. The process of changing skill requirements and technology use in a job is highly endogenous, and I cannot shed light on the factors that induce specific firms to make these decisions. However, I choose to model algorithmic technology as an outcome rather than a predictor so that I can assess its relationship with several skill measures simultaneously and avoid concerns of multiple hypothesis testing.

In addition, I estimate ordinary least square (OLS) regressions of the form

$$\ln\bar{w}_{yc} = \Gamma' SKILL_{yc} + \xi_{yc}, \tag{3.2}$$

using the aggregated data on job postings linked to wages, where $\ln\bar{w}_{yc}$ is the log of the mean hourly wage in year y and MSA c in the OES data, $SKILL_{yc}$ is a vector of aggregated skill variables from the job postings data, and ξ_{yc} is a mean-zero residual. Again, I estimate these specifications separately for each occupation.

3.5 Results

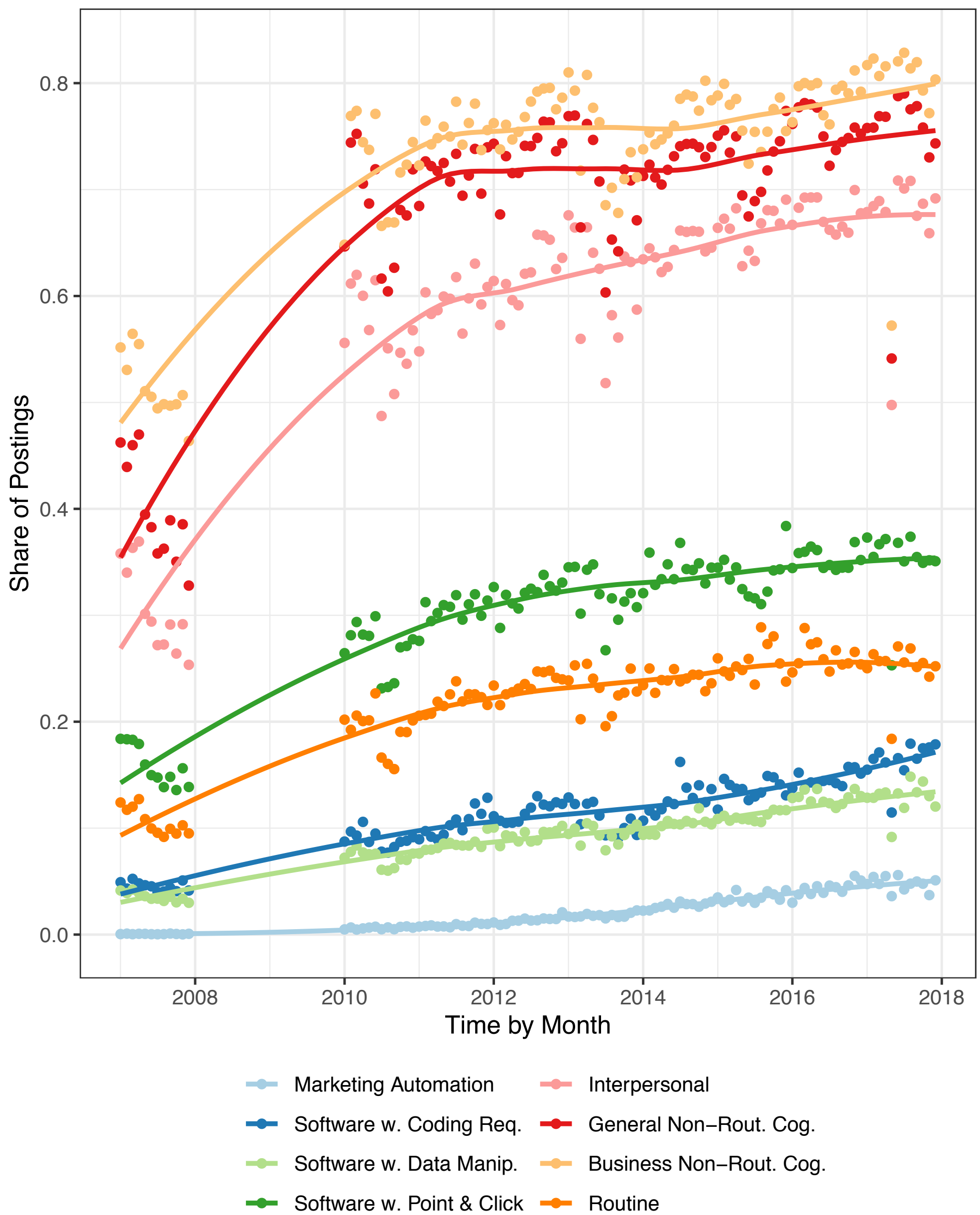
3.5.1 Marketing Managers

Figure 3.3 shows a scatterplot of the proportion of job postings for marketing managers that list marketing automation software and the skill categories as requirements for each month in the data.[16] The lines represent fitted loess curves. The share of postings mentioning marketing automation has grown from less than 0.5 percent in 2010 to around 5 percent by the end of 2017.[17] The share of posts requiring the skill categories has also increased steadily over time, for all categories. Growth is particularly noticeable for the software skill categories and for interpersonal skills.

To assess which skill measures are associated with marketing automation controlling for

[16]Figure C.1 displays a comparable plot for each of the skill composites.

[17]BGT has changed its data processing algorithms over time, likely leading to the distinct trend breaks for most of the skill measures between 2007 and 2010. Due to these changes, more weight should be given to time trends beginning in 2010.



Notes: This figure shows the share of job postings for marketing managers that mention marketing automation software and the skill categories by month in 2007 and 2010-2017 (dots), along with fitted local polynomial regressions (lines).

Figure 3.3: *Marketing Managers - Skill Categories Over Time*

these time trends and other job characteristics, I estimate versions of Equation 3.1. Table C.3 reports the results of this analysis using the skill composites.[18] Column (1) shows coefficients from regressing the marketing automation indicator on the skill composites, controlling only for the total number of skills in the posting and time fixed effects in months. Column (2) adds fixed effects for employer, and clusters standard errors at the employer level. Column (3) adds controls for education and experience requirements. Finally, Column (4) includes MSA fixed effects. Table C.4 shows coefficients from a comparable set of regressions that replace the skill composites with the broader skill categories.

Results are remarkably similar across specifications in Tables C.3 and C.4. Therefore, I focus on my preferred specification in Column (4) in both tables, which includes the full set of fixed effects and education and experience controls. Figures 3.4 and 3.5 plot the coefficients on the skill composites and categories, respectively, from these specifications.

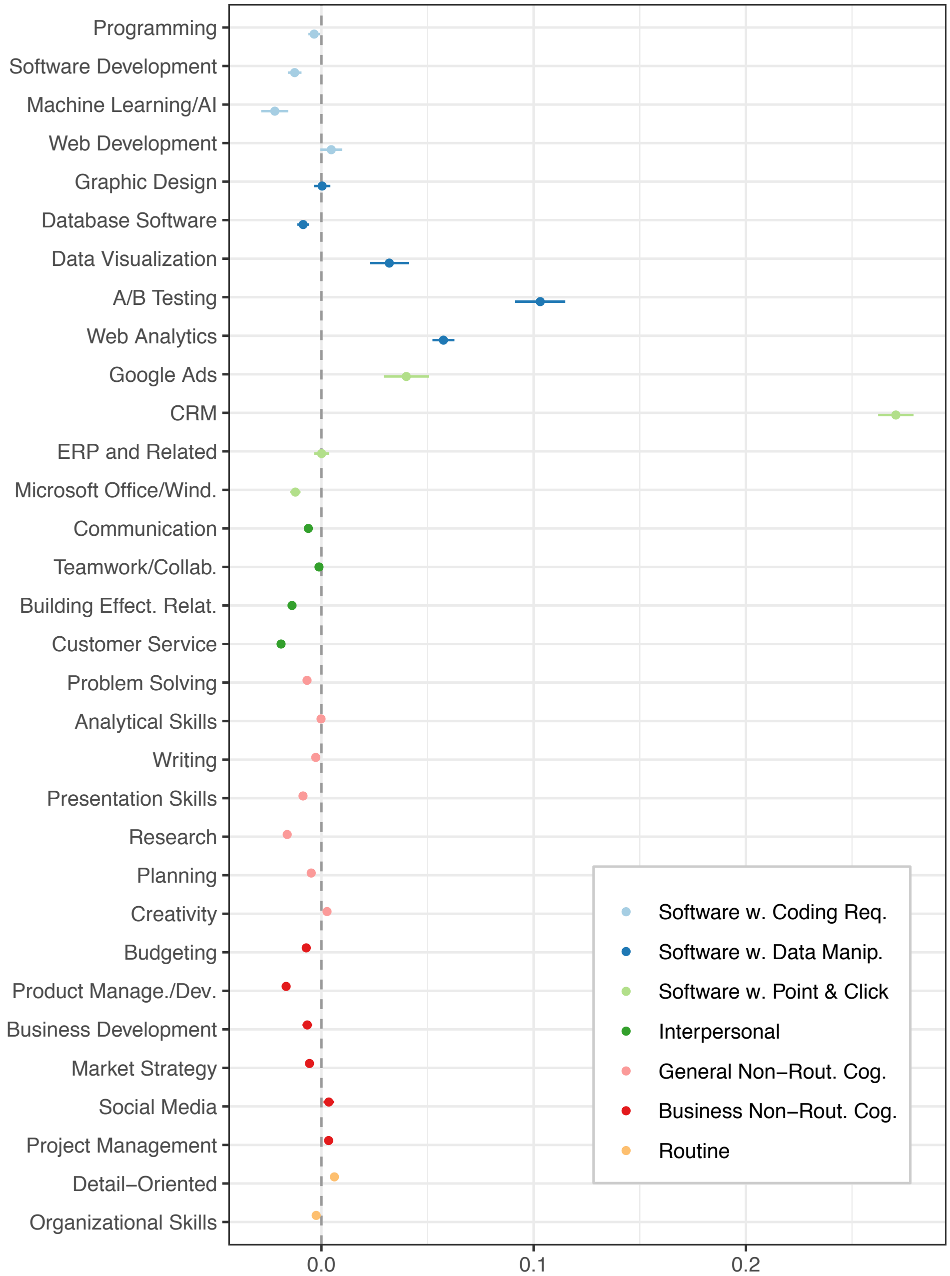
The coefficient plots indicate that the mention of marketing automation technology is negatively predicted by advanced software-related skills that require coding, including programming languages, skills related to software development, and machine learning and artificial intelligence.[19] On average, including at least one software skill that involves coding in a post is associated with a 0.9 percentage point decrease in the probability of mentioning marketing automation technology. Given that 2.4 percent of posts mention marketing automation overall, this coefficient represents a sizable effect.

By contrast, software skills that require limited or no coding are associated with an increase in the probability of a posting listing marketing automation technology of 4.4 and 3.3 percentage points, respectively. In particular, skills related to data visualization, A/B testing, web analytics, Google Ads, and CRM platforms all positively predict the mention of marketing automation technology.[20] CRM software has an especially strong relationship

[18] Table C.2 shows summary statistics for the variables in the regressions in Tables C.3 and C.4.

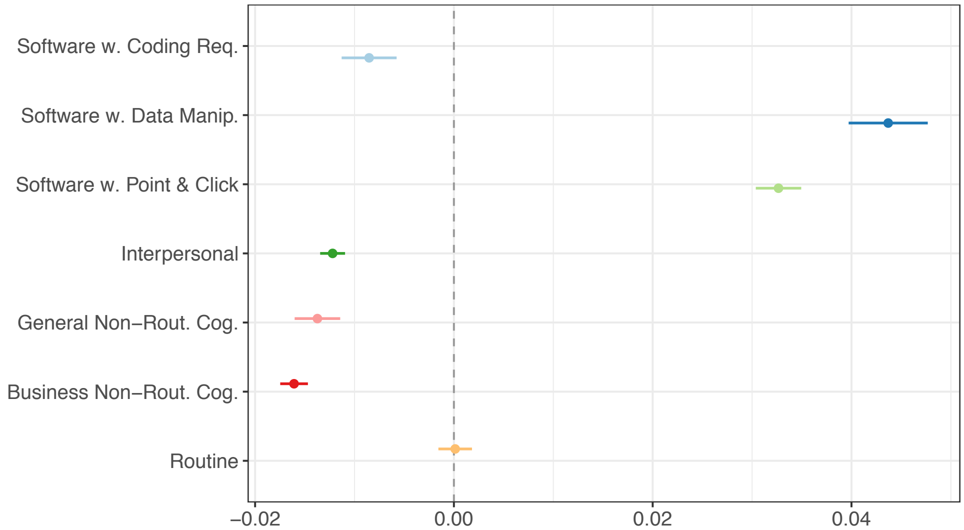
[19] The exception is web development, which has a positive and marginally significant relationship with marketing automation, perhaps because marketing campaigns are often designed designed to increase traffic to a company's website.

[20] One notable exception to this pattern is that the mention of Microsoft Office applications is negatively



Notes: The dots in this figure represent the coefficients on the skill composites from Column 4 of Table C.3, in which the outcome is the marketing automation indicator. The lines represent 95% confidence intervals.

Figure 3.4: *Marketing Managers - Coefficients on Skill Composites*



Notes: The dots in this figure represent the coefficients on the skill categories from Column 4 of Table C.4, in which the outcome is the marketing automation indicator. The lines represent 95% confidence intervals.

Figure 3.5: *Marketing Managers - Coefficients on Skill Categories*

with marketing automation, likely because marketing automation tools are often integrated with a CRM system to enable companies to manage customer relationships in a streamlined fashion.

These results are consistent with the notion that marketing managers follow a *data dashboard* model in their use of marketing automation technology. Job postings that explicitly mention marketing automation are likely to mention other tools that allow marketing managers to identify and interact with potential customers in a data-driven way, as well as measure the success of their efforts. However, marketing managers do not appear to need programming knowledge or expertise in statistics to exploit marketing automation software or related technologies.

Figures 3.4 and 3.5 also indicate that the mention of marketing automation is negatively predicted by the majority of the non-routine cognitive skills that are most frequently

related to marketing automation, perhaps because Office is an older technology that is not associated with "big data" in the way these other tools are.

listed in job postings for marketing managers. These skills include interpersonal abilities such as communication and customer service, general human capital such as problem solving, presentation skills and research, and business- and marketing-related skills such as budgeting, product management and development, and market strategy.[21] Overall, the interpersonal, general non-routine cognitive, and business non-routine cognitive skill categories are each associated with a decrease in the probability of a post mentioning marketing automation of between 1 and 2 percentage points. Finally, the routine skill category is not significantly related to the mention of marketing automation.

These findings do not appear to be consistent with a large literature that suggests that the adoption of information technology is associated with rising demand for non-routine cognitive tasks and associated skills, and the displacement of workers performing routine tasks. However, it is important to note that the non-routine cognitive skills that I consider are among the most frequently required skills in marketing automation jobs, and if anything have become increasingly important in the period under study. For example, over 60 percent of job postings for marketing managers overall ask for interpersonal skills. Thus I interpret the negative coefficients on these skills to suggest that on the margin, employers who explicitly mention marketing automation software put relatively less weight on non-routine cognitive competencies, and more weight on complementary technical tools.

3.5.2 Financial Analysts

Figure 3.6 shows a scatterplot and fitted loess curves for the proportion of job postings for financial analysts that list programming languages and the skill categories as requirements for each month in the data, comparable to Figure 3.3.[22] The share of postings for financial analysts that mention programming rises moderately over time from 7.7 percent in 2010 to 10 percent in 2017, as does the share of financial analysts categorized as financial quantitative

[21] Exceptions are creativity, social media, and project management skills, which positively predict the mention of marketing automation.

[22] Figure C.2 displays a comparable plot for each of the skill composites.

analysts. The prevalence of most of the skill categories appears to increase as well.

Tables C.6 and C.7 present results from a set of linear probability models in which the indicator for programming skills in the financial analyst data is regressed on the skill composites and skill categories, respectively, comparable to the estimates for marketing managers in Tables C.3 and C.4.[23] Figures 3.7 and 3.8 display coefficients from my preferred specifications in Column (4) of each table, which include the full set of fixed effects and controls for education and experience requirements.

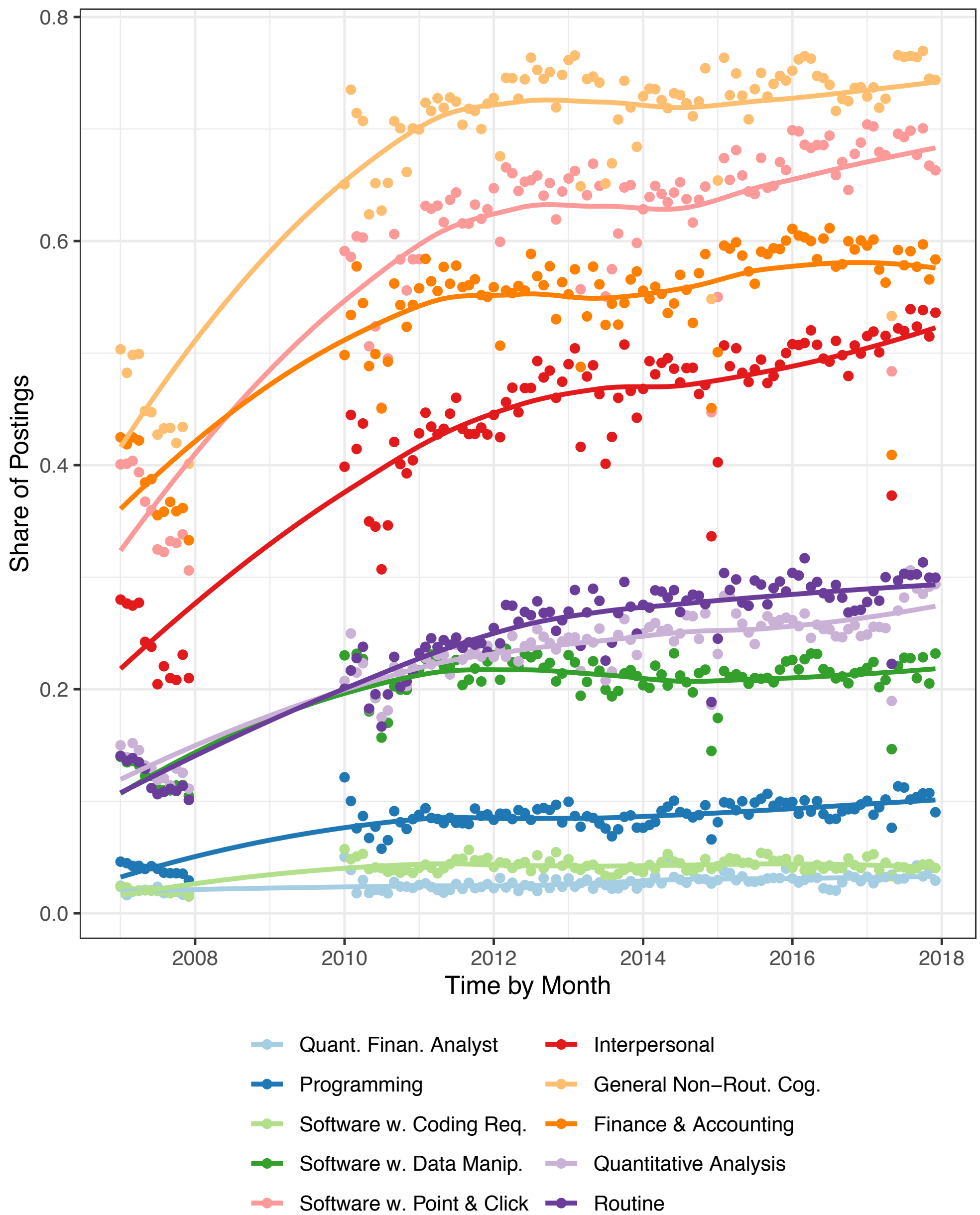
The coefficient plots indicate that, as might be expected, the mention of programming skills in postings for financial analysts is positively predicted by other skills related to coding, as well as skills that involve data manipulation and limited coding, such as Excel formulas and database software. These skill categories are associated with an increase in the probability of mentioning programming of 20 and 15 percentage points, respectively. However, "point and click" software including financial and accounting software is associated with a 2 percentage point decrease in the likelihood of listing programming skills.

In addition, the mention of programming skills is positively predicted by the mention of quantitative analytical skills such as economics, risk and credit, derivatives, and statistics and math. Listing these skills is associated with a 5 percentage point increase in the probability of a financial analyst posting requiring programming skills.

These results are consistent with the notion that financial analysts use algorithmic technology in a way that is similar to other *analyst* occupations in the sciences or math and computing. The use of sophisticated computational and statistical techniques is a core element of these jobs, and thus workers are likely to simply incorporate methodological advances into their analytical toolkit.

