March 23, 2020

Dear Members of the Selection Committee,

I am writing to apply for the Review of Economics and Statistics Fellowship for the National Bureau of Economic Research Summer Institute (SI). I am a female, African-American Assistant Professor at Florida State University in the Department of Economics who studies how stereotypes and discrimination affect job candidates, students, and other members of society. Therefore, my research interests fit well with the SI sessions on Labor Studies, Gender in the Economy, and Education.

My job market paper determines whether job applicants can accurately predict how much discrimination they will face when applying to stereotypically male jobs (e.g., engineers). Subjects in my experiment were asked to act as either workers or employers. Workers' profiles were sent to employers who were tasked to hire someone for a stereotypically male job. Before the employers saw the profiles, workers had the opportunity to pay to edit their profiles to either include or exclude gender. I find that female workers were willing to pay more to use profiles without gender than profiles with gender. The opposite was true for male workers. However, male workers, in particular, overestimated how much of a financial difference it made to change their profiles. Therefore, if misestimating the cost of being female causes men to select into stereotypically male jobs more than equally qualified women, then this could explain part of why occupational segregation occurs. This paper will be submitted this summer.

The rest of my work continues to study stereotypes and discrimination. I am currently developing a new experiment to measure perceptions of police bias and the impact these perceptions have on civilians' behavior. I am also finishing a paper on the effect of stereotypes on the test performance of black college students, which I plan to submit this spring.

Having had the privilege of attending the SI as a doctoral student, I would now benefit even more from attending as a new faculty member. I would be able to share my research findings with others, receive feedback on my work from top scholars in my field, find potential co-authors, and learn about current, ground-breaking research. This will help strengthen and advance both my research and my professional development.

I have attached my application materials to this file. Please feel free to contact me at 262-573-2830 or malston@fsu.edu if you require additional information. Thank you for your time and consideration.

Sincerely,

Mackenzie Alston

Name: Mackenzie Alston
Email Address: malston@fsu.edu
Current University Affiliation: Florida State University, Department of Economics
Date of Ph.D.: May 2019
Link to Personal Website: www.mackenziealston.com
Sessions Most Interested in Attending (in order): Labor Studies, Gender in the Economy, Education

March 2020

MACKENZIE ALSTON

CONTACT INFORMATION

Department of Economics
Florida State University
Bellamy Building 258
Tallahassee, FL 32306

www.mackenziealston.com
Office: 850-644-5002
Email: malston@fsu.edu
Citizenship: USA

EMPLOYMENT

Assistant Professor, Florida State University — August 2019 – present

EDUCATION

Ph.D. in Economics, Texas A&M University — May 2019
A.B. in Economics and Mathematics (cum laude), Washington University in Saint Louis — December 2013

FIELDS OF SPECIALIZATION

Experimental Economics, Behavioral Economics, Labor, Education

PUBLICATIONS

Alston, Mackenzie. "The Effect of Dropping Lowest Test Scores on Learning and Performance," *Equilibria*, 2015, 3(1): 13-18.

WORKING PAPERS

"The (Perceived) Cost of Being Female: An Experimental Investigation of Strategic Responses to Discrimination"

"The Effect of Stereotypes on Test Scores at a Historically Black University" with William Darity Jr., Catherine Eckel, Lawrence McNeil, and Rhonda Sharpe

"Fired Up?: How Past Unemployment Affects Future Work Motivation and Productivity" (Submitted)

"High-Capacity Donors' Preferences for Charitable Giving" with Catherine Eckel, Jonathan Meer, and Wei Zhan (NBER Working Paper 25290, Revise & Resubmit)

WORKS IN PROGRESS

"Identity, Stereotype Threat, and Black College Student Success: Evidence from a Longitudinal Survey" with William Darity Jr., Catherine Eckel, Lawrence McNeil, Mark Paul, and Rhonda Sharpe

"My Dream Police Officer: An Experimental Investigation of Perceptions of Police Bias"

PRESENTATIONS

*indicates participation in panel

2020: AEA/ASSA Annual Meeting (discussant)

2019: AEA/ASSA Annual Meeting, AEA Mentoring Pipeline Conference*, Southern Economic Association Conference, Science of Diversity & Inclusion Initiative Conference

2018: Virginia Tech Seminar, University of West Virginia Seminar, Southern Economic Association Conference*, UCSD Rady School of Management Seminar, AEA Mentoring Program Job Market Bootcamp,