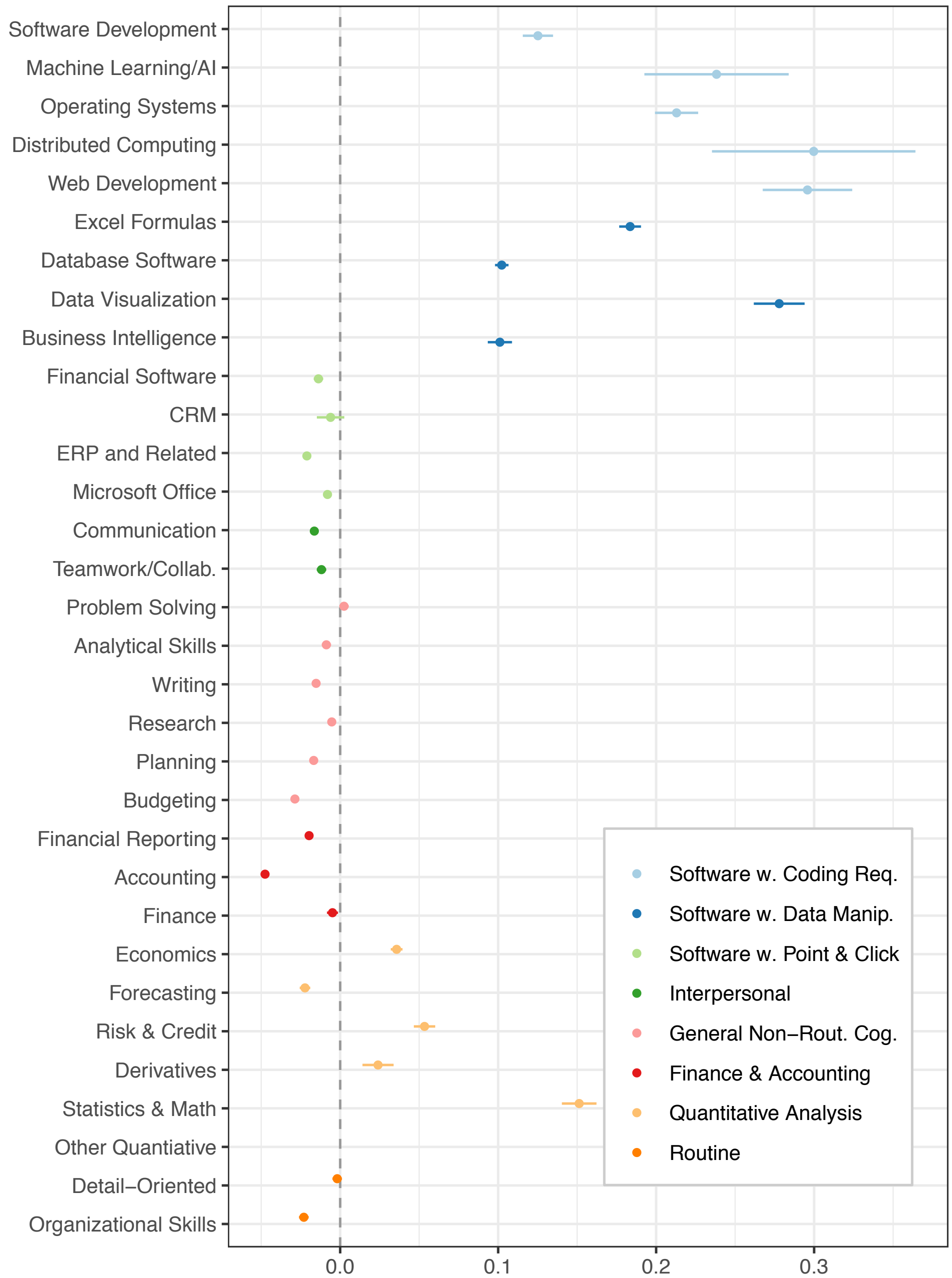
Lastly, I find that the mention of programming skills is negatively predicted by interpersonal, general non-routine cognitive and generic finance and accounting skills, similar to the results for marketing managers. Specifically, the mention of the interpersonal and general non-routine cognitive skill categories are each associated with approximately a 2 percentage

[23]Table C.5 shows summary statistics for the variables in the regressions in these tables.



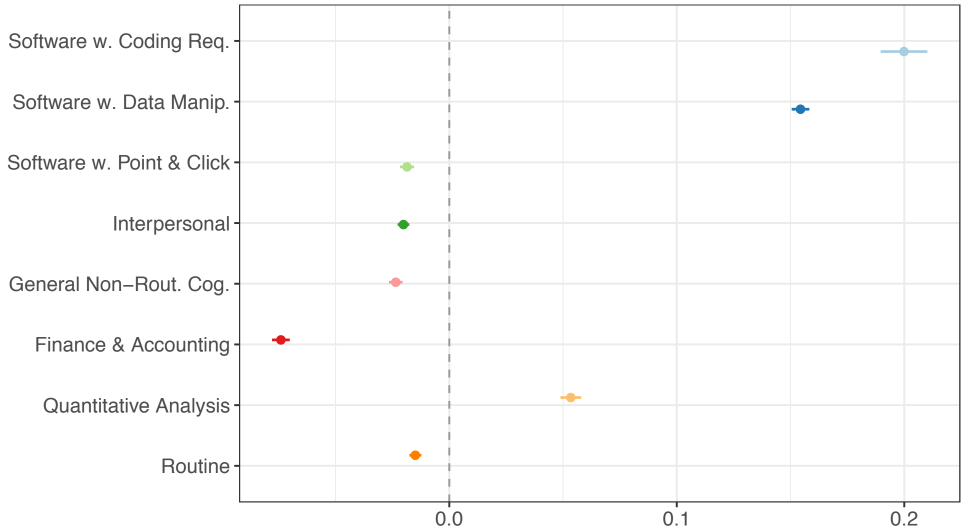
Notes: This figure shows the share of job postings for financial analysts that are quantified as financial quantitative analysts, the share that mention programming and the skill categories by month in 2007 and 2010-2017 (dots), along with fitted local polynomial regressions (lines).

Figure 3.6: *Financial Analysts - Skill Categories Over Time*



Notes: The dots in this figure represent the coefficients on the skill composites from Column 4 of Table C.6, in which the outcome is the programming indicator. The lines represent 95% confidence intervals.

Figure 3.7: *Financial Analysts - Coefficients on Skill Composites*



Notes: The dots in this figure represent the coefficients on the skill categories from Column 4 of Table C.7, in which the outcome is the programming indicator. The lines represent 95% confidence intervals.

Figure 3.8: *Financial Analysts - Coefficients on Skill Categories*

point decrease in the probability of listing programming skills, while the mention of finance and accounting is associated with a 7.4 percentage point decrease.[24]

3.5.3 Wages

Tables 3.1 and 3.2 report estimates of Equation 3.2 for marketing managers and financial analysts, respectively. Specifically, Table 3.1 displays coefficients from regressions of the log of mean wages by MSA and year in the OES data on aggregated skill variables from the job postings data for marketing managers. Column (1) includes only the measure of marketing automation and the average total number of skills in each posting. Column (2) adds the three skill categories related to software, while Column (3) adds the remaining skill categories.

In all specifications, the mention of marketing automation software in the job postings

[24]The mention of the routine skill category is also associated with a decrease in the probability of mentioning programming skills of 1.5 percentage points.

Table 3.1: *Marketing Managers - Wages by MSA & Year Regressed on Skills*

	Outcome:		
	Log of Mean Hourly Wage		
	(1)	(2)	(3)
Marketing Automation	0.385^{**}	0.385^{*}	0.473^{**}
	(0.146)	(0.150)	(0.146)
Total # of Skills	0.016^{**}	0.015^{**}	-0.003
	(0.001)	(0.002)	(0.003)
Software w. Coding Req.		0.378^{**}	0.292^{**}
		(0.065)	(0.064)
Software w. Data Manip.		-0.133^{+}	-0.043
		(0.071)	(0.070)
Software w. Point & Click		-0.035	-0.015
		(0.042)	(0.041)
Interpersonal			-0.019
			(0.040)
General Non-Rout. Cog.			0.201^{**}
			(0.045)
Business Non-Rout. Cog.			0.293^{**}
			(0.041)
Routine			-0.055
			(0.043)
Observations	1,957	1,957	1,957
R^2	0.079	0.095	0.148
Adjusted R^2	0.078	0.093	0.144

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table shows results from OLS regressions of the log of mean hourly wages in 2017 dollars by Metropolitan Statistical Area (MSA) and year on skill measures as indicated from the job postings data for marketing managers. The analysis uses the job postings data for the years 2010-2017, aggregated to the level of MSA by year. The data on wages are from the Occupational Employment Statistics (OES) metropolitan area estimates for the years 2010-2017.

Table 3.2: *Financial Analysts - Wages by MSA & Year Regressed on Skills*

	Outcome:		
	Log of Mean Hourly Wage		
	(1)	(2)	(3)
Programming	0.235**	0.198**	0.175**
	(0.061)	(0.064)	(0.065)
Total # of Skills	0.009**	0.001	−0.002
	(0.002)	(0.002)	(0.002)
Software w. Coding Req.		−0.089	−0.068
		(0.084)	(0.084)
Software w. Data Manip.		0.032	0.028
		(0.042)	(0.042)
Software w. Point & Click		0.232**	0.227**
		(0.032)	(0.033)
Interpersonal			−0.007
			(0.031)
General Non-Rout. Cog.			0.050
			(0.037)
Finance & Accounting			0.0004
			(0.033)
Quantitative Analysis			0.138**
			(0.037)
Routine			0.018
			(0.036)
Observations	1,955	1,955	1,955
R^2	0.029	0.057	0.065
Adjusted R^2	0.028	0.055	0.061

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table shows results from OLS regressions of the log of mean hourly wages in 2017 dollars by Metropolitan Statistical Area (MSA) and year on skill measures as indicated from the job postings data for financial analysts. The analysis uses the job postings data for the years 2010-2017, aggregated to the level of MSA by year. The data on wages are from the Occupational Employment Statistics (OES) metropolitan area estimates for the years 2010-2017.

data is positively related to wages. The estimate in Column (3) indicates that a 10 percentage point increase in the share of postings listing marketing automation is associated with a 5 percent increase in wages across MSA's and years. Interestingly, I find that software skills that require coding are also positively correlated with wages, as are general and business non-routine cognitive skills. These results suggest that while the mention of algorithmic technology is positively related to wages, so are other common skill categories that I do not find to be complementary to this technology.[25]

Table 3.2 reports estimates from a comparable set of specifications for financial analysts. Similar to the results for marketing managers, programming skills are positively related to wages in all columns. The coefficient in Column (3) indicates that a 10 percentage point increase in the share of postings listed programming skills is associated with a 2 percent increase in wages. I also find that wages are positively predicted by point and click software skills and quantitative analysis skills, although not by the software skill categories that I find to be complementary to programming.

Overall, these results suggest that the skill measures I construct using the job postings data for marketing managers and financial analysts have some power in explaining variation in wages for these occupations across time and space. In addition, I find that the mention of algorithmic technology positively predicts wages.

3.6 Conclusion

This paper explores how skill requirements in job postings for two cognitive occupations—marketing managers and financial analysts—are related to technology that facilitates data-driven decision-making. I define a set of skill composites separately by occupation, and estimate the probability of mentioning algorithmic technology as a function of the skill measures, controlling for time, employer and geographic area fixed effects.

[25]By contrast, the coefficients on software skills that involve limited or no coding are negative and insignificant, suggesting that these skills are not associated with wages when marketing automation is included in the regression.

I find that the mention of algorithmic technology is positively associated with complementary technical skills, but that the nature of these skills varies by occupation. My results suggest that marketing managers are an example of a *data dashboard* occupation in which workers use simple and sometimes complex algorithms to automate tasks and improve the quality of decision-making, but do so with the assistance of software platforms with a user-friendly, no coding interface. By contrast, financial analysts are an example of an *analyst* occupation in which all workers are quantitative and the most sophisticated are those who incorporate cutting-edge computation and statistical techniques into their analytical toolkit. Notably, in both occupations the mention of algorithmic technology is negatively predicted by frequently-required non-routine cognitive skills, including interpersonal abilities and general human capital.

I also find that the mention of these algorithmic technologies is positively associated with realized worker wages across geographic area and time for both occupations. These results suggest that data from online job postings can be valuable in understanding how technology use is related to skill requirements and wages in cognitive occupations.

References

Abraham, K. and Amaya, A. (2018). *Probing for informal work activity*. Working Paper 24880, National Bureau of Economic Research, Cambridge, MA.

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In O. C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4B, Elsevier, pp. 1043–1171.

— and — (2012). What does human capital do? A review of Goldin and Katz's "The Race between Education and Technology". *Journal of Economic Literature*, **50** (2), 426–463.

— and Restrepo, P. (2018). The race between machine and man: Implications of technology for growth, factor shares, and employment. *American Economic Review*, **108** (6), 1488–1542.

— and — (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, **128** (6), 2188–2244.

Agrawal, A., Gans, J. and Goldfarb, A. (2019). Artificial intelligence: The ambiguous labor market impact of automating prediciton. *Journal of Economic Perspectives*, **33** (2), 31–50.

Akerlof, G. and Kranton, R. (2000). Economics and identity. *Quarterly Journal of Economics*, **115** (3), 715–753.

Akerlof, G. A. and Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, **105** (2), 255–283.

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

Altonji, J. G. and Blank, R. M. (1999). Race and gender in the labor market. In O. C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 3, Elsevier, pp. 3143–3259.

— and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, **116** (1), 313–350.

Autor, D. (2014a). *Polanyi's paradox and the shape of employment growth*. Working Paper 20485, National Bureau of Economic Research, Cambridge, MA.

— (2014b). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, **344** (6186), 843–852.

— and Dorn, D. (2013a). The growth of low skill service jobs and the polarization of the U.S. labor market. *American Economic Review*, **103** (5), 1553–1597.

—, Katz, L. and Kearney, M. (2006). The polarization of the U.S. labor market. *The American Economic Review*, **96** (2), 189–194.

—, Levy, F. and Murnane, R. (2002). Upstairs, downstairs: Computers and skills on two floors of a large bank. *Industrial and Labor Relations Review*, **55** (3), 432–447.

—, — and — (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, **118** (4), 1279–1333.

Autor, D. H. and Dorn, D. (2013b). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, **103** (5), 1553–1597.

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

Baker, M. and Cornelson, K. (2018). Gender-based occupational segregation and sex differences in sensory, motor, and spatial aptitudes. *Demography*, **55** (5), 1749–1775.

Barach, M. A., Golden, J. M. and Horton, J. J. (forthcoming). Steering in online markets: The role of platform incentives and credibility. *Management Science*.

— and Horton, J. (2020). *How do employers use compensation history? Evidence from a field experiment*. Working Paper 26627, National Bureau of Economic Research.

Bartel, A., Casey, I. and Shaw, K. (2003). "New technology" and its impact on the jobs of high school educated workers: A look deep inside three manufacturing industries. In E. Appelbaum, A. Bernhardt and R. Murnane (eds.), *Low-Wage America: How Employers Are Reshaping Opportunity in the Workplace*, Russell Sage Foundation.

—, Ichniowski, C. and Shaw, K. (2004). Using "insider econometrics" to study productivity. *American Economic Review: Papers and Proceedings*, **94** (2), 214–223.

—, — and — (2007). How does information technology affect productivity? Plant-level comparisons of production innovation, process improvement, and worker skills. *Quarterly Journal of Economics*, **122** (4), 1721–1758.

Beaudry, P., Green, D. and Sand, B. (2014). The declining fortunes of the young since 2000. *American Economic Review*, **104** (5), 381–386.

—, — and — (2016). The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics*, **34** (S1), S199–S147.

— and Lewis, E. (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.

Bell, A. (2019). *Job amenities & earnings inequality*. Job Market Paper, Harvard University.

Bertrand, M. (2011). New perspectives on gender. In O. C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4B, Elsevier, pp. 1543–1590.

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

Blau, F., Brummund, P. and Liu, A. (2013). Trends in occupation segregation by gender 1970-2009: Adjusting for the impact of changes in the occupational coding system. *Demography*, **50** (471-492).

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

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, **4**, 436–455.

Borghans, L., ter Weel, B. and Weinberg, B. A. (2014). People skills and the labor-market outcomes of underrepresented groups. *Industrial & Labor Relations Review*, **67** (2), 287–334.

Breiman, L. (2001). Random forests. *Machine Learning*, **45** (5-32).

Brown, C. (1980). Equalizing differences in the labor market. *Quarterly Journal of Economics*, **94** (1), 113–134.

Brynjolfsson, E. and McAfee, A. (2011). *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversably transforming employment and the economy*. Lexington, MA: Digital Frontier Press.

— and — (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York, NY: W. W. Norton & Company.

Buhrmester, M., Kwang, T. and Gosling, S. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, **6** (1), 3–5.

Caldwell, S. and Danieli, O. (2018). *Outside options in the labor market*. Working paper.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics*. Cambridge, UK: Cambridge University Press.

Catherine, W. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, **96** (4), 849–861.

Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *Quarterly Journal of Economics*, **129** (4), 1625–1660.

Cortes, G. M., Jaimovich, N. and Siu, H. (2018). *The end of men and the rise of women in the high-skilled labor market*. Working Paper 24274, National Bureau of Economic Research, Cambridge, MA.

Cortes, P. and Pan, J. (2018). Occupation and gender. In S. L. Averett, L. M. Argys and S. D. Hoffman (eds.), *Oxford Handbook of Women and the Economy*, Oxford University Press.

— and — (2019). When time binds: Substitutes for household production, returns to working long hours, and the skilled gender wage gap. *Journal of Labor Economics*, **37** (2), 351–398.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, **47** (2), 448–474.

Datta, N. (2019). *Willing to pay for security: A discrete choice experiment to analyse labour supply preferences*. Working Paper 1632, Centre for Economic Performance, London School of Economics and Political Science, London.

Deming, D. and Kahn, L. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, **36** (S1), 337–369.

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

Denning, J., Jacob, B., Lefgren, L. and vom Lehn, C. (2019). *The return to hours worked within and across occupations: Implications for the gender wage gap*. Working Paper 25739, National Bureau of Economic Research, Cambridge, MA.

Dillender, M. and Forsythe, E. (2019). *White collar technological change: Evidence from job posting data*. Working paper, W.E. Upjohn Institute for Employment Research.

England, P. (2005). Emerging theories of care work. *Annual Review of Sociology*, **31**, 381–399.

Felton, E., Raj, M. and Seamans, R. (2019). *The effect of artificial intelligence on human labor: An ability-based approach*. Working paper.

Filippas, A., Horton, J. and Golden, J. (2020). *Reputation inflation*. Working paper.

Flory, J., Leibbrandt, A. and List, J. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job-entry decisions. *Review of Economic Studies*, **82** (1), 122–155.

Folbre, N. (ed.) (2012). *For Love and Money: Care Provision in the United States*. New York: Russell Sage Foundation.

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

Frank, R. (1984). Are workers paid their marginal products? *American Economic Review*, **74** (4), 549–571.

Freeman, R. and Medoff, J. (1984). *What do unions do*. New York, NY: Basic Books.

Frey, C. B. and Osborne, M. (2013). *The future of employment: How susceptible are jobs to computerisation?* Working paper, University of Oxford.

Gartner (2020). *Marketing technology survey 2019: Marketers boost martech efficacy through disciplined planning and collaboration with IT*. Report.

GNEEZY, U. and LIST, J. A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, **74** (5), 1365–1384.

GOLDIN, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, **104** (4), 1091–1119.

— and KATZ, L. (2007). *The race between education and technology: The evolution of U.S. educational wage differentials, 1890-2005*. Working Paper 12984, National Bureau of Economic Research, Cambridge, MA.

— and — (2016). The most egalitarian of all professions: Pharmacy and the evolution of a family-friendly occupation. *Journal of Labor Economics*, **34** (3), 705–746.

—, — and KUZIEMKO, I. (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, **20** (4), 133–156.

GOOS, M., MANNING, A. and SALOMONS, A. (2014). Explaining job polarization: Routine-biased technological change and offshoring. *American Economic Review*, **104** (8), 2509–2526.

GROVE, W., HUSSEY, A. and JETTER, M. (2011). The gender pay gap beyond human capital: Heterogeneity in noncognitive skills and in labor market tastes. *Journal of Human Resources*, **46** (4), 827–874.

HERSHBEIN, B. and KAHN, L. (2018). Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. *American Economic Review*, **108** (7), 1737–1772.

HORTON, J. (2020). *No evidence of rent-sharing in wages: The ruble collapse in an online labor market*. Working paper, MIT Sloan.

HORTON, J. J. (2010). Online labor markets. *Internet and Network Economics: 6th International Workshop, WINE 2010, Stanford, CA, USA, December 13-17, 2010. Proceedings.*

— (2017). The effects of algorithmic labor market recommendations: Evidence from a field experiment. *Journal of Labor Economics*, **35** (2), 345–385.

HOXBY, C. and LEIGH, A. (2004). Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States. *American Economic Review: Papers and Proceedings*, **94** (2), 236–240.

KATZ, L. and AUTOR, D. (1999). Changes in the wage structure and earnings inequality. *Handbook of Labor Economics*, **3**, 1463–1555.

— and KRUEGER, A. (2019). *Understanding trends in alternative work arrangements in the United States*. Working Paper 25425, National Bureau of Economic Research, Cambridge, MA.

— and MURPHY, K. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *Quarterly Journal of Economics*, **107** (1), 35–78.

KLEINBERG, J., LUDWIG, J., MULLAINATHAN, S. and SUNSTEIN, C. (2018). Discrimination in the age of algorithms. *Journal of Legal Analysis*, **10** (113-174).

Krueger, A. B. and Summers, L. H. (1988). Efficienc wages and the inter-industry wage structure. *Econometrica*, **56** (2), 259–293.

Kuhn, P. and Shen, K. (2013). Gender discrimination in jobs ads: Evidence from China. *Quarterly Journal of Economics*, **128** (1), 287–336.

— and Villeval, M. C. (2015). Are women more attracted to co-operation than men? *Economic Journal*, **125** (582), 115–140.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, **25** (1), 1–35.

Lazear, E. (1989). Pay equality and industrial politics. *Journal of Political Economy*, **97** (3), 561–580.

Levy, F. (2008). Computers and the supply of radiology services: Anatomy of a disruptive technology. *Journal of the American College of Radiology*, **5** (10), 1067–1072.

Lordan, G. and Pischke, J.-S. (2018). *Does Rosie like riveting? Male and female occupational choices*. Working paper, London School of Economics and Political Science.

Macneil, I. (1974). The many futures of contracts. *Southern California Law Review*, **47** (691-816).

Maestas, N., Mullen, K., Powell, D., von Wachter, T. and Wenger, J. (2018). *The value of working conditions in the United States and implications for the structure of wages*. Working Paper 25204, National Bureau of Economic Research, Cambridge, MA.

Manning, A. (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, **107** (12), 3722–3759.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142**, 698–714.

McKinsey (2019). *Beyond the Rubicon: North American asset management in an era of unrelenting change*. Report.

Mincer, J. (1974). *Schooling, experience, and earnings*. Columbia University Press.

Mortensen, D. T. (2011). Markets with search friction and the DMP model. *American Economic Review*, **101** (4), 1073–1031.

Mullainathan, S. and Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, **31** (2), 87–106.

Ngai, R. and Petrongolo, B. (2017). Gender gaps and the rise of the service economy. *American Economic Journal: Macroeconomics*, **9** (4), 1–44.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, **122** (3), 1067–1101.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, **14**, 693–709.

Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review*, **104** (11), 3565–3599.

Pan, J. (2015). Gender segregation in occupations: The role of tipping and social interactions. *Journal of Labor Economics*, **33** (2), 365–408.

Pinker, S. (2008). *The sexual paradox: Troubled boys, gifted girls and the real difference between the sexes*. New York: Macmillan.

Remus, D. and Levy, F. (2017). Can robots be lawyers? Computers, lawyers, and the practice of law. *Georgetown Journal of Legal Ethics*, **30** (3).

Romer, D. (1992). Why do firms prefer more able workers? *Working Paper*.

Rosen, S. (1981). The economics of superstars. *American Economic Review*, **71** (5), 845–858.

— (1986). The theory of equalizing differences. In O. C. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, vol. 1, Elsevier, pp. 641–692.

Rothwell, J. (2014). *Still searching: Job vacancies and STEM skills*. Working paper, Brookings Institution, Washington, DC.

Roy, A. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Paper*, **3** (2), 135–146.

Sasser Modestino, A., Shoag, D. and Ballance, J. (2016). Downskilling: changes in employer skill requirements over the business cycle. *Labour Economics*, **41**, 333–347.

—, — and Ballance, J. (forthcoming). Upskilling: Do employers demand greater skill when workers are plentiful? *Review of Economics and Statistics*.

Snowberg, E. and Yariv, L. (2018). *Testing the waters: Behavior across participant pools*. Working Paper 24781, National Bureau of Economic Research, Cambridge, MA.

Su, R., Rounds, J. and Armstrong, P. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, **135** (6), 859–884.

Terviö, M. (2008). The difference that CEOs make: An assignment model approach. *American Economic Review*, **98** (3), 642–668.

Terviö, M. (2009). Superstars and mediocrities: Market failure in the discovery of talent. *The Review of Economic Studies*, **76** (2), 829–850.

Wasserman, M. (2019). *Hours constraints, occupational choice and fertility: Evidence from medical residens*. Working paper, UCLA Anderson School of Management.

Webb, M. (2020). *The impact of artificial intelligence on the labor market*. Working paper, Stanford University.

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

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

Wiswall, M. and Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *Quarterly Journal of Economics*, **133** (1), 457–507.

Zoega, G. and Booth, A. L. (2005). *Worker heterogeneity, intra-firm externalities and wage compression*. Tech. rep., Birkbeck University of London.

Appendix A

Appendix to Chapter 1

A.1 Data Appendix

A.1.1 Selecting Gender-Typical Tasks

This section describes the process of selecting the gender-typical job tasks that are examined in the hypothetical choice experiment.

ACS Data

I use data from the Integrated Public Use Microdata five-year ACS sample for 2012-2016 to measure the occupational female share. I restrict the pooled sample to workers aged 18 and older who are currently employed and are not classified as unpaid family workers. I calculate the proportion of workers who are female in each of the ACS occupation codes, weighting by the ACS person weight.

I also use the ACS data to measure the mean log hourly wage in each occupation, the mean usual hours of work per week in each occupation, and the share of workers in each occupation with at least a four-year college degree. To calculate hourly wages, I divide wage and salary income from the last 12 months by annual hours of work. I measure annual hours of work as usual hours of work per week multiplied by weeks worked in the last 12 months. Beginning in 2008, ACS respondents report weeks worked in intervals (1-13 weeks, 14-26 weeks, 27-39 weeks, 40-47 weeks, 48-49 weeks and 50-52 weeks). Therefore, I use data from the 2005-2007 ACS to impute the actual number of weeks worked in the 2012-2016 sample as mean weeks worked among respondents within each interval in the earlier period. I inflate wages to 2018 dollars using the Consumer Price Index from the Bureau of Labor Statistics (BLS), and exclude hourly wages observations that are less than $3 and greater than $200.

O*NET Data

I use data from the O*NET 21.1, released in November 2016, to measure the frequency of task performance by occupation. I focus on variables from the *generalized work activities* and *work styles* domains, which include 41 and 16 elements, respectively, that I interpret as providing information about categories of job tasks. The surveys on which the O*NET data are based ask workers to indicate the importance of each work activity and work style element to the performance of their current job, on a scale of 1 to 5. The published O*NET data report the mean importance value for each element in each occupation.

The O*NET surveys also asks respondents to report the level of each work activity required to perform their current job, on a scale of 0 to 7. I use measures of the importance of work activities rather than the level for consistency across domains, and because I interpret importance to map more closely onto the notion of the frequency with which an activity is performed. By contrast, the examples provided to guide workers in answering the questions about the level of an activity seem explicitly intended to capture information about worker skills and education requirements. For instance, the example for level two of *making decisions and solving problems* is "determine the meal selection for a cafeteria," while the level five example is "make the final decision about a company's 5-year plan." I conduct robustness tests in which I use the level value and the level value multiplied by the importance value (with the variables first rescaled to fall between 0 and 10) for the work activity measures, and find similar results.

Occupational Crosswalk

The O*NET 21.1 uses occupation codes based on the BLS 2010 Standard Occupation Classification (SOC), but with some additional detailed codes. To link the O*NET variables to the female share data from the ACS, I first collapse the data to the mean value within each SOC code. I then use a crosswalk available from the Census Bureau to map each SOC code to an ACS occupation code and collapse the data to the mean value within each ACS code, weighting by the number of workers employed in each SOC occupation in 2012-2016, based on the BLS Occupational Employment Statistics (OES) data. The O*NET data can be matched to 464 ACS occupation codes, out of a total of 478 codes in the ACS data. After applying the crosswalk, I standardize the O*NET measures to have a mean of zero and a standard deviation of one.

Random Forest Analysis

As part of the process of selecting tasks, I use a random forest algorithm (Breiman 2001) to identify O*NET work activities and work styles that are highly predictive of the occupational female share. The outcome in this analysis is the female share, and the predictors are the O*NET variables. I conduct the analysis separately for the two O*NET domains. I choose the tuning parameters via cross-validation, following Mullainathan and Spiess (2017). Specifically, I use eight-fold cross-validation to select the number of trees (400, 500, 600 or 700), the minimum node size to which each tree is grown (3, 5, 7, 10 or 15), and the proportion of predictors available to be chosen at each internal node in each tree (0.2, 0.3 or 0.4).

While the primary purpose of the random forest technique is predictive, the algorithm also calculates two types of variable importance scores, which rank each predictor according to its contribution to prediction accuracy. The first type of variable importance score is calculated using the *permutation method*, which estimates the improvement in prediction accuracy out-of-sample produced by including a variable in the algorithm. Specifically, the permutation method importance score for predictor X_j is calculated as follows: For each tree t, randomly permute the values of X_j in the out-of-bag (OOB) data not used to construct the tree, such that variable X_j becomes pure noise. Calculate the mean squared prediction error (MSPE) generated by 1) running the true OOB data down tree t, and 2) running the permuted OOB data down tree t. Subtract the first MSPE from the second. Average these differences over all trees t in the forest, then divide by the standard deviation of the difference.

The second type of variable importance score is calculated using the *node purity method*, which yields the improvement in prediction accuracy in-sample produced by including a particular variable in the algorithm. Specifically, the node purity importance score for predictor X_j is calculated as follows: For each tree t and each internal node d at which X_j is chosen as the splitting variable, compute the MSPE for cases that reach node d generated by 1) splitting on X_j, with no further branches in the tree, and 2) a constant prediction at node d. Subtract the first MSPE from the second. Sum these differences over all nodes d in tree t, then average over all trees t in the forest.

I designate a variable as highly predictive if it is among the top-scoring variables using either importance score method. Specifically, for the analysis using the work activities domain, I classify a variable as highly predictive if its importance score places it among the top ten variables using either method. For the analysis using the work styles, I classify a variable as highly predictive if it is ranked among the top four variables using either importance score. This methodology results in 11 highly predictive work activities and 4 highly predictive work styles.

Figures A.1 and A.2 display the two types of variable importance scores generated by the random forest analysis for the O*NET work activities and work styles. The importance scores are standardized to have a mean of zero and a standard deviation of one within each domain, and are sorted according to the permutation scores. The measures designated as highly predictive are shown in dark green.

Selecting Tasks

As described in the main text, I use a hybrid quantitative and qualitative approach to select tasks to include in the experiment. Tables A.1 and A.2 show the results of the quantitative analysis for work activities and work styles, respectively. First, I regress the female share on each O*NET measure, standardized to have a mean of zero and standard deviation of one, and rank the coefficients from most positive to most negative (Column 1). Next, I estimate multivariate regressions in which I regress the female share on all O*NET variables in the work activities or work styles domain (Column 2). In Column 3, I regress the female share on the six work activities (two work styles) with the most positive bivariate coefficients and the six work activities (two work styles) with the most negative bivariate coefficients. Finally, in Column 4 I regress the female share on the variables designated as highly predictive by the random forest algorithm.

The results demonstrate that a large proportion of the O*NET measures (29 of 41 work activities and 12 of 16 work styles) have a significant bivariate relationship with the female share. It is also clear that many of the O*NET variables are correlated with each other, as the coefficients change substantially between the bivariate regressions in Column 1 and the multivariate regression in Column 2.

Tables A.3 and A.4 show that results are similar when I repeat this set of regression analyses including controls for occupation cluster, mean log hourly wages in each occupation and the share of workers with a college degree or more in each occupation in the ACS. The occupation clusters, which are constructed to resemble the broad categories in Acemoglu and Autor (2011), are: 1) managerial, professional and technical occupations ("professional"), 2) sales and office and administrative support occupations ("sales and clerical"), 3) production, construction, extraction, transportation, and installation, maintenance and repair occupations ("blue collar"), 4) healthcare support, protective service, food preparation, building and grounds cleaning and maintenance, and personal care occupations ("service"), and 5) agriculture and military occupations. I also find similar results when I repeat the entire analysis, including the random forest algorithm, using the female share based only on workers with and only on workers without a college degree, and using work activity variables that indicate the level rather than importance value and the level multiplied by the importance value, as described above (results not shown).

I follow a qualitative approach to select a final set of gender-typical tasks. Specifically, I look for work activities and work styles that are statistically significant and consistent in sign across multiple specifications in Tables A.1, A.2, A.3 and A.4. I focus on measures that are highly ranked based on the bivariate OLS coefficients, highly predictive in the random forest algorithm, or ideally both. I eliminate some measures that might be difficult for participants to understand, specifically *drafting, layinging out, and specifying equipment* and *estimating quantifiable characteristics*.

The names of the O*NET measures that I choose are displayed in bold in Tables A.1, A.2, A.3 and A.4. In some cases, I combine measures of related activities into a single gender-typical task that I include in the experiment, as documented in Table 1.1. In particular, I combine the work activities *operating vehicles, mechanized devices, or equipment*; *repairing and maintaining mechanical equipment*; and *repairing and maintaining electronic equipment*. I also combine the work styles *social orientation* and *cooperation*.

Descriptive Results for Selected Tasks

Figures A.3 and A.4 display the female share (scaled from 0 to 100) and the mean levels of each of the selected gender-typical tasks by major occupation group. For the tasks that combine multiple O*NET variables, I average the variables to create a single measure.

The tasks are rescaled to represent percentiles weighted by employment in each occupation in the 2012-2016 ACS, including all currently employed individuals aged 18 and older. Thus the weighted mean of each task across all occupations is 50. It is clear that the female-typical (male-typical) tasks are rated as more important in majority-female (majority-male) occupation categories, as expected. For example, *helping and caring for others* is rated as most important in female-dominated health, personal care, social services and education occupations, as well as male-dominated protective service occupations. By contrast, *operating* *and repairing equipment* is rated as most important in male-dominated occupations, including construction, maintenance, transportation and production.

Tables A.5, A.6, A.7, A.8 and A.9 display the ten occupations with the highest and the ten occupations with the lowest levels of each selected task measure, standardized to have a mean of zero and standard deviation of one, along with the female share in each displayed occupation.

A.1.2 Implication for Gender Gaps

This section describes the data used in Table A.17, which shows robustness checks for the results on the implications of the WTP estimates for gender differences in sorting and segregation.

In all data sources described below, I inflate wages to 2018 dollars using the Consumer Price Index from the Bureau of Labor Statistics (BLS), and exclude hourly wages observations that are less than \$3 and greater than \$200. In addition, I use the O*NET measures described above to classify occupations as involving a high or low level of the gender-typical tasks. In each dataset, I rescale the O*NET measures to reflect percentiles weighted by employment, and choose a cutoff percentile such that the share of workers in a high-task k job in that dataset matches the share in a high-task job in the experiment sample.

CPS MORG Data

The CPS is a monthly household survey conducted by the U.S. Census Bureau that is designed to represent the civilian, non-institutional U.S. population. Surveyed households are interviewed for four consecutive months, then excluded from the sample for eight months, then interviewed for an additional four consecutive months. Respondents are only asked to report weekly and hourly earnings in their fourth and eighth months in the survey; the files on these "outgoing" households comprise the MORG data.

The analysis in Panel A of Table A.17 uses data from the 2012-2018 CPS MORG. I restrict the sample to currently employed individuals aged 18-64 who are not self-employed, have valid wages, and work in an occupation that can be matched to the O*NET. To construct the wage measure, I use data on hourly wages for workers who report being paid hourly, and weekly earnings divided by usual hours of work per week for others. I multiply weekly earnings values by 1.5 for individuals who report a top-coded weekly earnings value of $2,884.61 and exclude allocated wage observations, following Acemoglu and Autor (2011). The CPS occupation codes are very similar to the ACS codes, with some additional detailed occupations. Therefore, I use the crosswalk between the SOC codes and the ACS occupation codes to link the O*NET measures to the CPS MORG data, after adjusting the CPS codes to match the ACS codes.

The regression specification that I use to estimate the task wage differentials in the CPS MORG includes controls for race and ethnicity, educational attainment, potential experience and its square, year, region, metropolitan area status and union coverage. The race and ethnicity variables consist of mutually exclusive indicators for Black, Hispanic, and other race, with White non-Hispanic as the omitted category. The education variables consist of indicators for less than a high school diploma, high school diploma, some college, associate degree and graduate degree, with bachelor's degree as the omitted category. Potential experience is measured as age minus years of education minus six. Union coverage is measured as an indicator for being a member of a union or covered by a union or employee association contract. I weight the regression by the CPS earnings weight.

PSID Data

The PSID is a longitudinal survey that follows a representative sample of U.S. households first surveyed in 1968 and their descendants. The analysis in Panel C of Table A.17 uses data from the 2007, 2009 and 2011 PSID waves that are available in the replication package for Blau and Kahn (2017). The sample in the replication package is restricted to observations on individuals aged 25-64 who are not self-employed and have valid data on hourly wages. I also exclude individuals who are missing data on region of residence. The PSID data use occupation codes from the 2000 Census, which I link to the ACS codes and thus to the

O*NET using a set of aggregated Census/ACS occupation codes that are consistent over time. These codes were developed by Autor and Dorn (2013b) and updated by Deming (2017) and for this project.

The regression specification that I use to estimate the task wage differentials in the PSID includes measures constructed by Blau and Kahn of years of full-time work experience, years of part-time work experience, and the squares of full-time and part-time experience. The authors define full-time work as at least 1,500 hours per year, and part-time work as less than 1,500 hours but greater than zero hours. To construct the work experience measures, the authors use the longitudinal structure of the data, and impute missing values as necessary.

The wage regression also includes controls for race and ethnicity, educational attainment, year, region and metropolitan area status. The race and ethnicity variables consist of mutually exclusive indicators for Black, Hispanic, and other race, with White non-Hispanic as the omitted category. The educational attainment variables consist of years of education and indicators for having exactly a bachelor's degree and exactly a graduate degree. I weight the regression by the PSID family weight.

NLSY79

The NLSY79 is a longitudinal study of a nationally representative sample of individuals who were aged 14 to 22 in 1979, the first year of the survey. The analysis in Panel D of Table A.17 uses data from the 1979 to 2016 waves of the NLSY79. I restrict the sample to observations on individuals aged 18-64 with valid data on hourly wages, region of residence, educational attainment, and cognitive, non-cognitive and social skills. The NLSY79 uses occupation codes from the 1970 Census for the waves conducted in 1979 to 1981, occupation codes from the 1980 Census for the 1982 to 2000 waves, and occupation codes from the 2000 Census for the 2002 to 2016 waves. I link the Census occupation codes to the ACS codes and the O*NET using the set of aggregated Census/ACS occupation codes developed by Autor and Dorn (2013b).

The regression specification that I use to estimate the task wage differentials in the NLSY79 includes measures of cognitive, non-cognitive and social skills that I construct following Deming (2017). Specifically, I measure cognitive skill using a standardized version of scores on the Armed Forces Qualifying Test (AFQT) that have been adjusted by Altonji *et al.* (2012) to be comparable across the NLSY79 and NLSY97 (see further information on the NLSY97 below). The measure of non-cognitive skill consists of the standardized mean of scores on the Rotter Locus of Self-Control and Rosenberg Self-Esteem Scale. Finally, the measure of social skill comprises the standardized mean of sociability in adulthood (self-reported in 1985), sociability at age 6 (self-reported retrospectively in 1985), the number of clubs the respondent participated in during high school, and an indicator for whether the respondent participated in high school sports.

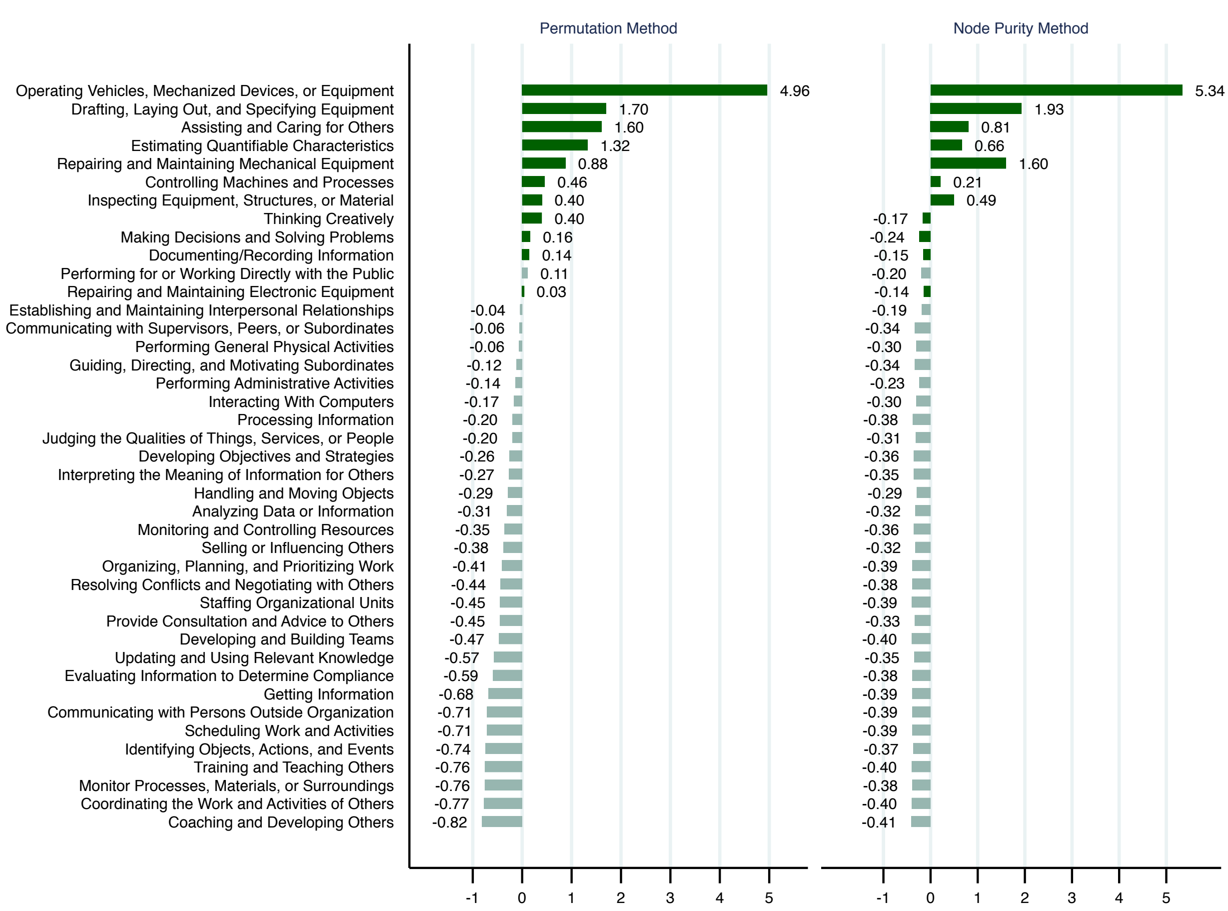
The wage regression also controls for race and ethnicity, educational attainment, potential experience and its square, year, region, metropolitan area status, and residence in an urban versus rural area. The race and ethnicity variables consist of mutually exclusive indicators for Black and Hispanic, with White non-Hispanic as the omitted category. The educational attainment variables consist of years of education and indicators for having exactly a bachelor's degree and exactly a graduate degree. I weight the regression using a weight variable constructed by Altonji *et al.* (2012).

NLSY97

The NLSY97 is a longitudinal survey of a nationally representative sample of individuals aged 12 to 16 at the end of 1996 who were first interviewed in 1997. The analysis in Panel E of Table A.17 uses data from the 1997 to 2015 waves of the NLSY97. I restrict the sample to observations on individuals aged 18-64 with valid data on hourly wages, region of residence, educational attainment, and cognitive, non-cognitive and social skills. The NLSY97 uses occupation codes from the 2000 Census. I link the Census occupation codes to the ACS codes and the O*NET using the set of aggregated Census/ACS occupation codes developed by Autor and Dorn (2013b).

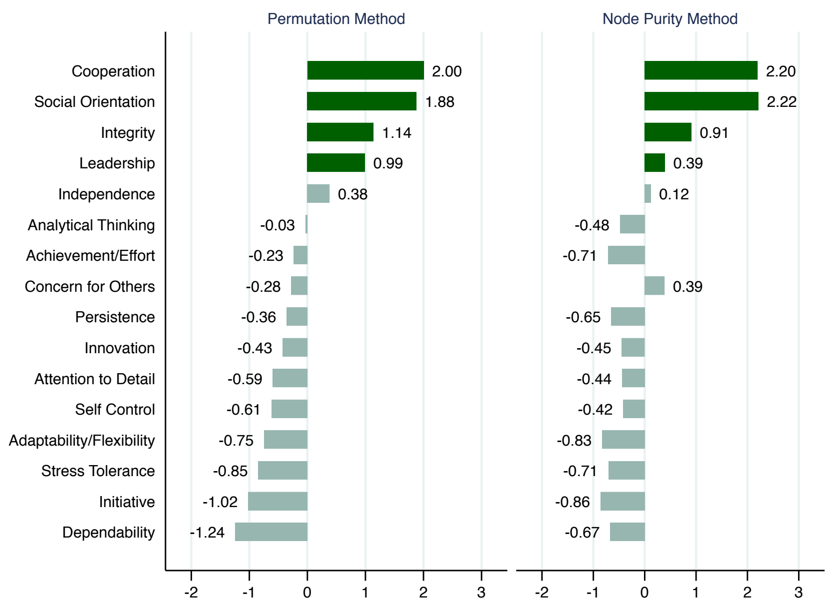
The regression specification that I use to estimate the task wage differentials in the NLSY97 includes measures of cognitive, non-cognitive and social skills that I construct following Deming (2017). Specifically, I measure cognitive skill using a standardized version of scores on the Armed Forces Qualifying Test (AFQT) that have been adjusted by Altonji *et al.* (2012), as described above. The measure of non-cognitive skill is the standardized mean of the self-reported personality traits organization, conscientiousness, dependability, thoroughness, trust, and discipline. Lastly, the measure of social skill comprises the standardized mean of the self-reported personality traits extraversion and animation (i.e. a negative score on a measure of being reserved or quiet).