Norms and Behavioral Change Workshop, AEA Mentoring Pipeline Conference, World Economic Science Association Conference, Spring School in Behavioral Economics Poster Conference

2017: Southern Economic Association Conference, Science of Philanthropy Initiative Conference, AEA Mentoring Pipeline Conference, World Economic Science Association Conference, Texas Experimental Association Symposium

2016: North American Economic Science Association Conference, American Society of Hispanic Economists Conference

2013: Ohio Undergraduate Research Symposium Poster Conference

2012: AEA Mentoring Pipeline Conference

FELLOWSHIPS AND AWARDS

Undergraduate Research Opportunities Program Materials Grant ($500)	February 2020
Provost's Faculty Travel Grant, Florida State University	November 2019
AEA Mentoring Program Travel Awards, AEA Pipeline Mentoring	2017 - 2018
Graduate Diversity Fellowship, Texas A&M University	2014 - 2017
Equity and Inclusion Fellowship, Association for Public Policy Analysis & Management	November 2017
John Van Huyck Fellowship, Texas A&M University	October 2017
Teaching Fellowship, AEA Summer Training Program, Michigan State University	July 2017
M.J. Grove Fellowship, Private Enterprise Research Center, Texas A&M University	Summer 2016
Outstanding July Teaching Fellow Award (Foundation Level), AEA Summer Training Program, Michigan State University	July 2016
Teaching Fellowship, AEA Summer Training Program, Michigan State University	July 2016
Lechner Graduate Grant, Texas A&M University	Fall 2014

RESEARCH EXPERIENCE

Research Assistant for Catherine Eckel Texas A&M University	2014 - 2018
Laboratory Assistant Missouri Social Science Experimental Laboratory, Washington University in Saint Louis	2012 - 2013

	Research Assistant for Katie Coffman The Ohio State University	Summer 2013
TEACHING EXPERIENCE	*Instructor, Florida State University* Labor Economics (Undergraduate) [4.59/5]	Fall 2019
	Instructor, Texas A&M University Experimental Economics (Undergraduate) [4.92/5]	Summer 2018
	Co-Instructor, AEA Summer Training Program, Michigan State University Foundations of Mathematical Methods for Economic Analysis (Undergraduate) [4.82/5]	July 2017
	Teaching Fellow, AEA Summer Training Program, Michigan State University Foundations of Mathematical Methods for Economic Analysis (Undergraduate) [4.86/5]	July 2016
	Teaching Assistant, Texas A&M University Economic Data Analysis (Undergraduate) Experimental Economics (Graduate) Research in Experimental Economics (Undergraduate) Experimental Economics (Undergraduate)	 Fall 2018 Fall 2016 Spring 2016 Fall 2015
WORKSHOP PARTICIPATION	Spring School in Behavioral Economics, University of California, San Diego	March 2018
	6th Lindau Meeting on Economic Sciences (Lindau, Germany)	August 2017
	Meta-Analysis Psychology Summer Workshop, Texas A&M University	May 2017
	Barcelona BESLab Experimental Economics Summer School in Macroeconomics	July 2016
	The International Foundation for Research in Experimental Economics (IFREE) 21st Annual Visiting Graduate Student Workshop in Experimental Economics	January 2016
	AEA Summer Training Program, University of New Mexico	Summer 2012
REFEREE EXPERIENCE	*Quarterly Journal of Economics*, *Industrial Relations*, *Southern Economic Journal*	

The (Perceived) Cost of Being Female: An Experimental Investigation of Strategic Responses to Discrimination

Mackenzie Alston*

Version Date: April 9, 2019

Current Version Available at JMP.mackenziealston.com

Abstract

Occupational gender segregation and wage gaps may be caused by women applying to certain jobs less than men or women being discriminated against during the hiring process. I estimate how much discrimination job applicants expect to face by giving them the opportunity to include or exclude gender from their resumes in an online experiment where subjects perform a stereotypically male task. On average, men were willing to forfeit 9.4% of their earnings from being hired to use resumes that included gender while women were willing to forfeit 10.0% of their earnings to use resumes that excluded gender. However, workers did not earn significantly more when they included rather than excluded their gender. Therefore, workers, men especially, overestimated the cost of being female in this study, suggesting that workers may not be applying for stereotypically male jobs at optimal rates.

JEL: J71, M51, C91

*Florida State University, Department of Economics, Bellamy Building 258, Tallahassee, Florida 32306, malston@fsu.edu. I am extremely grateful for the guidance and support given to me by Catherine Eckel and Katie Coffman. I would also like to thank Bethany Patterson, whose attention to detail was invaluable for this project. This project was supported by the National Science Foundation, the AEA Mentoring Program (NSF Awards #1357478 & 1730651), and the John Van Huyck Award.

1. Introduction

Of the 337 occupations listed in the 2017 United States Current Population Survey, 56% of them have a 50% or higher share of men. But it is not clear from these statistics whether the source of this occupational gender segregation is the result of discrimination against women or women's occupational preferences. Previous research has tried to understand women's demand for female-dominated jobs through differences in ability or preferences (e.g., Paglin and Rufolo, 1990; Niederle and Vesterlund, 2007). However, past literature has also shown that women have been discriminated against when applying for high-paying and male-oriented jobs (e.g., Riach and Rich, 2006; Neumark, Bank, and Van Nort, 1996). This paper explores the role of *anticipated* discrimination as a factor that could explain why there is an underrepresentation of females in certain fields. That is, men and women may make their decision to pursue fields like science, which are male-dominated, based on their perception of whether or not employers are gender biased. If there is a perception that men will be preferred when they apply for these careers, then men may choose to pursue these fields at a higher rate than women, even after controlling for differences in ability and preferences. If people's perceptions are false, then this becomes particularly problematic.

To test whether job applicants can accurately predict how much discrimination exists when applying for stereotypically male jobs, I conducted two experiments using Amazon Mechanical Turk (MTurk), an online job platform hosted by Amazon. I recruited participants from MTurk to act as workers and hiring managers. The workers completed a stereotypically male task and answered questions that were used to build two resume-like profiles. These profiles included information like their age and past volunteer activity. The profiles also included a noisy signal of the workers' ability, a measure of how well they performed on half of the task. The two profiles were identical, except one did not include the worker's gender.

Each male worker was paired with a female worker who had a similar signal of ability. Their profiles were then shared with a participant from the hiring manager experiment, who had to choose which of the two workers to hire. Workers who were hired were paid $0.40 for each correct answer on the task. If they were not hired, they were only paid $0.10 for each correct answer.

Before the managers made their decision, I elicited the workers' preference for the managers to see the profiles with or without gender. The workers were asked to consider a

situation in which their manager was originally supposed to receive review the profiles that included gender. The workers were then asked how much they would be willing to pay so that the managers saw the profiles *without* gender instead. Next, the workers were asked to consider the case where managers were initially supposed to see profiles that excluded gender. They then indicated how much they would be willing to pay so that the managers saw the profiles *with* gender instead.

The workers' maximum willingness to pay to change the profiles was used to measure how much discrimination they believed existed in this study. If the workers believed that the managers would discriminate against women on the basis of their gender, then the workers had an incentive to change which version of the profiles the manager received. Female workers should be willing to pay to remove gender from their profiles if they believed that hiding gender would increase their chances of being hired. On the other hand, male workers should be interested in using profiles that included gender.

I find that the workers believed that female workers would be disadvantaged during the hiring process due to their gender. In a post-experimental survey, nearly 40 percent of the workers expected men to be chosen more than women when gender was revealed to the managers, and only 8 percent believed that women would be chosen more. When managers did not know the gender of the workers, only 14 percent of workers expected men would be chosen more. More importantly, women were willing to forfeit 10.0 percent of their earnings from being hired to use gender-free profiles. Meanwhile, men were willing to forfeit 9.4 percent of their earnings from being hired to use profiles with gender. In other words, women were willing to pay to make the hiring process gender-free, and men were willing to pay for the opposite. This suggests that both men and women anticipated that women would become targets of discrimination because of their gender.

I determined whether the managers were discriminating based on gender by comparing their hiring decisions when they saw two profiles with gender to when they saw two profiles without gender. Men were hired 66 percent of the time when the managers saw gender on the profiles and 61 percent of the time when the managers did not see gender. This difference is not statistically significant ($p=0.38$; two-tailed t-test). However, there were cases when it appeared as if the managers did let gender influence their decision, especially when they were forced to choose between two workers with similar signals of productivity.

Regardless, workers were both overestimated how much their gender would affect their earnings. The amount that the workers earned from the task when gender was revealed was not significantly different than the amount they earned when their gender was unknown. Therefore, workers should not have been willing to pay to change the profiles, yet I find that workers were interested in paying for the opportunity. In particular, men appeared to overestimate how much money they should spend to include gender on the profiles. This finding reflects the importance of capturing information about job applicants' perceptions of discrimination and not just the actual levels of discrimination that exists.