The wage regression also controls for race and ethnicity, educational attainment, potential experience and its square, year, region, metropolitan area status, and residence in an urban versus rural area. The race and ethnicity variables consist of mutually exclusive indicators for Black and Hispanic, with White non-Hispanic as the omitted category. The education variables consist of indicators for less than a high school diploma, high school diploma, some college, associate degree and graduate degree, with bachelor's degree as the omitted category. I weight the regression using a weight variable constructed by Altonji *et al.* (2012).



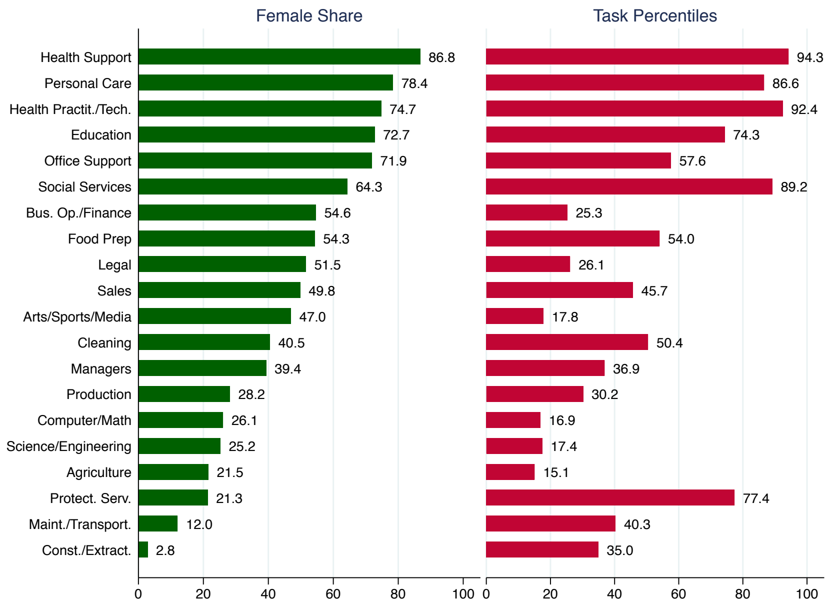
Notes: This figure displays the variable importance scores for the O*NET work activities generated by the random forest analysis. The importance scores are standardized to have a mean of zero and a standard deviation of one and are sorted according to the permutation scores. The measures designated as highly predictive are shown in dark green.

Figure A.1: *Variable Importance Scores - Work Activities*

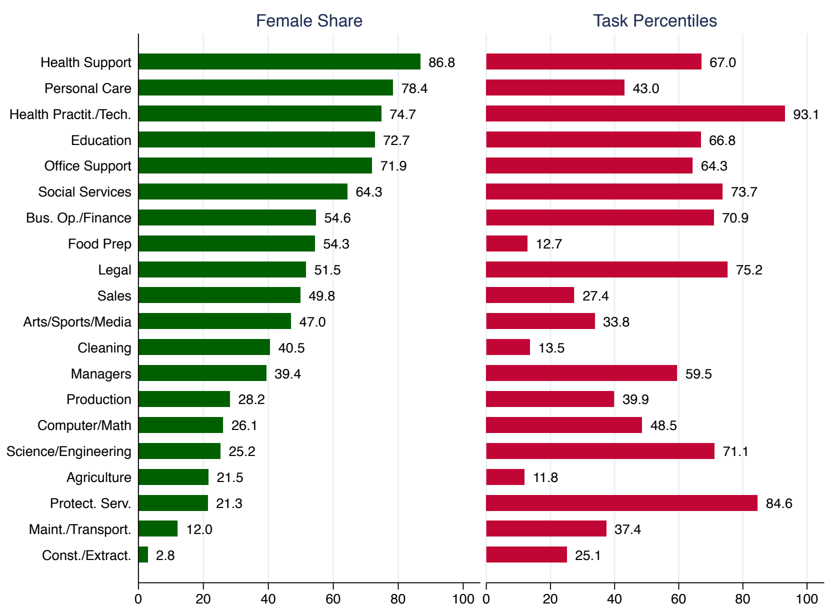


Notes: This figure displays the variable importance scores for the O*NET work styles generated by the random forest analysis. The importance scores are standardized to have a mean of zero and a standard deviation of one and are sorted according to the permutation scores. The measures designated as highly predictive are shown in dark green.

Figure A.2: *Variable Importance Scores - Work Styles*



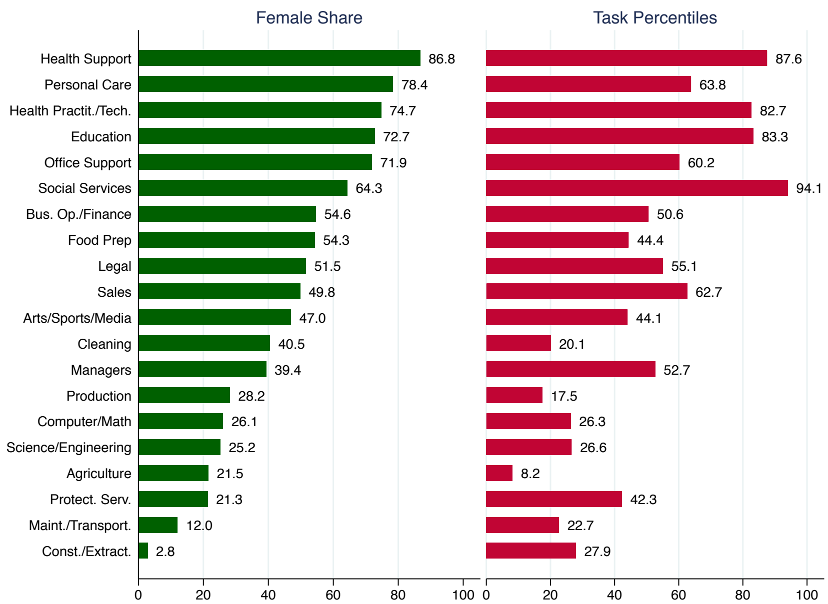
(a) Helping and Caring for Others



(b) Documenting and Recording Information

Figure A.3: *Female-Typical Tasks and Female Share by Occupation*

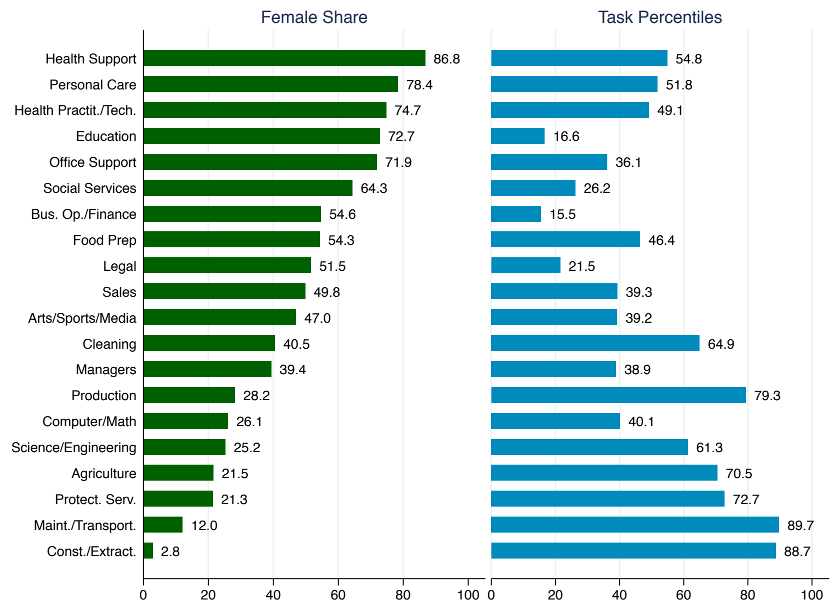
Continued on next page



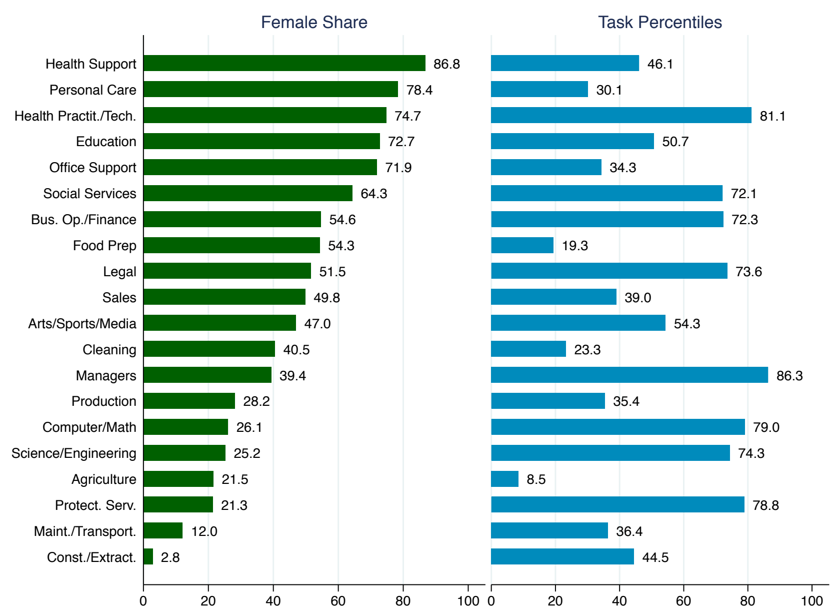
(c) Working and Communicating with Others

Notes: This figure displays the mean occupational female share in the 2012-2016 ACS (scaled from 0 to 100) and the mean levels of the selected female-typical tasks measured using the O*NET variables listed in Table 1.1, by major occupation group. Tasks based on multiple O*NET variables are constructed by averaging the component variables. Each task is rescaled to reflect percentiles weighted by employment in the 2012-2016 ACS.

Figure A.3: *(Continued) Female-Typical Tasks by Occupation*



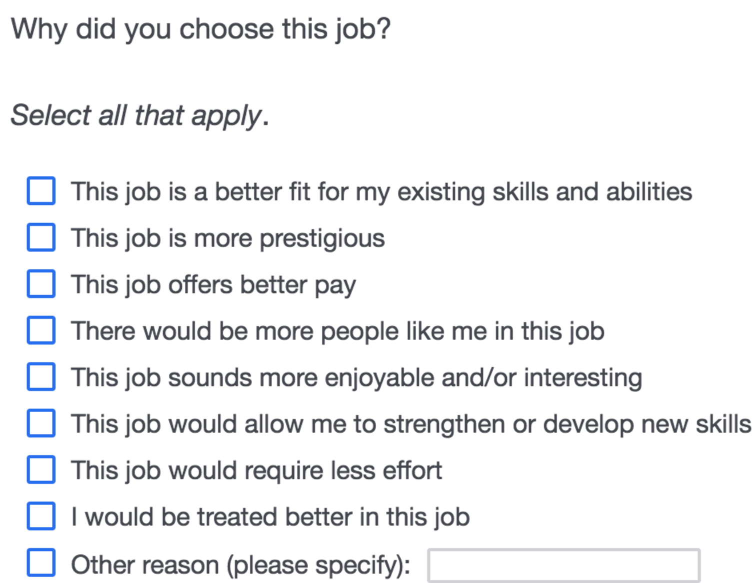
(a) Operating and Repairing Equipment



(b) Making Decisions and Solving Problems

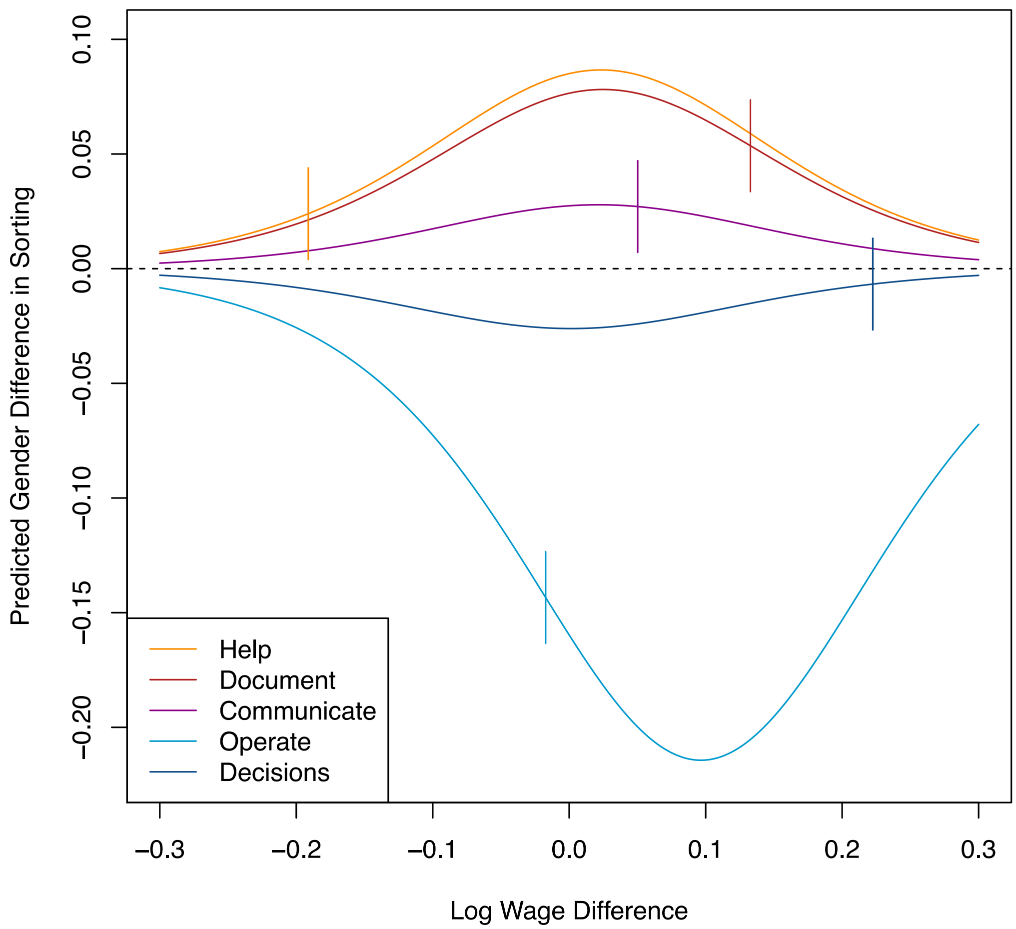
Notes: This figure displays the mean occupational female share in the 2012-2016 ACS (scaled from 0 to 100) and the mean levels of the selected male-typical tasks measured using the O*NET variables listed in Table 1.1, by major occupation group. Tasks based on multiple O*NET variables are constructed by averaging the component variables. Each task is rescaled to reflect percentiles weighted by employment in the 2012-2016 ACS.

Figure A.4: *Male-Typical Tasks and Female Share by Occupation*



Notes: This figure displays the wording of the response options available to experiment participants in the questions about the reasons for the choices made in the hypothetical scenarios. The order of responses is randomized, and participants may make multiple entries.

Figure A.5: *Reasons for Choices - Survey Text*



Notes: This figure plots the predicted gender differences in sorting on the gender-tyipcal tasks ($\hat{Q}_k$) as a function of ω_k, the log difference in wage offers between the high-task and low-task jobs, using the coefficients from the specification in Table 1.10. The vertical lines show the estimated task wage differentials ($\hat{\omega}_k = b_k$) in the ACS from Table 1.9.

Figure A.6: *Predicted Sorting Differences by ω_k*

Table A.1: *Female Share Regressed on Work Activities*

	(1)	(2)	(3)	(4)
Establishing and Maintaining Interpersonal Relationships	0.125**	0.002	-0.030*	
	(0.011)	(0.017)	(0.013)	
Performing Administrative Activities	0.104**	0.027*	0.028*	
	(0.012)	(0.013)	(0.012)	
Assisting and Caring for Others	0.104**	0.070**	0.083**	0.088**
	(0.011)	(0.015)	(0.013)	(0.010)
Performing for or Working Directly with the Public	0.094**	0.023	0.012	
	(0.012)	(0.015)	(0.010)	
Interacting With Computers	0.086**	0.010	-0.025	
	(0.013)	(0.018)	(0.017)	
Documenting/Recording Information	0.077**	0.016	0.031*	0.053**
	(0.013)	(0.014)	(0.013)	(0.009)
Resolving Conflicts and Negotiating with Others	0.067**	0.001		
	(0.013)	(0.016)		
Communicating with Persons Outside Organization	0.065**	-0.013		
	(0.012)	(0.016)		
Interpreting the Meaning of Information for Others	0.061**	0.027		
	(0.012)	(0.017)		
Organizing, Planning, and Prioritizing Work	0.055**	0.036*		
	(0.013)	(0.016)		
Getting Information	0.045**	-0.001		
	(0.013)	(0.016)		
Communicating with Supervisors, Peers, or Subordinates	0.044**	0.013		
	(0.013)	(0.015)		
Processing Information	0.040**	0.000		
	(0.013)	(0.020)		
Updating and Using Relevant Knowledge	0.033**	-0.012		
	(0.012)	(0.018)		
Selling or Influencing Others	0.026*	0.005		
	(0.012)	(0.012)		
Developing and Building Teams	0.020	0.020		
	(0.013)	(0.019)		
Coaching and Developing Others	0.014	-0.007		
	(0.013)	(0.023)		
Staffing Organizational Units	0.013	-0.010		
	(0.012)	(0.015)		
Analyzing Data or Information	0.012	0.008		
	(0.011)	(0.021)		
Provide Consultation and Advice to Others	0.006	-0.029*		
	(0.013)	(0.014)		

Continued on next page

Table A.1 – *Continued from previous page*

	(1)	(2)	(3)	(4)
Training and Teaching Others	-0.003	0.029^{+}		
	(0.014)	(0.015)		
Scheduling Work and Activities	-0.004	0.005		
	(0.012)	(0.015)		
Developing Objectives and Strategies	-0.006	-0.024		
	(0.013)	(0.019)		
Thinking Creatively	-0.011	-0.004		0.001
	(0.012)	(0.016)		(0.012)
Evaluating Information to Determine Compliance	-0.012	-0.002		
	(0.013)	(0.013)		
Coordinating the Work and Activities of Others	-0.018	-0.020		
	(0.012)	(0.016)		
Judging the Qualities of Things, Services, or People	-0.024^{+}	0.014		
	(0.013)	(0.012)		
Monitoring and Controlling Resources	-0.027^{*}	-0.005		
	(0.012)	(0.015)		
Guiding, Directing, and Motivating Subordinates	-0.028^{*}	-0.033^{+}		
	(0.012)	(0.019)		
Making Decisions and Solving Problems	-0.031^{*}	-0.041^{*}		-0.059^{**}
	(0.012)	(0.017)		(0.012)
Identifying Objects, Actions, and Events	-0.050^{**}	-0.042^{**}		
	(0.012)	(0.014)		
Monitor Processes, Materials, or Surroundings	-0.068^{**}	0.040^{**}		
	(0.012)	(0.014)		
Performing General Physical Activities	-0.104^{**}	-0.007		
	(0.012)	(0.025)		
Handling and Moving Objects	-0.110^{**}	0.037		
	(0.011)	(0.026)		
Estimating Quantifiable Characteristics	-0.121^{**}	-0.010		-0.017
	(0.011)	(0.015)		(0.011)
Repairing and Maintaining Electronic Equipment	-0.137^{**}	-0.005	0.004	-0.004
	(0.009)	(0.015)	(0.014)	(0.014)
Controlling Machines and Processes	-0.149^{**}	-0.008	0.039^{*}	0.024
	(0.010)	(0.023)	(0.019)	(0.018)
Inspecting Equipment, Structures, or Material	-0.160^{**}	-0.027	-0.041^{*}	-0.016
	(0.010)	(0.020)	(0.016)	(0.017)
Drafting, Laying Out, and Specifying Equipment	-0.167^{**}	-0.035^{**}	-0.060^{**}	-0.044^{**}
	(0.010)	(0.011)	(0.009)	(0.010)
Repairing and Maintaining Mechanical Equipment	-0.179^{**}	-0.031	-0.052^{*}	-0.045^{*}
	(0.008)	(0.021)	(0.021)	(0.022)
Operating Vehicles, Mechanized Devices, or Equipment	-0.192^{**}	-0.128^{**}	-0.137^{**}	-0.129^{**}
	(0.008)	(0.016)	(0.014)	(0.013)

Continued on next page

Table A.1 – *Continued from previous page*

	(1)	(2)	(3)	(4)
N		464	464	464
R^2		0.768	0.699	0.726
Adjusted R^2		0.745	0.691	0.719

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table displays coefficients from regressions of the occupational female share in the 2012-2016 ACS on O*NET variables from the work activities domain, standardized to have a mean of zero and standard deviation of one. Column 1 shows coefficients from bivariate regressions of the female share on each variable separately; the variables in this table are sorted by these coefficients, in descending order. Column 2 shows coefficients from the multivariate regression of the female share on all work activities. Column 3 shows coefficients from the female share regressed on the six work activities with the most positive bivariate coefficients and the six work activities with the most negative bivariate coefficients. Column 4 shows coefficients from the female share regressed on the variables designated as highly predictive by a random forest algorithm, as described in Appendix A.1. Selected tasks are displayed in bold.