If job applicants' perceptions of gender discrimination are not consistent with the true levels of discrimination, then this could result in men applying for stereotypically male jobs at a higher rate than equally qualified women. Consequently, there would be an underrepresentation of females in certain tracks. Finding a way to adjust job applicants' anticipated discrimination to better reflect the true level of discrimination could then affect the gender composition in science, technology, engineering, and math (STEM) and other male-dominated fields. This may be of particular interest for companies, schools, and organizations that wish to increase diversity.[1]

2. Gender Discrimination in the Labor Market

Previous literature has tried to answer the question of whether or not men and women are treated differently by hiring managers. A common strategy researchers have used to study discrimination in the labor market is audit and correspondence studies. In these studies, researchers create fictitious resumes that are sent out to real-world companies. These resumes are designed such that they are nearly identical to each other aside from a particular characteristic of the job applicant. In correspondence studies examining gender discrimination, resumes differ by having female names like Ann and male names like Michael. Then researchers compare the percent of female job applicants who received calls back from the companies to the percent of male job applicants who received callbacks. If the callback rate for one gender is higher than the other, then that is evidence that companies were discriminating. This is indeed what Riach and Rich (2006) found for male-oriented jobs like engineering. Neumark et al. (1996) found that women were less likely to receive callbacks for high-paying jobs like wait staff at high-priced restaurants. Other studies like Goldin and Rouse (2000) have also shown that women can face

[1] Adding diversity to these fields could potentially bring several benefits like reducing gender-based differences in earnings (e.g., Kirkeboen et al., 2016; Arcidicano, 2004) and improving group performance (e.g., Woolley et al., 2010; Hoogendorn, Oosterbeck, and Van Praag, 2013).

labor discrimination using quasi-experimental data. Thus, there is evidence that there are cases in which employers favor men over women.

Because audit and correspondence studies focus on the employers' decisions, they are unable to provide insight into the beliefs of the job applicants. Using experiments and survey questions involving job applicants can help fill this gap. Kang et al. (2016) interviewed black and Asian university students who were searching for jobs and internships. They found that 31% of blacks and 40% of Asians admitted to "whitening" their resumes by modifying how their names were written and changing how they described their race- or culture-based extracurricular activities and experiences so that they sounded white or race-neutral. One common reason stated as to why they whitened their resume was to "even the playing field" or "get their foot in the door." This implies that they felt that other racial groups had an advantage over them and suspected that some employers would discriminate against them if their race were revealed.

In their third study, Kang et. al (2016) conducted a correspondence study where they varied the race of the fictitious job applicant as well as whether or not the resumes employers saw were whitened. Completely whitening one's resume increased black applicants' callback rates by 12.5 percentage points and Asian applicants' callback rate by 9.5 percentage points. Thus, this paper suggests that racial minorities may be correct in their belief that employers care about race. However, the study is unable to say whether their predictions about how much discrimination they will face are accurate. Neither does it discuss gender differences in the behavior of job applicants.

Zafar (2013) used data on sophomores at Northwestern University to understand how men and women choose their majors, which is an important step to determining their career. He found that women believed that they would be treated more poorly than men for careers in natural sciences, math and computer sciences, and engineering. But this was related to their beliefs about discrimination on-the-job instead of during the application process.

To my knowledge, this is the first paper to study whether job applicants anticipate that females will become targets of discrimination and determine whether or not their expectations are correct. By using two experiments involving both workers and hiring managers, I studied the supply and demand side of the same labor market. Consequently, these experiments allowed me to collect information that revealed how much discrimination workers believed existed in this

particular labor market as well as who was hired from that same market, a unique feature of this study.

Workers' beliefs about discrimination were captured through their responses to survey questions and their willingness to pay to have (or not have) a gender-blind hiring process. In addition, each hiring decision was recorded, so it was possible to calculate the likelihood that a female was hired if gender was hidden and when gender was revealed. This gave me the ability to determine whether workers were over- or underestimating how biased managers truly were.

Another advantage of this design was that I avoided some sample selection issues. If I asked workers applying for real-world stereotypically male jobs, then I would be restricted to analyzing individuals who had already decided that the benefit of applying for the job outweighed the costs (e.g., the risk of being discriminated against). In this study, the workers were given limited information about the experiment before agreeing to participate and thus I was able to observe the behavior of workers who might have otherwise perceived the cost of applying to be too high. Therefore, this study will add to the literature on discrimination and gender differences by providing information on whether or not job applicants are able to accurately predict to the level of discrimination that exists in stereotypically male fields.