Table A.2: *Female Share Regressed on Work Styles*

	(1)	(2)	(3)	(4)
Cooperation	0.139**	0.086**	0.125**	0.095**
	(0.012)	(0.022)	(0.020)	(0.020)
Social Orientation	0.131**	0.108**	0.106**	0.093**
	(0.011)	(0.022)	(0.020)	(0.017)
Concern for Others	0.124**	0.000		
	(0.011)	(0.019)		
Integrity	0.114**	0.078**		0.082**
	(0.012)	(0.017)		(0.013)
Self Control	0.103**	-0.056**		
	(0.013)	(0.021)		
Independence	0.092**	0.060**		
	(0.013)	(0.013)		
Adaptability/Flexibility	0.091**	0.026		
	(0.012)	(0.020)		
Dependability	0.090**	-0.021		
	(0.014)	(0.019)		
Stress Tolerance	0.086**	0.011		
	(0.013)	(0.020)		
Attention to Detail	0.052**	0.024		
	(0.013)	(0.016)		
Achievement/Effort	0.043**	0.042+		
	(0.012)	(0.025)		
Initiative	0.041**	0.040		
	(0.012)	(0.026)		
Persistence	0.018	-0.061*		
	(0.012)	(0.024)		
Innovation	0.004	-0.040*		
	(0.012)	(0.017)		
Leadership	0.003	-0.121**	-0.137**	-0.136**
	(0.012)	(0.016)	(0.015)	(0.013)
Analytical Thinking	-0.001	-0.038*	0.033*	
	(0.011)	(0.019)	(0.014)	
N		464	464	464
R^2		0.505	0.396	0.441
Adjusted R^2		0.487	0.391	0.436

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table displays coefficients from regressions of the occupational female share in the 2012-2016 ACS on O*NET variables from the work styles domain, standardized to have a mean of zero and standard deviation of one. Column 1 shows coefficients from bivariate regressions of the female share on each variable separately; the variables in this table are sorted by these coefficients, in descending order. Column 2 shows coefficients from the multivariate regression of the female share on all work styles. Column 3 shows coefficients from the female share regressed on the two work styles with the most positive bivariate coefficients and the two work styles with the most negative bivariate coefficients. Column 4 shows coefficients from the female share regressed on the variables designated as highly predictive by a random forest algorithm, as described in Appendix A.1. Selected tasks are displayed in bold.

Table A.3: *Work Activities with Controls*

	(1)	(2)	(3)	(4)
Establishing and Maintaining Interpersonal Relationships	0.062**	0.003	0.001	
	(0.012)	(0.016)	(0.013)	
Performing Administrative Activities	0.048**	0.020	0.010	
	(0.010)	(0.012)	(0.010)	
Assisting and Caring for Others	0.079**	0.068**	0.075**	0.075**
	(0.011)	(0.014)	(0.012)	(0.010)
Performing for or Working Directly with the Public	0.024*	0.010	-0.003	
	(0.011)	(0.014)	(0.009)	
Interacting With Computers	0.032*	-0.000	-0.002	
	(0.015)	(0.019)	(0.017)	
Documenting/Recording Information	0.063**	0.020	0.055**	0.063**
	(0.012)	(0.014)	(0.012)	(0.010)
Resolving Conflicts and Negotiating with Others	0.017	-0.005		
	(0.011)	(0.016)		
Communicating with Persons Outside Organization	-0.022*	-0.003		
	(0.011)	(0.015)		
Interpreting the Meaning of Information for Others	0.047**	0.024		
	(0.014)	(0.016)		
Organizing, Planning, and Prioritizing Work	0.022+	0.030+		
	(0.012)	(0.016)		
Getting Information	0.045**	0.007		
	(0.012)	(0.016)		
Communicating with Supervisors, Peers, or Subordinates	0.027*	0.009		
	(0.011)	(0.015)		
Processing Information	0.027*	-0.014		
	(0.014)	(0.020)		
Updating and Using Relevant Knowledge	0.027+	0.010		
	(0.014)	(0.017)		
Selling or Influencing Others	-0.032**	0.002		
	(0.010)	(0.012)		
Developing and Building Teams	0.007	0.022		
	(0.010)	(0.018)		
Coaching and Developing Others	0.006	-0.011		
	(0.010)	(0.021)		
Staffing Organizational Units	-0.015	-0.009		
	(0.010)	(0.014)		
Analyzing Data or Information	0.001	0.021		
	(0.017)	(0.021)		
Provide Consultation and Advice to Others	-0.015	-0.011		
	(0.013)	(0.014)		

Continued on next page

Table A.3 – *Continued from previous page*

	(1)	(2)	(3)	(4)
Training and Teaching Others	0.007	0.011		
	(0.010)	(0.014)		
Scheduling Work and Activities	-0.027**	0.000		
	(0.010)	(0.014)		
Developing Objectives and Strategies	-0.040**	-0.021		
	(0.012)	(0.018)		
Thinking Creatively	-0.053**	-0.011		-0.003
	(0.011)	(0.015)		(0.012)
Evaluating Information to Determine Compliance	0.037**	0.008		
	(0.010)	(0.013)		
Coordinating the Work and Activities of Others	-0.010	-0.010		
	(0.010)	(0.016)		
Judging the Qualities of Things, Services, or People	-0.010	0.009		
	(0.010)	(0.011)		
Monitoring and Controlling Resources	-0.038**	-0.012		
	(0.010)	(0.014)		
Guiding, Directing, and Motivating Subordinates	-0.020*	-0.019		
	(0.010)	(0.018)		
Making Decisions and Solving Problems	-0.023+	-0.025		-0.014
	(0.012)	(0.016)		(0.012)
Identifying Objects, Actions, and Events	-0.008	-0.034*		
	(0.010)	(0.014)		
Monitor Processes, Materials, or Surroundings	0.017+	0.037**		
	(0.010)	(0.014)		
Performing General Physical Activities	-0.030+	-0.009		
	(0.016)	(0.025)		
Handling and Moving Objects	-0.026	0.022		
	(0.017)	(0.026)		
Estimating Quantifiable Characteristics	-0.067**	-0.013		-0.012
	(0.011)	(0.014)		(0.010)
Repairing and Maintaining Electronic Equipment	-0.046**	-0.012	-0.007	-0.010
	(0.011)	(0.014)	(0.014)	(0.013)
Controlling Machines and Processes	-0.046**	0.004	0.023	0.020
	(0.017)	(0.023)	(0.018)	(0.017)
Inspecting Equipment, Structures, or Material	-0.055**	-0.016	-0.014	-0.007
	(0.015)	(0.019)	(0.015)	(0.016)
Drafting, Laying Out, and Specifying Equipment	-0.084**	-0.029**	-0.037**	-0.031**
	(0.009)	(0.011)	(0.008)	(0.009)
Repairing and Maintaining Mechanical Equipment	-0.099**	-0.022	-0.044*	-0.040*
	(0.015)	(0.020)	(0.019)	(0.019)
Operating Vehicles, Mechanized Devices, or Equipment	-0.132**	-0.120**	-0.129**	-0.127**
	(0.012)	(0.017)	(0.014)	(0.013)

Continued on next page

Table A.3 – *Continued from previous page*

	(1)	(2)	(3)	(4)
N		464	464	464
R^2		0.795	0.769	0.772
Adjusted R^2		0.772	0.759	0.763

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table displays coefficients from regressions of the occupational female share in the 2012-2016 ACS on O*NET variables from the work activities domain, standardized to have a mean of zero and standard deviation of one. Column 1 shows coefficients from regressions of the female share on each variable separately. Column 2 shows coefficients from the regression of the female share on all work activities. Column 3 shows coefficients from the female share regressed on the six work activities with the most positive bivariate coefficients and the six work activities with the most negative bivariate coefficients in Table A.1. Column 4 shows coefficients from the female share regressed on the variables designated as highly predictive by a random forest algorithm, as described in Appendix A.1. Selected tasks are displayed in bold. All regressions control for broad occupation cluster, mean log hourly wages and the share of workers with a college degree or more in each occupation in the ACS.

Table A.4: *Work Styles with Controls*

	(1)	(2)	(3)	(4)
Cooperation	0.079**	0.032+	0.066**	0.059**
	(0.011)	(0.017)	(0.015)	(0.015)
Social Orientation	0.074**	0.069**	0.076**	0.069**
	(0.011)	(0.018)	(0.016)	(0.015)
Concern for Others	0.078**	0.026		
	(0.010)	(0.018)		
Integrity	0.068**	0.005		0.037**
	(0.012)	(0.015)		(0.012)
Self Control	0.061**	-0.010		
	(0.010)	(0.018)		
Independence	0.063**	0.054**		
	(0.010)	(0.011)		
Adaptability/Flexibility	0.054**	0.011		
	(0.012)	(0.016)		
Dependability	0.054**	-0.016		
	(0.011)	(0.014)		
Stress Tolerance	0.048**	0.003		
	(0.011)	(0.017)		
Attention to Detail	0.056**	0.046**		
	(0.010)	(0.011)		
Achievement/Effort	0.029*	0.031		
	(0.013)	(0.020)		
Initiative	0.021+	0.019		
	(0.013)	(0.019)		
Persistence	0.002	-0.048*		
	(0.013)	(0.019)		
Innovation	-0.004	-0.041**		
	(0.011)	(0.013)		
Leadership	-0.001	-0.071**	-0.088**	-0.086**
	(0.011)	(0.014)	(0.013)	(0.013)
Analytical Thinking	0.004	0.002	0.021	
	(0.014)	(0.016)	(0.014)	
N		464	464	464
R^2		0.689	0.632	0.638
Adjusted R^2		0.674	0.624	0.630

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table displays coefficients from regressions of the occupational female share in the 2012-2016 ACS on O*NET variables from the work styles domain, standardized to have a mean of zero and standard deviation of one. Column 1 shows coefficients from regressions of the female share on each variable separately. Column 2 shows coefficients from the regression of the female share on all work styles. Column 3 shows coefficients from the female share regressed on the two work styles with the most positive bivariate coefficients and the two work styles with the most negative bivariate coefficients in Table A.2. Column 4 shows coefficients from the female share regressed on the variables designated as highly predictive by a random forest algorithm, as described in Appendix A.1. Selected tasks are displayed in bold. All regressions control for broad occupation cluster, mean log hourly wages and the share of workers with a college degree or more in each occupation in the ACS.

Table A.5: *Occupations with Highest and Lowest Levels -*

Helping and Caring for Others

Rank	Occupation	Female Share
1	Nurse anesthetists	0.592
2	Nurse practitioners and nurse midwives	0.911
3	Podiatrists	0.248
4	Licensed practical and licensed vocational nurses	0.887
5	Registered nurses	0.898
6	Dental assistants	0.942
7	Nursing, psychiatric, and home health hides	0.869
8	Recreational therapists	0.765
9	Medical assistants	0.914
10	Personal care aides	0.836
455	Fence erectors	0.017
456	Surveying and mapping technicians	0.093
457	Record keeping weighers, measurers, checkers, and samplers	0.462
458	Miscellaneous legal support workers	0.730
459	Economists	0.334
460	Computer programmers	0.218
461	Proofreaders and copy markers	0.712
462	Automotive glass installers and repairers	0.022
463	Actuaries	0.348
464	Financial analysts	0.388

Notes: This table displays the ten occupations ranked the highest and the ten occupations ranked the lowest in terms of the level of the task *helping and caring for others*, along with the share of workers who are female in each occupation based on the 2012-2016 ACS. Tasks are measured using the O*NET variables listed in Table 1.1 and are standardized to have a mean of zero and standard deviation of one.

Table A.6: *Occupations with Highest and Lowest Levels -*

Documenting and Recording Information

Rank	Occupation	Female Share
1	Licensed practical and licensed vocational nurses	0.887
2	Nurse anesthetists	0.592
3	Nurse practitioners and nurse midwives	0.911
4	Registered nurses	0.898
5	Medical assistants	0.914
6	Physical therapists	0.696
7	Personal care aides	0.836
8	Optometrists	0.420
9	Medical transcriptionists	0.940
10	Respiratory therapists	0.645
455	Structural iron and steel workers	0.027
456	Shoe and leather workers	0.313
457	Sewing machine operators	0.743
458	Drywall installers, ceiling tile installers, and tapers	0.025
459	Food preparation workers	0.586
460	Transportation attendants, except flight attendants	0.586
461	Pressers, textile, garment, and related materials	0.667
462	Graders and sorters of agricultural products	0.690
463	Laundry and dry-cleaning workers	0.595
464	Tire builders	0.089

Notes: This table displays the ten occupations ranked the highest and the ten occupations ranked the lowest in terms of the level of the task *documenting and recording information*, along with the share of workers who are female in each occupation based on the 2012-2016 ACS. Tasks are measured using the O*NET variables listed in Table 1.1 and are standardized to have a mean of zero and standard deviation of one.

Table A.7: *Occupations with Highest and Lowest Levels -*

Working and Communicating with Others

Rank	Occupation	Female Share
1	Flight attendants	0.776
2	Dental assistants	0.942
3	Elementary and middle school teachers	0.790
4	Actors	0.432
5	Reservation and transportation ticket agents and travel clerks	0.595
6	Meeting, convention, and event planners	0.767
7	Respiratory therapists	0.645
8	Social workers	0.810
9	Special education teachers	0.854
10	Counselors	0.715
455	Couriers and messengers	0.164
456	Printing press operators	0.204
457	Metal furnace operators, tenders, pourers, and casters	0.076
458	Library technicians	0.755
459	Computer hardware engineers	0.160
460	Machinists	0.044
461	Motion picture projectionists	0.186
462	Graders and sorters of agricultural products	0.690
463	Economists	0.334
464	Computer, automated teller, and office machine repairers	0.111

Notes: This table displays the ten occupations ranked the highest and the ten occupations ranked the lowest in terms of the level of the task *working and communicating with others*, along with the share of workers who are female in each occupation based on the 2012-2016 ACS. Tasks are measured using the O*NET variables listed in Table 1.1 and are standardized to have a mean of zero and standard deviation of one.

Table A.8: *Occupations with Highest and Lowest Levels -*

Operating and Repairing Equipment

Rank	Occupation	Female Share
1	Elevator installers and repairers	0.011
2	Heating, air conditioning, and refrigeration mechanics and installers	0.012
3	Aircraft mechanics and service technicians	0.053
4	Computer, automated teller, and office machine repairers	0.111
5	Coin, vending, and amusement machine servicers and repairers	0.172
6	Electrical power-line installers and repairers	0.011
7	Automotive service technicians and mechanics	0.015
8	Heavy vehicle / mobile equipment service technicians and mechanics	0.012
9	Electrical and electronics repairers	0.070
10	Machinery maintenance workers	0.037
455	Insurance underwriters	0.654
456	Financial analysts	0.388
457	Judicial law clerks	0.543
458	Operations research analysts	0.485
459	Management analysts	0.420
460	Actuaries	0.348
461	Economists	0.334
462	Compensation and benefits managers	0.770
463	Brokerage clerks	0.703
464	Compensation, benefits, and job analysis specialists	0.767

Notes: This table displays the ten occupations ranked the highest and the ten occupations ranked the lowest in terms of the level of the task *operating and repairing equipment*, along with the share of workers who are female in each occupation based on the 2012-2016 ACS. Tasks are measured using the O*NET variables listed in Table 1.1 and are standardized to have a mean of zero and standard deviation of one.

Table A.9: *Occupations with Highest and Lowest Levels -*

Making Decisions and Solving Problems

Rank	Occupation	Female Share
1	Nurse anesthetists	0.592
2	Social and community service managers	0.681
3	Actuaries	0.348
4	Lawyers, and judges, magistrates, and other judicial workers	0.363
5	Air traffic controllers and airfield operations specialists	0.216
6	Management analysts	0.420
7	Physicians and surgeons	0.353
8	Operations research analysts	0.485
9	Biomedical and agricultural engineers	0.144
10	Chief executives and legislators	0.245
455	Pressers, textile, garment, and related materials	0.667
456	Roasting, baking, and drying machine operators and tenders	0.315
457	Textile knitting and weaving machine setters, operators, and tenders	0.598
458	Postal service mail carriers	0.392
459	Miscellaneous personal appearance workers	0.854
460	Shoe and leather workers	0.313
461	Miscellaneous agricultural workers including animal breeders	0.208
462	Cleaners of vehicles and equipment	0.151
463	Interviewers, except eligibility and loan	0.796
464	Graders and sorters of agricultural products	0.690

Notes: This table displays the ten occupations ranked the highest and the ten occupations ranked the lowest in terms of the level of the task *making decisions and solving problems*, along with the share of workers who are female in each occupation based on the 2012-2016 ACS. Tasks are measured using the O*NET variables listed in Table 1.1 and are standardized to have a mean of zero and standard deviation of one.

Table A.10: *Summary Statistics - Employed and Non-Employed*

	Experiment			ACS		
	All	Women	Men	All	Women	Men
Female	0.534	1.000	0.000	0.514	1.000	0.000
Age	34.4	35.0	33.7	47.0	47.9	46.0
White	0.720	0.770	0.663	0.650	0.648	0.652
Black	0.079	0.070	0.090	0.119	0.124	0.114
Hispanic	0.109	0.077	0.146	0.152	0.147	0.157
Other Race	0.091	0.083	0.101	0.079	0.081	0.076
HS or less	0.091	0.092	0.089	0.411	0.393	0.430
Some college	0.238	0.232	0.245	0.235	0.239	0.230
Associate's degree	0.124	0.143	0.103	0.078	0.086	0.069
Bachelor's degree	0.400	0.372	0.433	0.176	0.180	0.171
Graduate degree	0.146	0.161	0.130	0.101	0.101	0.100
Employed	0.902	0.889	0.917	0.602	0.553	0.654
N	1,931	1,031	900	12,330,760	6,393,549	5,937,211

Notes: This table shows summary statistics in the experiment sample compared with the 2012-2016 ACS, including all experiment participants and all individuals in the ACS aged 18 and older.

Table A.11: *WTP for Tasks - College vs. Non-College Workers*

(a) College degree or more

	Help	Document	Communic.	Operate	Decisions
All	-0.023**	-0.021**	-0.014*	-0.126**	0.019**
	(0.008)	(0.005)	(0.006)	(0.011)	(0.006)
Women	-0.011	-0.009	-0.014+	-0.160**	0.020**
	(0.011)	(0.006)	(0.008)	(0.016)	(0.008)
Men	-0.037**	-0.036**	-0.016+	-0.088**	0.017+
	(0.010)	(0.009)	(0.008)	(0.014)	(0.009)
Diff (W-M)	0.026+	0.028**	0.002	-0.072**	0.004
	(0.015)	(0.011)	(0.011)	(0.022)	(0.012)
N	1,050	1,050	1,050	1,050	1,050

(b) Less than college degree

	Help	Document	Communic.	Operate	Decisions
All	-0.029**	-0.021**	-0.027**	-0.107**	-0.026**
	(0.009)	(0.006)	(0.007)	(0.012)	(0.006)
Women	-0.009	-0.010	-0.019*	-0.144**	-0.035**
	(0.011)	(0.008)	(0.009)	(0.015)	(0.009)
Men	-0.055**	-0.033**	-0.038**	-0.051**	-0.015+
	(0.014)	(0.009)	(0.011)	(0.015)	(0.009)
Diff (W-M)	0.045**	0.022+	0.019	-0.093**	-0.020
	(0.017)	(0.012)	(0.014)	(0.022)	(0.013)
N	872	872	872	872	872

Continued on next page

Table A.11: *(Continued) WTP - College vs. Non-College*

(c) Difference (College − Non-College)

	Help	Document	Communic.	Operate	Decisions
All	0.006	-0.001	0.013	-0.019	0.045**
	(0.011)	(0.008)	(0.009)	(0.016)	(0.009)
Women	-0.002	0.002	0.006	-0.016	0.056**
	(0.015)	(0.010)	(0.012)	(0.022)	(0.012)
Men	0.018	-0.004	0.022+	-0.037+	0.032*
	(0.017)	(0.012)	(0.013)	(0.021)	(0.013)
Diff (W-M)	-0.020	0.006	-0.017	0.020	0.024
	(0.023)	(0.016)	(0.018)	(0.031)	(0.017)
N	1,922	1,922	1,922	1,922	1,922

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates for experiment participants with a four-year college degree or more in Panel A, WTP estimates for participants with less than a college degree in Panel B, and the difference between the estimates in Panel C.

Table A.12: *WTP for Tasks - No Inattention*

	Help	Document	Communic.	Operate	Decisions
All	-0.026**	-0.017**	-0.025**	-0.119**	0.000
	(0.006)	(0.004)	(0.004)	(0.008)	(0.004)
Women	-0.010	-0.005	-0.021**	-0.151**	-0.003
	(0.008)	(0.005)	(0.006)	(0.011)	(0.006)
Men	-0.043**	-0.033**	-0.030**	-0.079**	0.005
	(0.008)	(0.006)	(0.006)	(0.010)	(0.007)
Diff (W-M)	0.033**	0.028**	0.009	-0.072**	-0.008
	(0.011)	(0.008)	(0.009)	(0.015)	(0.009)
N	1,588	1,588	1,588	1,588	1,588

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates excluding participants who are inattentive. Inattention is measured by asking participants at the end of the survey to indicate the decisions they made in the hypothetical choice experiment for a randomly selected two of the five gender-typical tasks. A participant is considered inattentive if they answer either question incorrectly.

Table A.13: *WTP for Tasks - Employed*

	Help	Document	Communic.	Operate	Decisions
All	-0.025**	-0.021**	-0.021**	-0.119**	-0.000
	(0.006)	(0.004)	(0.005)	(0.009)	(0.004)
Women	-0.008	-0.007	-0.018**	-0.156**	-0.006
	(0.008)	(0.005)	(0.006)	(0.012)	(0.006)
Men	-0.043**	-0.036**	-0.023**	-0.074**	0.006
	(0.008)	(0.006)	(0.007)	(0.011)	(0.007)
Diff (W-M)	0.035**	0.029**	0.005	-0.082**	-0.012
	(0.012)	(0.008)	(0.009)	(0.016)	(0.009)
N	1,742	1,742	1,742	1,742	1,742

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates excluding participants who are not currently employed.

Table A.14: *WTP for Tasks - No Existing Skills Motivation*

	Help	Document	Communic.	Operate	Decisions
All	-0.024*	-0.047**	-0.033**	-0.060**	-0.017*
	(0.010)	(0.007)	(0.008)	(0.010)	(0.007)
Women	-0.016	-0.032**	-0.028*	-0.080**	-0.017+
	(0.015)	(0.009)	(0.011)	(0.013)	(0.010)
Men	-0.032*	-0.060**	-0.039**	-0.036*	-0.017
	(0.015)	(0.011)	(0.012)	(0.015)	(0.011)
Diff (W-M)	0.015	0.028*	0.010	-0.043*	-0.000
	(0.021)	(0.014)	(0.016)	(0.020)	(0.015)
N	461	544	523	496	532

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates excluding data from hypothetical scenarios for which participants cite a better fit for existing skills as a motivation for their choice. This table also restricts the data to the randomly selected two tasks for which each participant is asked to give a reason for their choice.

Table A.15: *WTP for Tasks - No Develop New Skills Motivation*

	Help	Document	Communic.	Operate	Decisions
All	-0.045**	-0.034**	-0.049**	-0.130**	-0.024**
	(0.011)	(0.006)	(0.008)	(0.015)	(0.007)
Women	-0.028*	-0.018*	-0.039**	-0.167**	-0.024*
	(0.014)	(0.008)	(0.010)	(0.021)	(0.010)
Men	-0.070**	-0.049**	-0.062**	-0.078**	-0.025*
	(0.020)	(0.009)	(0.013)	(0.019)	(0.011)
Diff (W-M)	0.042^{+}	0.031*	0.022	-0.089**	0.001
	(0.024)	(0.012)	(0.016)	(0.028)	(0.015)
N	617	672	632	659	598

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows WTP estimates excluding data from hypothetical scenarios for which participants cite developing new skills as a motivation for their choice. This table also restricts the data to the randomly selected two tasks for which each participant is asked to give a reason for their choice.

Table A.16: *WTP for Tasks - Difference in Differences*

Help − Operate	0.113**
	(0.019)
Document − Operate	0.106**
	(0.017)
Communicate − Operate	0.088**
	(0.018)
Help − Decisions	0.041**
	(0.014)
Document − Decisions	0.034**
	(0.012)
Communicate − Decisions	0.017
	(0.012)
N	1,931

Notes: Robust standard errors in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. This table shows the difference in differences in the WTP estimates across gender and tasks. The point estimates correspond to the difference across tasks in the estimates of β_k reported in the last row of Table 1.4.

Table A.17: *Sorting and Segregation - Robustness Checks*

(a) CPS MORG

	Q	I	b	$\hat{Q}$	$\hat{Q}/Q$	$\hat{I}$	$\hat{I}/I$
Help	0.276		-0.222	0.017	0.063		
Document	0.147		0.131	0.054	0.369		
Communic.	0.264		0.045	0.027	0.104		
Operate	-0.312		-0.011	-0.149	0.478		
Decisions	-0.006		0.238	-0.006	0.956		
Index		0.389				0.161	0.413
N			648,149				

(b) Additional Amenities in ACS

	Q	I	b	$\hat{Q}$	$\hat{Q}/Q$	$\hat{I}$	$\hat{I}/I$
Help	0.270		-0.087	0.058	0.216		
Document	0.141		0.065	0.074	0.524		
Communic.	0.259		0.016	0.028	0.107		
Operate	-0.305		-0.039	-0.122	0.402		
Decisions	-0.011		0.079	-0.021	1.887		
Index		0.389				0.154	0.395
N			6,419,869				

(c) PSID

	Q	I	b	$\hat{Q}$	$\hat{Q}/Q$	$\hat{I}$	$\hat{I}/I$
Help	0.268		-0.153	0.034	0.128		
Document	0.197		0.127	0.056	0.282		
Communic.	0.286		0.014	0.028	0.097		
Operate	-0.311		-0.068	-0.097	0.312		
Decisions	-0.072		0.244	-0.005	0.075		
Index		0.451				0.127	0.282
N			17,576				

Continued on next page

Table A.17: *(Continued) Sorting and Segregation - Robustness Checks*

(d) NLSY79

	Q	I	b	$\hat{Q}$	$\hat{Q}/Q$	$\hat{I}$	$\hat{I}/I$
Help	0.279		-0.127	0.043	0.156		
Document	0.122		0.161	0.044	0.358		
Communic.	0.257		0.022	0.028	0.108		
Operate	-0.353		0.035	-0.190	0.538		
Decisions	-0.095		0.106	-0.018	0.191		
Index		0.400				0.192	0.479
N			170,211				

(e) NLSY97

	Q	I	b	$\hat{Q}$	$\hat{Q}/Q$	$\hat{I}$	$\hat{I}/I$
Help	0.282		-0.157	0.033	0.118		
Document	0.100		0.163	0.043	0.432		
Communic.	0.304		-0.036	0.025	0.082		
Operate	-0.312		0.025	-0.182	0.582		
Decisions	-0.045		0.177	-0.010	0.232		
Index		0.380				0.185	0.488
N			63,230				

Notes: This table shows observed and predicted gender differences in sorting on the gender-typical tasks and task-based segregation in the CPS MORG (Panel A), the ACS with additional job amenities included in the wage regression (Panel B), the PSID (Panel C), the NLSY79 (Panel D), and the NLSY97 (Panel E). See Appendix A.1 for details of sample and variable construction in the additional datasets. Q ($\hat{Q}$) is the observed (predicted) gender difference in the share of workers sorting into the high-task job. b is the coefficient on the task from the wage regression. I ($\hat{I}$) is the value of the segregation index based on observed (predicted) sorting.

Appendix B

Appendix to Chapter 2

B.1 RD Validity Checks

This section probes the assumptions underlying our RD design, using the tests suggested by Imbens and Lemieux (2008). Assumption 4 implies that firm beliefs about worker value and the components of the score including the wage bid do not change discontinuously at the cutoff. This assumption might not hold if applicants are able to precisely manipulate the score, or if threshold crossing is for some other reason associated with factors that affect worker productivity or the wage, implying a violation of the independence of the instrument.

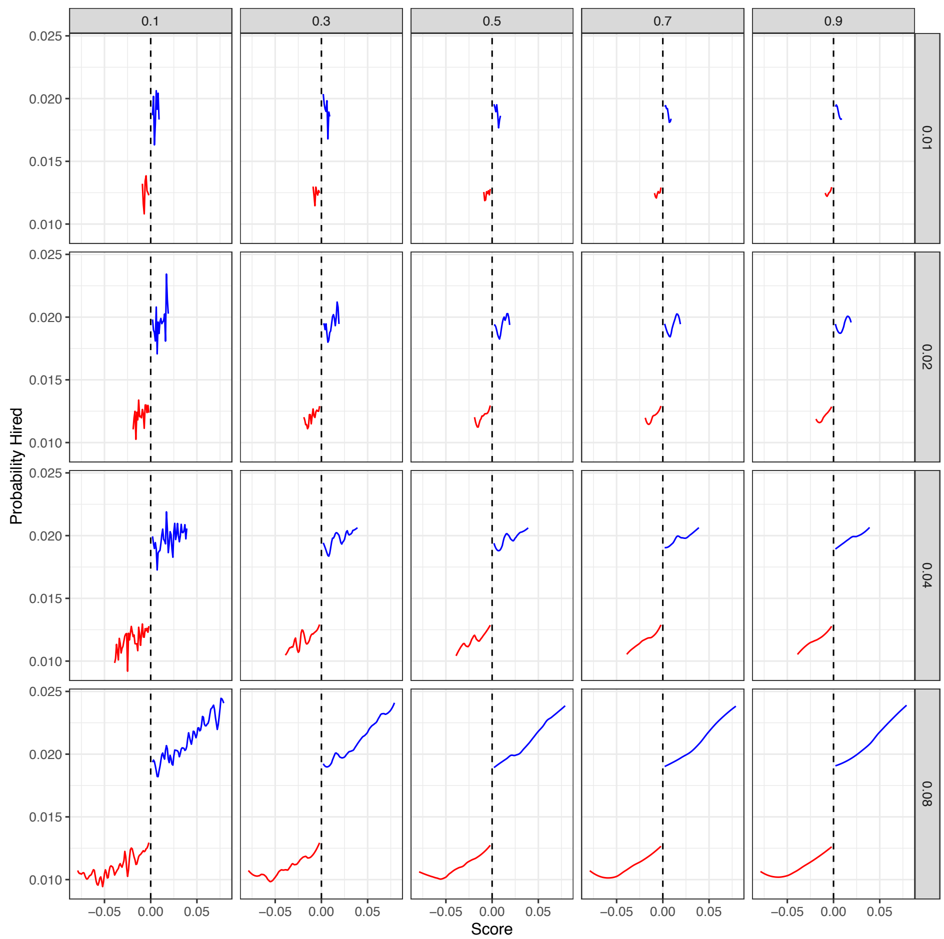
A common method of assessing whether applicants may be manipulating the score in order to cross the threshold is by testing for a discontinuity in the density of the score at the recommendation threshold. Figure B.2 plots the results of the density discontinuity test proposed by McCrary (2008). The test reports a significant break in the density estimate at the cutoff. However, as McCrary notes, the density test may fail even in the absence of manipulation. In our context, workers cannot view their recommendation status and neither employers nor workers can view the score or are aware of its existence. Therefore, it is unlikely that job applicants can sort over the threshold by adjusting their score or conditioning their wage bid on the score. Moreover, Figure B.2 shows that the test finds significant discontinuities in the score density at a range of points within the bandwidth other than the cutoff. This suggests that the discontinuity at the threshold may be a product of the algorithm used by the platform to construct the score, rather than the result of deliberate action by job applicants.

Another common test of the validity of a regression discontinuity design is to assess whether worker attributes that should not be affected by recommendation status change discontinuously at the threshold. As mentioned in Section 2.6.1, Figure 2.1 demonstrates visually that the hourly wage bid, hourly profile wage, and the mean worker rank in the applicant pool do not appear to jump at the cutoff. Table B.3 provides further evidence that the recommendation discontinuity is not significantly related to applicant wage rates. Column (1) reports results from a model in which the log hourly wage bid is regressed on the recommendation indicator. Column (2) adds a linear control for the score. Column (3) includes a second-degree polynomial in the score. Columns (4) through (6) repeat the specifications in the first three columns, with the log hourly wage listed on the applicant's profile as the outcome. As in Table 2.1, all specifications restrict the applications data to a bandwidth of 0.1 around the score cutoff, include opening-by-month fixed effects, and cluster standard errors by worker.

Results in Table B.3 indicate that there is no significant relationship between wages rates and recommendation status after conditioning on the score. When controls for the score are included in Columns (2)-(3) and Columns (5)-(6), the coefficient on the recommendation indicator is insignificant and close to zero. However, the score has a strong positive relationship with wages, suggesting that employers believe higher-wage workers to have other correlates of productivity, such as positive employer feedback or more experience, that make them more valuable despite their higher cost.

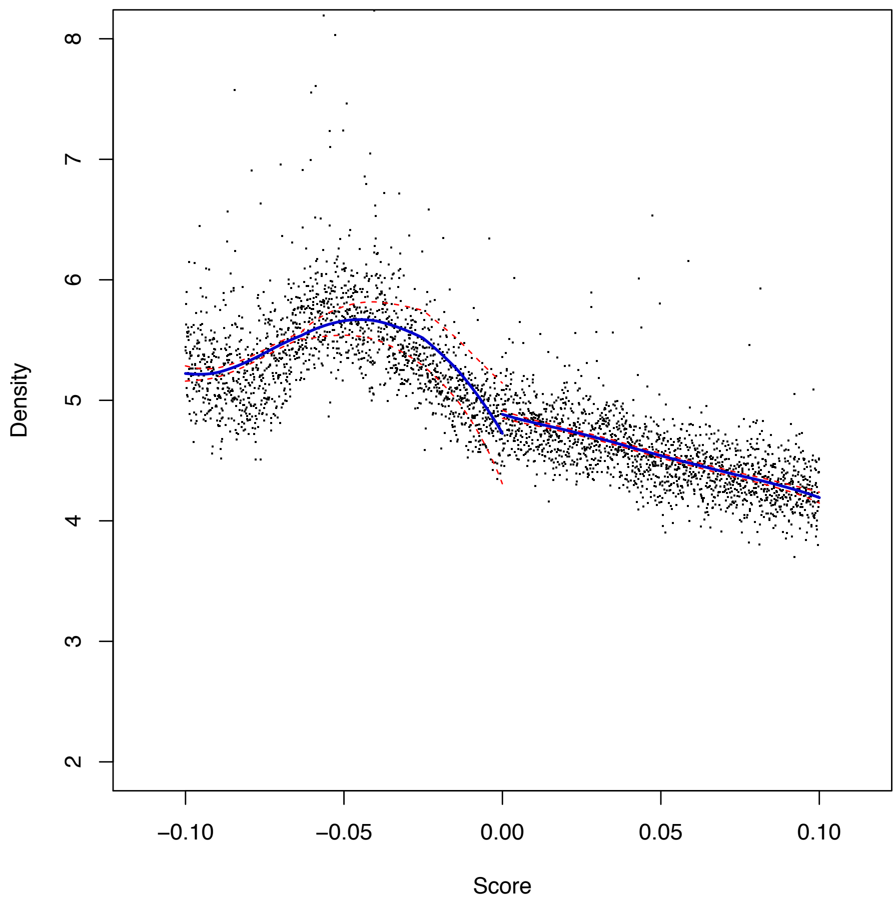
A final RD check that is typically performed is to test whether hiring, the first-stage outcome, changes discontinuously at score cutoffs other than the true recommendation thresholds. Table B.4 shows estimates of the effect of fake recommendation indicators on the probability of being hired, using specifications equivalent to Column (3) of Table 2.1.

Specifically, we show results for cutoffs of $s^* \in \{0.3, 0.4, 0.6, 07\}$. All specifications restrict to a bandwidth of 0.1 around the fake recommendation threshold. The coefficients on the fake recommendation indicators are statistically significant, but they are much smaller in magnitude than the estimates at the true cutoff. Furthermore, at three of the four fake cutoffs hiring actually decreases rather than increases.



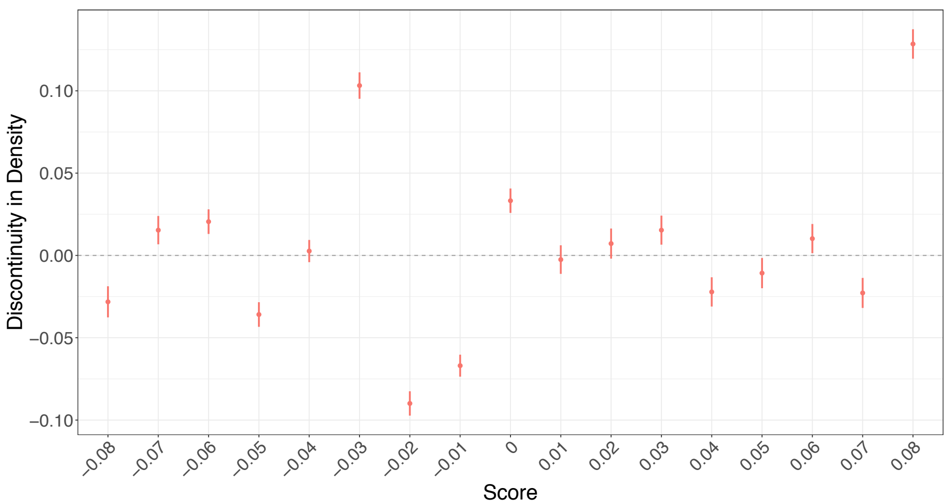
Note: This figure uses the applications data to fit local polynomial regressions of degree two that predict the probability of hiring as a function of the score, separately on each side of the threshold and using a grid of parameters. The columns indicate the smoothing parameter, α, which controls the proportion of observations used to fit the regression at each point. The rows indicate the bandwidth, δ, such that the subplots include observations with score values in the interval $(s^* - \delta, s^* + \delta)$.

Figure B.1: *Predicted Hiring Around the Threshold*



Notes: This figure plots the results of a test for a discontinuity in the density of the score at the recommendation threshold, using the methodology proposed by McCrary (2008). The data are restricted to a bandwidth of 0.1 around the recommendation threshold. The black dots show the height of the first-step histogram. The blue lines show the smoothed density estimates using local linear regression, separately above versus below the cutoff, while the dashed red lines show 95 percent confidence intervals.

Figure B.2: *Discontinuity Test for Score Density*



Notes: This figure plots results of the McCrary (2008) test for a discontinuity in the density of the score for a range of score values surrounding the true recommendation threshold. The data are restricted to a bandwidth of 0.1 around the true cutoff. The point estimate is the log difference in the estimate of the density above versus below the score value, while the lines represent 95 percent confidence intervals.

Figure B.3: *Density Discontinuity Estimates at Fake Thresholds*

Table B.1: *Summary Statistics - Applications Data*

Statistic	N	Mean	St. Dev.
Hired	3,648,710	0.014	0.117
Recommended	3,648,710	0.455	0.498
Score	3,648,710	0.495	0.057
Hourly Wage Bid	3,648,710	13.699	12.618
Hourly Profile Wage	3,648,710	13.517	12.015

Table B.2: *Summary Statistics - Job Openings*

Statistic	N	Mean	St. Dev.
Δ_j	117,624	0.077	0.517
Hired Wage	117,624	12.892	11.587
Hours Worked	117,624	77.049	239.048
Total Wage Bill	117,624	797.145	3,220.327

Table B.3: *Effect of Recommendation on Wage Bid and Profile Wage*

	Outcome:					
	Log Hourly Wage Bid			Log Hourly Profile Wage		
	(1)	(2)	(3)	(4)	(5)	(6)
Recommended	0.076**	0.001	−0.0001	0.071**	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Score		0.762**			0.699**	
		(0.022)			(0.024)	
Score Polynom.	N	N	2nd deg	N	N	2nd deg
Observations	3,648,710	3,648,710	3,648,710	3,648,710	3,648,710	3,648,710
R^2	0.656	0.657	0.657	0.611	0.611	0.611
Adjusted R^2	0.604	0.605	0.605	0.552	0.552	0.552

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports OLS estimates, using the applications data and restricting to a bandwidth of 0.1 around the recommendation threshold. All columns include opening-by-month fixed effects and cluster standard errors by worker.

Table B.4: *Effect of Fake Recommendation Thresholds on Hiring*

	Outcome:			
	Hired			
	$s^* = 0.3$	$s^* = 0.4$	$s^* = 0.6$	$s^* = 0.7$
	(1)	(2)	(3)	(4)
Recommended	-0.0005^{*}	-0.0004^{*}	0.001^{**}	-0.003^{**}
	(0.0002)	(0.0002)	(0.0004)	(0.001)
Score	0.033^{**}	0.035^{**}	0.079^{**}	0.127^{**}
	(0.002)	(0.002)	(0.004)	(0.006)
Recommended*Score	0.0004	-0.009^{*}	0.013^{+}	-0.091^{**}
	(0.003)	(0.004)	(0.007)	(0.011)
Observations	4,139,893	4,655,449	3,173,804	1,972,695
R^2	0.295	0.294	0.300	0.411
Adjusted R^2	0.163	0.164	0.115	0.171

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports linear probability estimates equivalent to Column (3) of Table 2.1, but using fake recommendation thresholds based on values of the score other than 0.5, as indicated. All columns restrict to a bandwidth of 0.1 around the fake recommendation threshold, include opening-by-month fixed effects and cluster standard errors by worker.

Table B.5: *Effect of Delta Instrument on Hours Worked and Wage Bill*

	Outcome:			
	Log of Hours Worked		Log of Total Wage Bill	
	(1)	(2)	(3)	(4)
Δ_j	-0.028^{**}	-0.028	0.034^{**}	0.114^{**}
	(0.010)	(0.018)	(0.012)	(0.023)
Restrict to Openings w/ Hired Worker in BW	N	Y	N	Y
Observations	117,624	36,036	117,624	36,036
R^2	0.0001	0.0001	0.0001	0.001

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports OLS estimates of the reduced form effect of Δ_j on hours worked and the wage bill. The data are aggregated to the job opening level. Columns (1) and (3) include all job openings. Columns (2) and (4) restrict to openings where the hired worker is within the bandwidth.

Table B.6: *Effect of Log Hourly Wage Bid on Hiring*

	Outcome:					
	Hired					
	Below Threshold			Above Threshold		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Wage Bid	-0.001^{**}	-0.001^{**}	-0.001^{**}	-0.005^{**}	-0.005^{**}	-0.005^{**}
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Score		0.026^{**}			0.086^{**}	
		(0.003)			(0.004)	
Score Polynom.	N	N	2nd deg	N	N	2nd deg
Observations	1,989,238	1,989,238	1,989,238	1,659,341	1,659,341	1,659,341
R^2	0.365	0.365	0.365	0.371	0.371	0.371
Adjusted R^2	0.175	0.175	0.175	0.136	0.136	0.136

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table reports linear probability estimates, using the applications data and restricting to a bandwidth of 0.1 around the recommendation threshold. Columns (1)-(3) includes only applicants below the cutoff, while Columns (4)-(6) includes only applicants above the cutoff. All columns include opening-by-month fixed effects and cluster standard errors by worker.

Appendix C

Appendix to Chapter 3

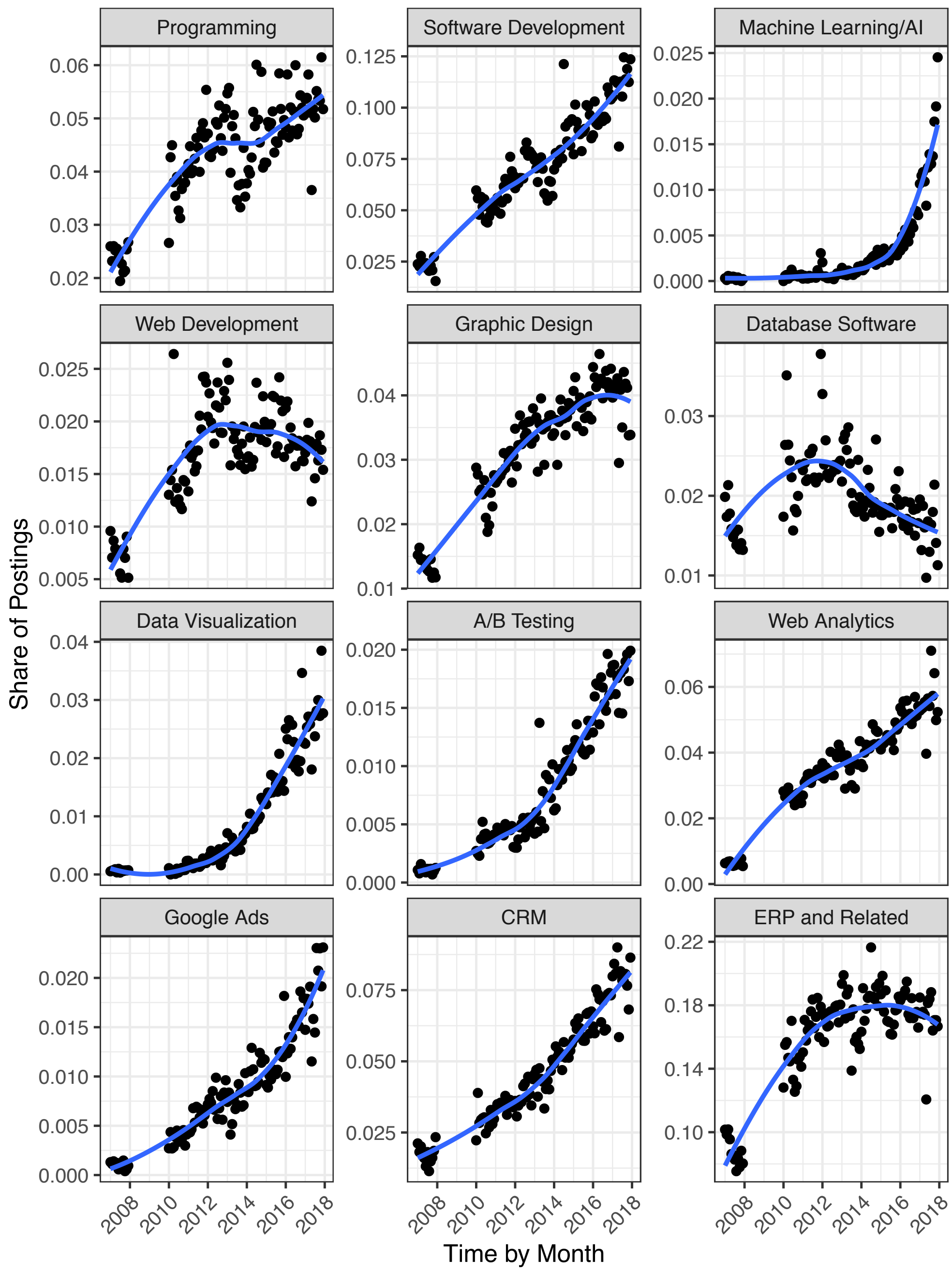


Figure C.1: *Marketing Managers - Skill Composites Over Time*

Continued on next page

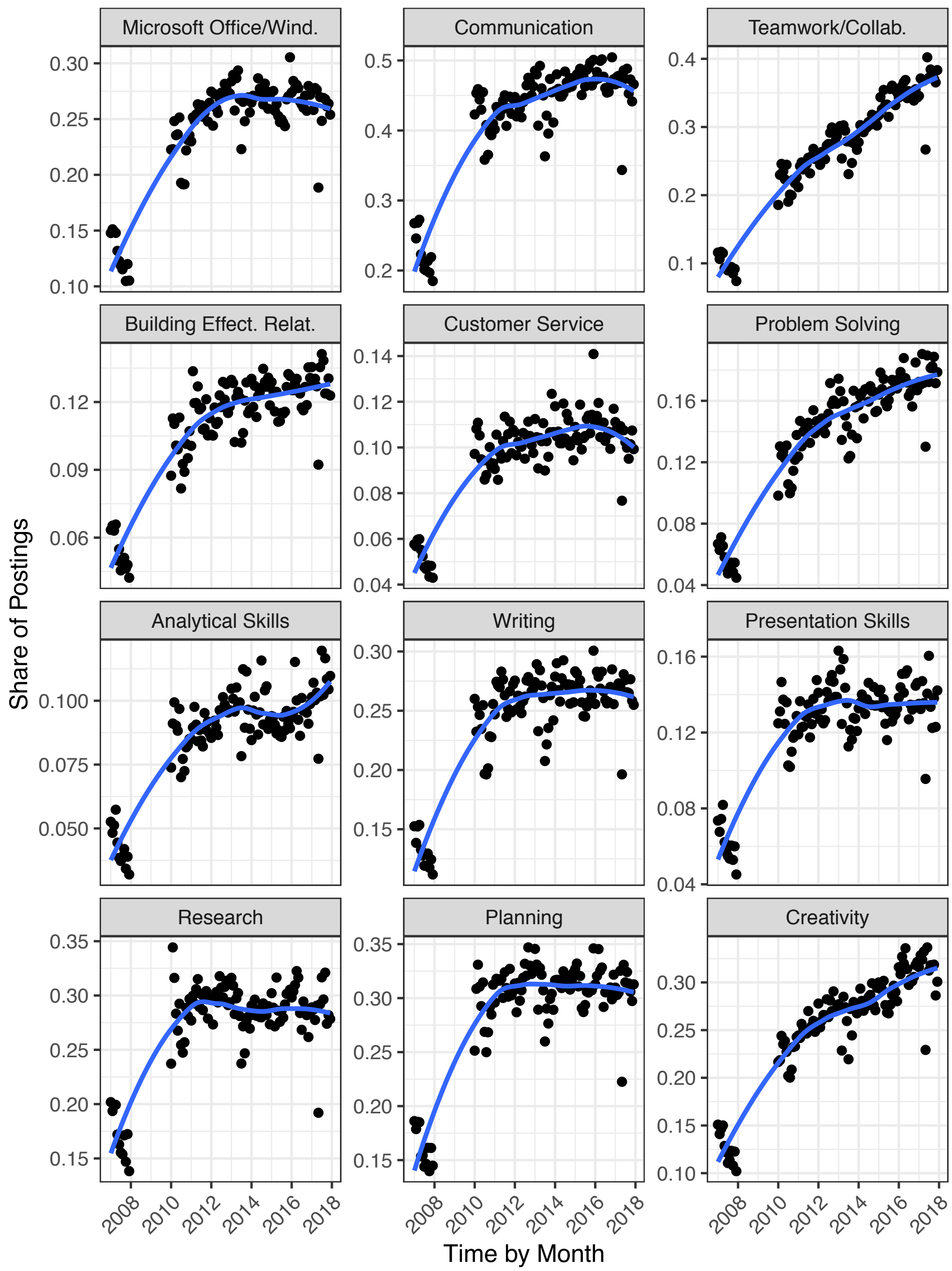
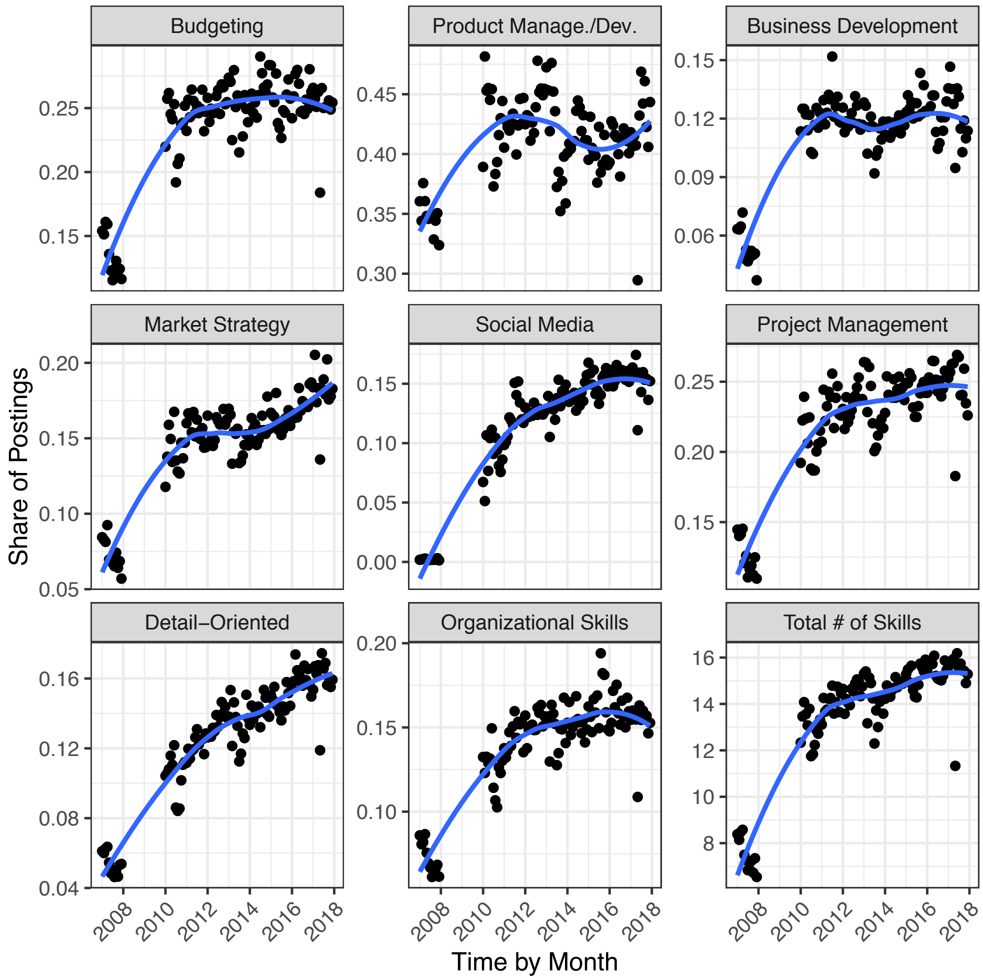


Figure C.1: *(Continued) Marketing Managers - Skill Composites Over Time*

Continued on next page



Notes: This figure shows the share of job postings for marketing managers that mention each of the skill composites and the total number of skills by month in 2007 and 2010-2017 (dots), along with fitted local polynomial regressions (lines).

Figure C.1: *(Continued) Marketing Managers - Skill Composites Over Time*

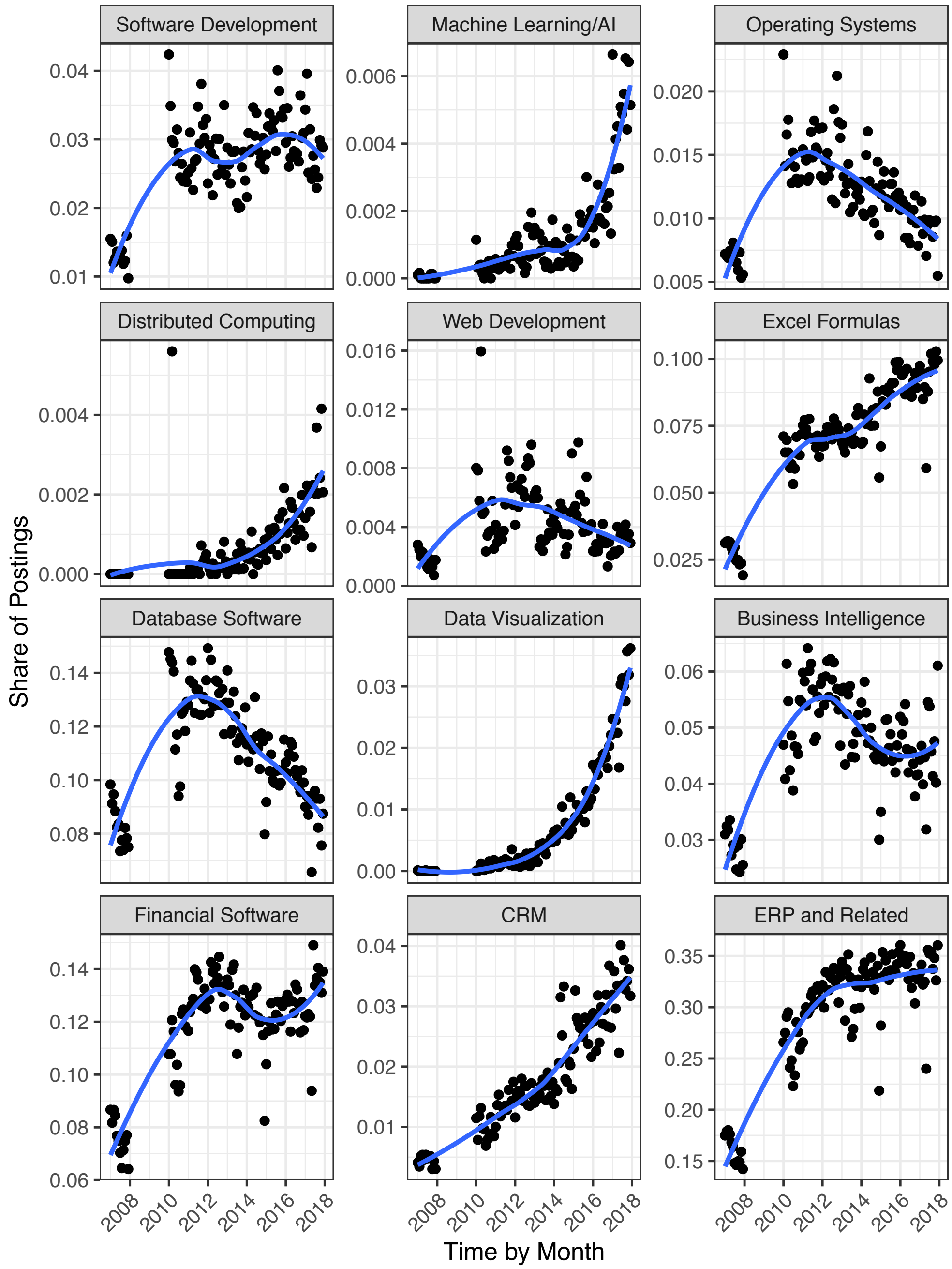


Figure C.2: *Financial Analysts - Skill Composites Over Time*

Continued on next page

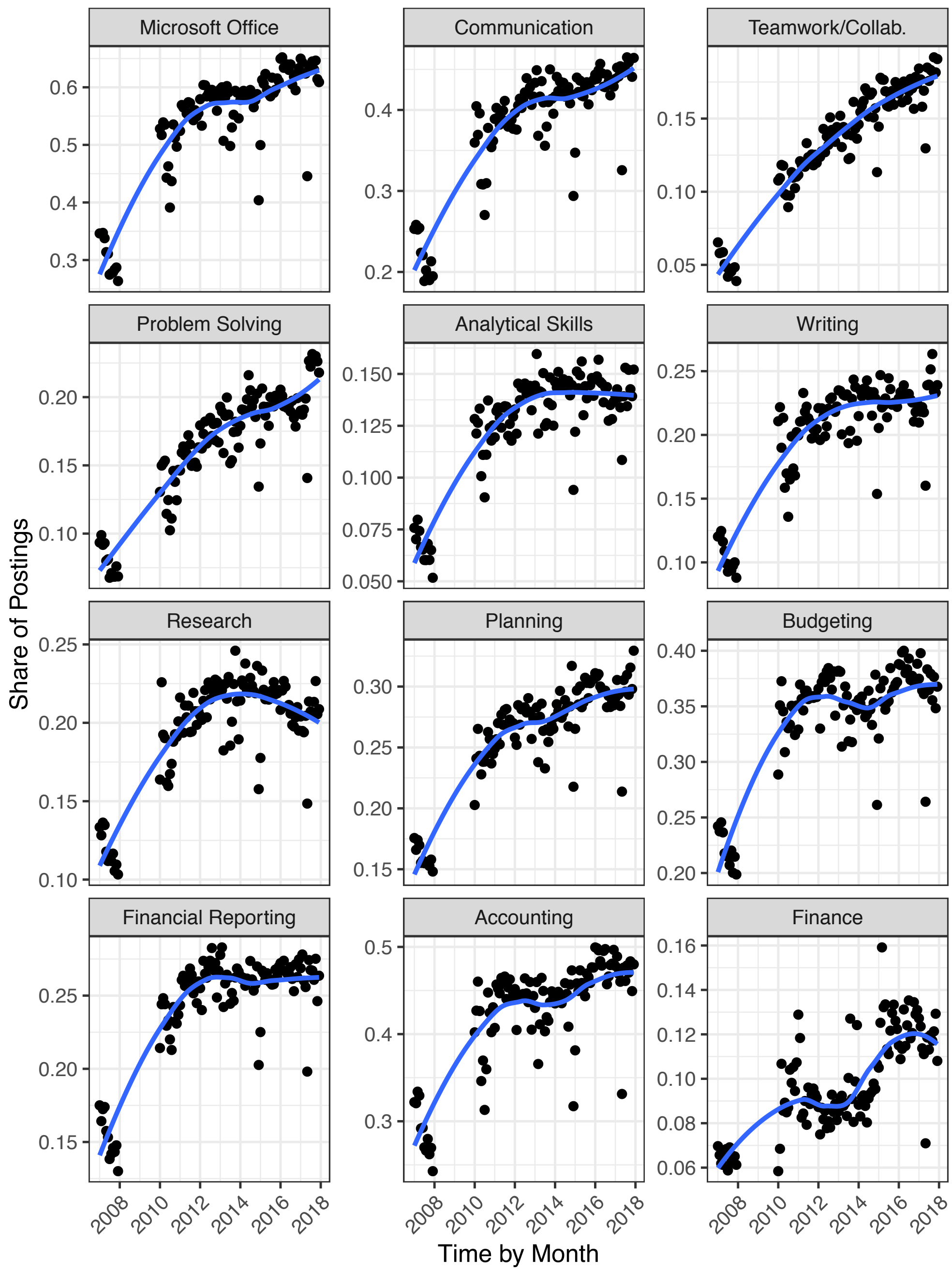
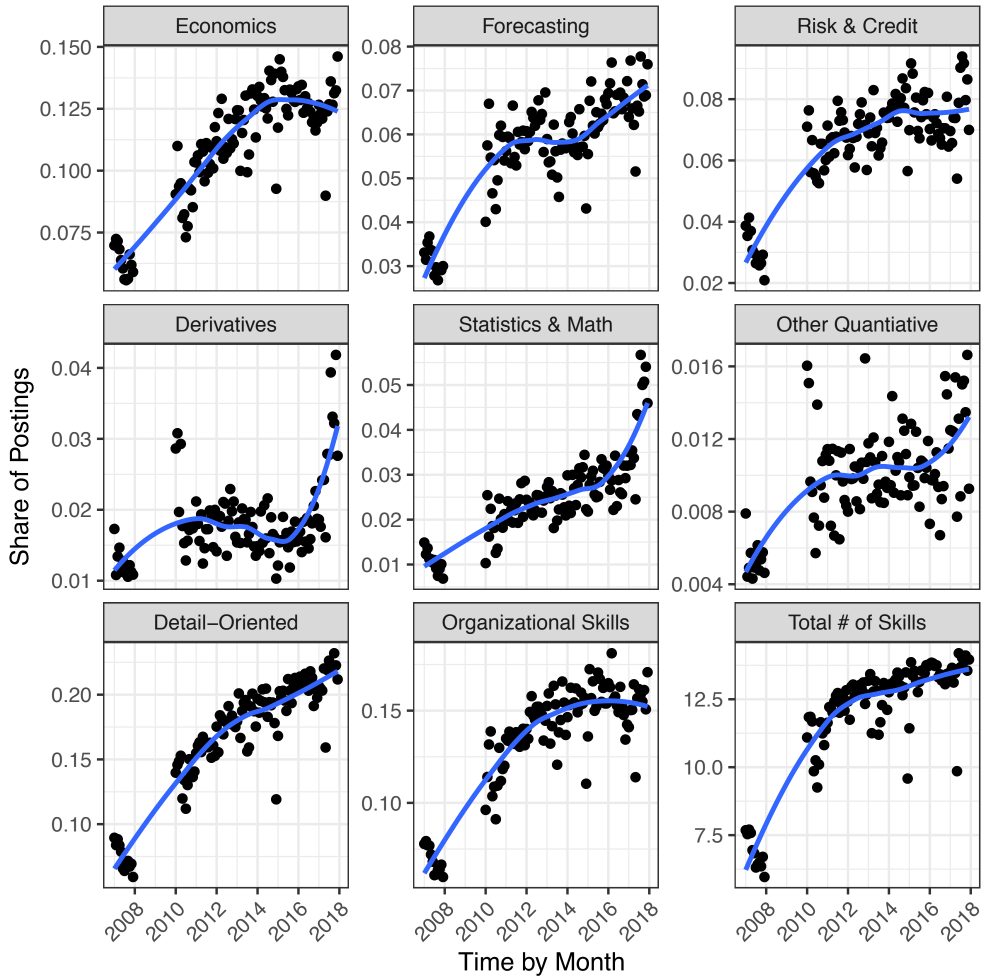


Figure C.2: *(Continued) Financial Analysts - Skill Composites Over Time*

Continued on next page



Notes: This figure shows the share of job postings for financial analysts that mention each of the skill composites and the total number of skills by month in 2007 and 2010-2017 (dots), along with fitted local polynomial regressions (lines).

Figure C.2: *(Continued) Financial Analysts - Skill Composites Over Time*

Table C.1: *Construction of Skill Composites*

Skill Composite	Skills	Occupation
Algorithmic Technology		
Marketing Automation	Marketo, Eloqua, Hubspot, Salesforce Marketing Cloud	Marketing Managers
Programming	SQL, SAS, SPSS, STATA, MATLAB, Python, R, C++, S-Plus, Perl, Ruby, Scala, C#, Debugging, .NET Programming, Statistical Programming	Financial Analysts
Software with Coding Required		
Programming	*See above*	Marketing Managers
Software Development	Software Development, Scrum, Agile Development, Software Engineering, Atlassian JIRA, Systems Development Life Cycle, Systems Analysis, Platform as a Service (PaaS), Object-Oriented Development Software, Software Architecture	Both
Machine Learning/AI	Machine Learning, Artificial Intelligence	Both
Operating Systems	Unix, Linux, Shell Scripting, Microsoft Windows	Financial Analysts
Distributed Computing	Apache Hadoop, Mapreduce, Apache Hive, Apache Spark, Apache Impala	Financial Analysts
Web Development	Javascript, XML, HTML, CSS, Ericsson, PHP, User Interface Design	Both
Software with Data Manipulation		
Graphic Design	Adobe Photoshop, Adobe Indesign, Adobe Acrobat, Adobe Creative Suite, Adobe Illustrator, QuarkXPress	Marketing Managers
Excel Formulas	Pivot Tables, Macros, Visual BAsic, Vlookup	Financial Analysts
Database Software	Database Software, Sybase, Microsoft Access	Both
Data Visualization	Tableau, Data Visualization, Infographics, Qlikview	Both
A/B Testing	A/B Testing	Marketing Managers
Web Analytics	Web Analytics, Google Analytics, Omniture	Marketing Managers
Business Intelligence	Cognos Impromptu, IBM Cognos, Crystal Reports, BusinessObjects	Financial Analysts
Software with Point & Click		
Google Ads	Google Adwords	Marketing Managers
Financial Software	Hyperion, Peoplesoft, Accounting Software, Oracle Financials, Quickbooks, Factset	Financial Analysts
CRM	CRM, Salesforce, Siebel, Microsoft Dynamics, Netsuite, Software as a Service (SaaS), Sales Automation Software	Both
ERP and Related	Oracle, SAP, ERP, JD Edward, Confluence	Both

Continued on next page

Table C.1 – *Continued from previous page*

Skill Composite	Skills	Occupation
Microsoft Office/Wind.	Microsoft Office, Microsoft Excel, Microsoft Powerpoint, Microsoft Word, Microsoft Outlook, Microsoft Sharepoint, Microsoft Visio, Microsoft Project, Word Processing, Microsoft Windows (*Marketing Managers only*)	Both
Interpersonal		
Communication	Communication Skills	Both
Teamwork/Collab.	Teamwork, Collaboration	Both
Building Effect. Relat.	Building Effective Relationships	Marketing Managers
Customer Service	Customer Service	Marketing Managers
General Non-Routine Cognitive		
Problem Solving	Problem Solving	Both
Analytical Skills	Analytical Skills	Both
Writing	Writing, Written Communication	Both
Presentation Skills	Presentation Skills	Marketing Managers
Research	Research, Market Research (*Marketing Managers only*)	Both
Planning	Planning, Marketing Planning (*Marketing Managers only*), Financial Planning (*Financial Analysts only*)	Both
Creativity	Creativity	Marketing Managers
Budgeting	*See below*	Financial Analysts
Business Non-Routine Cognitive		
Budgeting	Budgeting	Marketing Managers
Product Manag./Dev.	Product Management, Product Development	Marketing Managers
Market Strategy	Market Strategy	Marketing Managers
Social Media	Facebook Youtube, LinkedIn, Social Media	Marketing Managers
Project Management	Project Management	Marketing Managers
Business Development	Business Development	Marketing Managers
Finance & Accounting		
Financial Reporting	Financial Reporting, Financial Statements	Financial Analysts
Accounting	Accounting, Fund Accounting	Financial Analysts
Finance	Finance	Financial Analysts
Quantitative Analysis		
Economics	Economics	Financial Analysts
Forecasting	Forecasting	Financial Analysts
Risk & Credit	Risk Management, Model Risk Management, Market Risk, Risk Modeling, Risk Assessment, Risk Mitigation and Analysis, Risk Management Framework Credit Risk, Credit Risk Modeling, Stress Testing	Financial Analysts
Derivatives	Derivatives	Financial Analysts

Continued on next page

Table C.1 – *Continued from previous page*

Skill Composite	Skills	Occupation
Statistics & Math	Statistics, Econometrics, Mathemathics	Financial Analysts
Other Quantitative	Physics, Quantitative Research, Optimization	Financial Analysts
Routine		
Detail-Oriented	Detail-Oriented	Both
Organization Skills	Organizational Skills	Both

Table C.2: *Marketing Managers - Summary Statistics*

Statistic	N	Mean	St. Dev.
Marketing Automation	1,274,917	0.024	0.152
Software w. Coding Req.	1,274,917	0.117	0.322
Software w. Data Manip.	1,274,917	0.097	0.295
Software w. Point & Click	1,274,917	0.309	0.462
Interpersonal	1,274,917	0.601	0.490
General Non-Rout. Cog.	1,274,917	0.692	0.461
Business Non-Rout. Cog.	1,274,917	0.739	0.439
Routine	1,274,917	0.224	0.417
Programming	1,274,917	0.044	0.206
Software Development	1,274,917	0.075	0.263
Machine Learning/AI	1,274,917	0.003	0.057
Web Development	1,274,917	0.017	0.130
Graphic Design	1,274,917	0.033	0.179
Database Software	1,274,917	0.020	0.140
Data Visualization	1,274,917	0.010	0.102
A/B Testing	1,274,917	0.009	0.094
Web Analytics	1,274,917	0.038	0.191
Google Ads	1,274,917	0.009	0.096
CRM	1,274,917	0.049	0.216
ERP and Related	1,274,917	0.164	0.370
Microsoft Office/Wind.	1,274,917	0.247	0.431
Communication	1,274,917	0.430	0.495
Teamwork/Collab.	1,274,917	0.281	0.449
Building Effect. Relat.	1,274,917	0.113	0.316
Customer Service	1,274,917	0.099	0.299
Problem Solving	1,274,917	0.146	0.353
Analytical Skills	1,274,917	0.089	0.285
Writing	1,274,917	0.248	0.432
Presentation Skills	1,274,917	0.126	0.332
Research	1,274,917	0.277	0.448
Planning	1,274,917	0.294	0.456
Creativity	1,274,917	0.263	0.440
Budgeting	1,274,917	0.240	0.427
Product Manage./Dev.	1,274,917	0.410	0.492

Table C.2: *Marketing Managers - Summary Statistics*

Business Development	1,274,917	0.114	0.318
Market Strategy	1,274,917	0.152	0.359
Social Media	1,274,917	0.123	0.329
Project Management	1,274,917	0.226	0.418
Detail-Oriented	1,274,917	0.131	0.337
Organizational Skills	1,274,917	0.143	0.351
Total # of Skills	1,274,917	13.879	8.806
No Educ. or <BA	1,274,917	0.381	0.486
BA Degree	1,274,917	0.564	0.496
Master's Deg. or More	1,274,917	0.055	0.227
No Experience Req.	1,274,917	0.000	0.000
0-5 Years Experience	1,274,917	0.409	0.492
5-10 Years Experience	1,274,917	0.211	0.408
10-15 Years Experience	1,274,917	0.022	0.147
Experience Not Mentioned	1,274,917	0.358	0.479

Notes: This table shows summary statistics for job postings for marketing managers (SOC code 11-2021) in the time period under study (2007 and 2010-2017).

Table C.3: *Marketing Managers - Coefficients on Skill Composites*

	Outcome: Marketing Automation			
	Time	Employer	Ed. & Exp.	MSA
	(1)	(2)	(3)	(4)
Programming	-0.004^{**}	-0.002	-0.002	-0.003^{*}
	(0.001)	(0.001)	(0.001)	(0.001)
Software Development	-0.013^{**}	-0.011^{**}	-0.011^{**}	-0.013^{**}
	(0.0005)	(0.002)	(0.002)	(0.002)
Machine Learning/AI	-0.020^{**}	-0.021^{**}	-0.020^{**}	-0.022^{**}
	(0.002)	(0.003)	(0.003)	(0.003)
Web Development	0.004^{**}	0.005^{*}	0.005^{+}	0.005^{+}
	(0.001)	(0.003)	(0.003)	(0.003)
Graphic Design	0.003^{**}	0.001	0.0004	0.0004
	(0.001)	(0.002)	(0.002)	(0.002)
Database Software	-0.009^{**}	-0.009^{**}	-0.010^{**}	-0.009^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Data Visualization	0.041^{**}	0.034^{**}	0.034^{**}	0.032^{**}
	(0.001)	(0.005)	(0.005)	(0.005)
A/B Testing	0.107^{**}	0.105^{**}	0.104^{**}	0.103^{**}
	(0.001)	(0.006)	(0.006)	(0.006)
Web Analytics	0.055^{**}	0.058^{**}	0.057^{**}	0.057^{**}
	(0.001)	(0.003)	(0.003)	(0.003)
Google Ads	0.036^{**}	0.043^{**}	0.042^{**}	0.040^{**}
	(0.001)	(0.005)	(0.005)	(0.005)
CRM	0.272^{**}	0.273^{**}	0.272^{**}	0.271^{**}
	(0.001)	(0.004)	(0.004)	(0.004)
ERP and Related	0.0001	0.0003	0.0003	0.0001
	(0.0004)	(0.002)	(0.002)	(0.002)
Microsoft Office/Wind.	-0.014^{**}	-0.013^{**}	-0.014^{**}	-0.012^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Communication	-0.006^{**}	-0.006^{**}	-0.006^{**}	-0.006^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Teamwork/Collab.	-0.001^{*}	-0.001	-0.0005	-0.001^{+}
	(0.0003)	(0.001)	(0.001)	(0.001)
Building Effect. Relat.	-0.014^{**}	-0.015^{**}	-0.014^{**}	-0.014^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)

Table C.3: *Marketing Managers - Coefficients on Skill Composites*

Customer Service	-0.020^{**}	-0.019^{**}	-0.019^{**}	-0.019^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Problem Solving	-0.007^{**}	-0.007^{**}	-0.007^{**}	-0.007^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Analytical Skills	-0.001^{+}	-0.00002	0.0001	-0.0002
	(0.0004)	(0.001)	(0.001)	(0.001)
Writing	-0.002^{**}	-0.002^{**}	-0.002^{**}	-0.003^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Presentation Skills	-0.009^{**}	-0.008^{**}	-0.008^{**}	-0.009^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Research	-0.016^{**}	-0.017^{**}	-0.016^{**}	-0.016^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Planning	-0.005^{**}	-0.005^{**}	-0.005^{**}	-0.005^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Creativity	0.004^{**}	0.003^{**}	0.003^{**}	0.003^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Budgeting	-0.009^{**}	-0.008^{**}	-0.008^{**}	-0.007^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Product Manage./Dev.	-0.015^{**}	-0.015^{**}	-0.015^{**}	-0.017^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Business Development	-0.008^{**}	-0.007^{**}	-0.006^{**}	-0.007^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Market Strategy	-0.005^{**}	-0.006^{**}	-0.006^{**}	-0.006^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Social Media	0.004^{**}	0.004^{**}	0.004^{**}	0.003^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Project Management	0.003^{**}	0.003^{**}	0.003^{**}	0.003^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Detail-Oriented	0.006^{**}	0.007^{**}	0.007^{**}	0.006^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Organizational Skills	-0.004^{**}	-0.002^{**}	-0.003^{**}	-0.002^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Total # of Skills	0.002^{**}	0.002^{**}	0.002^{**}	0.002^{**}
	(0.00003)	(0.0001)	(0.0001)	(0.0001)
Master's Deg. or More			-0.004^{**}	-0.005^{**}
			(0.001)	(0.001)

Table C.3: *Marketing Managers - Coefficients on Skill Composites*

No Educ. or <BA			0.002^{**}	0.003^{**}
			(0.001)	(0.001)
No Experience Req.			-0.005^{**}	-0.006^{**}
			(0.001)	(0.001)
5-10 Years Experience			-0.008^{**}	-0.009^{**}
			(0.002)	(0.002)
10-15 Years Experience			-0.004^{**}	-0.004^{**}
			(0.001)	(0.001)
Time Effects	X	X	X	X
Employer Effects		X	X	X
Education & Experience			X	X
MSA Effects				X
Observations	1,274,917	875,996	875,996	856,812
R^2	0.216	0.233	0.233	0.236
Adjusted R^2	0.216	0.229	0.229	0.231

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table shows linear probability estimates of Equation 3.1 using job postings for marketing managers and the skill composites. Columns (2)-(4) cluster standard errors by employer. The reference category for the education variables is requiring at minimum a bachelor's degree. The references category for the experience variables is requiring 0-5 years of experience.

Table C.4: *Marketing Managers - Coefficients on Skill Categories*

	Outcome: Marketing Automation			
	Time	Employer	Ed. & Exp.	MSA
	(1)	(2)	(3)	(4)
Software w. Coding Req.	-0.006^{**}	-0.004^{**}	-0.004^{**}	-0.009^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Software w. Data Manip.	0.046^{**}	0.045^{**}	0.044^{**}	0.044^{**}
	(0.0005)	(0.002)	(0.002)	(0.002)
Software w. Point & Click	0.033^{**}	0.032^{**}	0.032^{**}	0.033^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Interpersonal	-0.013^{**}	-0.012^{**}	-0.011^{**}	-0.012^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
General Non-Rout. Cog.	-0.013^{**}	-0.014^{**}	-0.013^{**}	-0.014^{**}
	(0.0004)	(0.001)	(0.001)	(0.001)
Business Non-Rout. Cog.	-0.015^{**}	-0.015^{**}	-0.014^{**}	-0.016^{**}
	(0.0003)	(0.001)	(0.001)	(0.001)
Routine	-0.001^{**}	0.001	0.0004	0.0001
	(0.0003)	(0.001)	(0.001)	(0.001)
Time Effects	X	X	X	X
Employer Effects		X	X	X
Education & Experience			X	X
MSA Effects				X
Observations	1,274,917	875,996	875,996	856,812
R^2	0.054	0.070	0.070	0.078
Adjusted R^2	0.054	0.064	0.065	0.071

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table shows linear probability estimates of Equation 3.1 using job postings for marketing managers and the skill categories. Columns (2)-(4) cluster standard errors by employer. The regressions also control for the total number of skills in each posting and for education and experience requirements as indicated.

Table C.5: *Financial Analysts - Summary Statistics*

Statistic	N	Mean	St. Dev.
Quant. Finan. Analyst	865,508	0.027	0.163
Programming	865,508	0.084	0.278
Software w. Coding Req.	865,508	0.041	0.197
Software w. Data Manip.	865,508	0.202	0.402
Software w. Point & Click	865,508	0.607	0.488
Interpersonal	865,508	0.444	0.497
General Non-Rout. Cog.	865,508	0.693	0.461
Finance & Accounting	865,508	0.543	0.498
Quantitative Analysis	865,508	0.233	0.423
Routine	865,508	0.252	0.434
Software Development	865,508	0.027	0.162
Machine Learning/AI	865,508	0.001	0.038
Operating Systems	865,508	0.012	0.107
Distributed Computing	865,508	0.001	0.028
Web Development	865,508	0.004	0.065
Excel Formulas	865,508	0.074	0.262
Database Software	865,508	0.109	0.312
Data Visualization	865,508	0.009	0.094
Business Intelligence	865,508	0.047	0.211
Financial Software	865,508	0.120	0.325
CRM	865,508	0.020	0.140
ERP and Related	865,508	0.302	0.459
Microsoft Office	865,508	0.551	0.497
Communication	865,508	0.391	0.488
Teamwork/Collab.	865,508	0.139	0.346
Problem Solving	865,508	0.170	0.376
Analytical Skills	865,508	0.129	0.335
Writing	865,508	0.206	0.405
Research	865,508	0.199	0.399
Planning	865,508	0.266	0.442
Budgeting	865,508	0.344	0.475
Financial Reporting	865,508	0.247	0.431
Accounting	865,508	0.430	0.495
Finance	865,508	0.101	0.301

Table C.5: *Financial Analysts - Summary Statistics*

Economics	865,508	0.113	0.316
Forecasting	865,508	0.058	0.235
Risk & Credit	865,508	0.067	0.251
Derivatives	865,508	0.018	0.133
Statistics & Math	865,508	0.026	0.159
Other Quantiative	865,508	0.010	0.100
Detail-Oriented	865,508	0.174	0.379
Organizational Skills	865,508	0.138	0.345
Total # of Skills	865,508	12.131	7.558
No Educ. or <BA	865,508	0.289	0.454
BA Degree	865,508	0.658	0.474
Master's Deg. or More	865,508	0.052	0.223
No Experience Req.	865,508	0.000	0.000
0-5 Years Experience	865,508	0.544	0.498
5-10 Years Experience	865,508	0.082	0.274
10-15 Years Experience	865,508	0.006	0.080
Experience Not Mentioned	865,508	0.368	0.482

Notes: This table shows summary statistics for job postings for financial analysts (SOC code 13-2051 and O*NET SOC code 13-2099.01) in the time period under study (2007 and 2010-2017).

Table C.6: *Financial Analysts - Coefficients on Skill Composites*

	Outcome: Programming			
	Time	Employer	Ed. & Exp.	MSA
	(1)	(2)	(3)	(4)
Software Development	0.126**	0.123**	0.125**	0.125**
	(0.002)	(0.005)	(0.005)	(0.005)
Machine Learning/AI	0.292**	0.250**	0.238**	0.238**
	(0.007)	(0.024)	(0.023)	(0.023)
Operating Systems	0.244**	0.212**	0.212**	0.213**
	(0.003)	(0.007)	(0.007)	(0.007)
Distributed Computing	0.374**	0.309**	0.309**	0.300**
	(0.010)	(0.033)	(0.032)	(0.033)
Web Development	0.317**	0.293**	0.293**	0.296**
	(0.004)	(0.014)	(0.014)	(0.014)
Excel Formulas	0.165**	0.187**	0.185**	0.184**
	(0.001)	(0.004)	(0.004)	(0.004)
Database Software	0.102**	0.103**	0.102**	0.102**
	(0.001)	(0.002)	(0.002)	(0.002)
Data Visualization	0.246**	0.282**	0.282**	0.278**
	(0.003)	(0.008)	(0.008)	(0.008)
Business Intelligence	0.105**	0.103**	0.103**	0.101**
	(0.001)	(0.004)	(0.004)	(0.004)
Financial Software	−0.014**	−0.012**	−0.013**	−0.014**
	(0.001)	(0.002)	(0.002)	(0.002)
CRM	−0.016**	−0.002	−0.004	−0.006
	(0.002)	(0.004)	(0.004)	(0.004)
ERP and Related	−0.023**	−0.021**	−0.021**	−0.021**
	(0.001)	(0.001)	(0.001)	(0.001)
Microsoft Office	−0.012**	−0.007**	−0.007**	−0.008**
	(0.001)	(0.001)	(0.001)	(0.001)
Communication	−0.015**	−0.014**	−0.016**	−0.016**
	(0.001)	(0.001)	(0.001)	(0.001)
Teamwork/Collab.	−0.011**	−0.011**	−0.012**	−0.012**
	(0.001)	(0.002)	(0.002)	(0.002)
Problem Solving	0.001	0.003*	0.003*	0.002
	(0.001)	(0.002)	(0.002)	(0.002)

Table C.6: *Financial Analysts - Coefficients on Skill Composites*

Analytical Skills	-0.008^{**}	-0.008^{**}	-0.008^{**}	-0.009^{**}
	(0.001)	(0.002)	(0.002)	(0.001)
Writing	-0.011^{**}	-0.014^{**}	-0.014^{**}	-0.015^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Research	-0.007^{**}	-0.003^{*}	-0.006^{**}	-0.005^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Planning	-0.020^{**}	-0.015^{**}	-0.016^{**}	-0.017^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Budgeting	-0.027^{**}	-0.028^{**}	-0.029^{**}	-0.029^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Financial Reporting	-0.019^{**}	-0.018^{**}	-0.019^{**}	-0.020^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Accounting	-0.048^{**}	-0.047^{**}	-0.048^{**}	-0.048^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Finance	-0.008^{**}	-0.003^{+}	-0.004^{+}	-0.005^{**}
	(0.001)	(0.002)	(0.002)	(0.002)
Economics	0.037^{**}	0.040^{**}	0.037^{**}	0.036^{**}
	(0.001)	(0.002)	(0.002)	(0.002)
Forecasting	-0.022^{**}	-0.021^{**}	-0.021^{**}	-0.022^{**}
	(0.001)	(0.002)	(0.002)	(0.002)
Risk & Credit	0.071^{**}	0.058^{**}	0.055^{**}	0.053^{**}
	(0.001)	(0.004)	(0.004)	(0.003)
Derivatives	0.024^{**}	0.031^{**}	0.027^{**}	0.024^{**}
	(0.002)	(0.005)	(0.005)	(0.005)
Statistics & Math	0.143^{**}	0.157^{**}	0.151^{**}	0.151^{**}
	(0.002)	(0.006)	(0.006)	(0.006)
Other Quantiative	0.407^{**}	0.312^{**}	0.289^{**}	0.287^{**}
	(0.003)	(0.008)	(0.008)	(0.008)
Detail-Oriented	-0.002^{**}	0.001	0.001	-0.002
	(0.001)	(0.002)	(0.002)	(0.002)
Organizational Skills	-0.024^{**}	-0.023^{**}	-0.023^{**}	-0.023^{**}
	(0.001)	(0.002)	(0.002)	(0.002)
Total # of Skills	0.007^{**}	0.006^{**}	0.006^{**}	0.007^{**}
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Master's Deg. or More			0.062^{**}	0.060^{**}
			(0.004)	(0.004)

Table C.6: *Financial Analysts - Coefficients on Skill Composites*

No Educ. or <BA			-0.009^{**}	-0.010^{**}
			(0.001)	(0.001)
No Experience Req.			-0.018^{**}	-0.017^{**}
			(0.002)	(0.002)
5-10 Years Experience			-0.035^{**}	-0.032^{**}
			(0.005)	(0.005)
10-15 Years Experience			-0.005^{**}	-0.005^{**}
			(0.001)	(0.001)
Time Effects	X	X	X	X
Employer Effects		X	X	X
Education & Experience			X	X
MSA Effects				X
Observations	865,508	532,675	532,675	519,799
R^2	0.219	0.262	0.265	0.271
Adjusted R^2	0.219	0.257	0.260	0.265

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table shows linear probability estimates of Equation 3.1 using job postings for financial analysts and the skill composites. Columns (2)-(4) cluster standard errors by employer. The reference category for the education variables is requiring at minimum a bachelor's degree. The references category for the experience variables is requiring 0-5 years of experience.

Table C.7: *Financial Analysts - Coefficients on Skill Categories*

	Outcome: Programming			
	Time	Employer	Ed. & Exp.	MSA
	(1)	(2)	(3)	(4)
Software w. Coding Req.	0.219^{**}	0.201^{**}	0.200^{**}	0.200^{**}
	(0.001)	(0.005)	(0.005)	(0.005)
Software w. Data Manip.	0.148^{**}	0.157^{**}	0.156^{**}	0.154^{**}
	(0.001)	(0.002)	(0.002)	(0.002)
Software w. Point & Click	-0.026^{**}	-0.020^{**}	-0.018^{**}	-0.019^{**}
	(0.001)	(0.002)	(0.002)	(0.002)
Interpersonal	-0.016^{**}	-0.017^{**}	-0.019^{**}	-0.020^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
General Non-Rout. Cog.	-0.024^{**}	-0.019^{**}	-0.022^{**}	-0.024^{**}
	(0.001)	(0.001)	(0.002)	(0.002)
Finance & Accounting	-0.078^{**}	-0.075^{**}	-0.074^{**}	-0.074^{**}
	(0.001)	(0.002)	(0.002)	(0.002)
Quantitative Analysis	0.066^{**}	0.061^{**}	0.056^{**}	0.053^{**}
	(0.001)	(0.003)	(0.002)	(0.002)
Routine	-0.019^{**}	-0.014^{**}	-0.013^{**}	-0.015^{**}
	(0.001)	(0.001)	(0.001)	(0.001)
Time Effects	X	X	X	X
Employer Effects		X	X	X
Education & Experience			X	X
MSA Effects				X
Observations	865,508	532,675	532,675	519,799
R^2	0.143	0.196	0.203	0.211
Adjusted R^2	0.143	0.191	0.198	0.204

Notes: $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$. This table shows linear probability estimates of Equation 3.1 using job postings for financial analysts and the skill categories. Columns (2)-(4) cluster standard errors by employer. The regressions also control for the total number of skills in each posting and for education and experience requirements as indicated.