3. Simple Framework for Job Applicants

When job applicants are deciding whether or not to apply to a job, there are many factors that may go into their decision-making process. For simplicity, consider the basic framework described in this section. Each applicant i has to make a choice to apply to a position or not. Denote this as $D \in \{0,1\}$ where $D = 0$ means that the applicant has chosen not to apply. If the applicant does apply (i.e, $D = 1$), then she pays a fixed cost, c, and will earn $P * x_i$. $P \in \{H, L\}$ can be thought of as a piece rate, and x_i is a measure of her ability or performance. If she is hired for the job, $P = H$. Otherwise, $P = L < H$. Without loss of generality, assume that $L \equiv 0$.

The applicant does not know the true probability in which she will be hired for a specific position. Let this probability be denoted as $0 \leq p(\boldsymbol{E}_i, G_i) \leq 1$, where $G_i \in \{M, F\}$. $G_i = M$ if the job applicant is male, and $G_i = F$ if the job applicant is female. $\boldsymbol{E}_i$ is a vector of variables in the hiring environment including the applicant's signal of ability $\tilde{\boldsymbol{x}}_i$ (e.g., information on her resume), the signal of ability sent by other individuals applying for the position $\tilde{\boldsymbol{x}}_j$, the other individuals' gender G_j, and other factors in the environment that may affect hiring decisions $\boldsymbol{O}$.

The applicant does not observe $p(\boldsymbol{E}_i, G_i)$; however, she has a belief about the probability that she will be hired. I model the belief as $\tilde{p}(\boldsymbol{E}_i, G_i) = p(\boldsymbol{E}_i, G_i) + \varepsilon$, where ε captures the noise.

The applicant's utility function can be written as follows:

$$(1)\ u(D, \boldsymbol{E}_i, G_i) = \begin{cases} \begin{cases} H * x_i - c, with\ probability\ p(\boldsymbol{E}_i, G_i) \\ L * x_i - c, with\ probability\ 1 - p(\boldsymbol{E}_i, G_i) \end{cases} & if\ D = 1 \\ L * x_i & if\ D = 0 \end{cases}$$

The applicant should apply to the position if $u(1, \boldsymbol{E}_i, G_i) > u(0, \boldsymbol{E}_i, G_i)$.

Without loss of generality, assume applicant *i* is female. Discrimination exists if $\tilde{\boldsymbol{x}}_i = \tilde{\boldsymbol{x}}_j$ but $p(\boldsymbol{E}_i, F) \neq p(\boldsymbol{E}_j, M)$. In other words, one gender is hired more frequently than another when all else is held equal. Applicant *i* perceives discrimination if $\tilde{\boldsymbol{x}}_i = \tilde{\boldsymbol{x}}_j$ but $\tilde{p}(\boldsymbol{E}_i, F) \neq \tilde{p}(\boldsymbol{E}_j, M)$. If, say, all applicants believe that $\tilde{p}(\boldsymbol{E}_i, F) < \tilde{p}(\boldsymbol{E}_j, M)$ when $\tilde{\boldsymbol{x}}_i = \tilde{\boldsymbol{x}}_j$, then more men will apply to that position than women, conditional on their signal of ability. This occupational segregation is not optimal if $\exists\ G_i$ s.t. $\tilde{p}(\boldsymbol{E}_i, G_i) \neq p(\boldsymbol{E}_i, G_i)$). If that is the case, then one gender is selecting out of the position more than they should.

To determine whether applicants can accurately predict how much discrimination exists when they apply for certain positions (i.e., $\exists\ G_i$ s.t. $\tilde{p}(\boldsymbol{E}_i, G_i) \neq p(\boldsymbol{E}_i, G_i)$), I designed an experiment where I give applicants the opportunity to hide their gender, thus allowing for $G_i \in \{\mathrm{M}, \mathrm{F}, \mathrm{N}\}$ where N represents the gender-neutral case where no one can tell if the applicant is male or female. N may seem artificial, but there are countries like France and Germany that are experimenting with anonymous job applications.[2] In this study, the gender-neutral case is relevant for understanding why there is an underrepresentation of females in certain careers. I ask whether people would mask their gender if it were possible. I use this to determine whether people overvalue or undervalue their gender in the labor market. However, there have been some places like

4. About the Mechanical Turk Sample

Subjects in this study were individuals recruited from Amazon Mechanical Turk (MTurk).[3] To be eligible to participate, they were required to be at least eighteen years old, be

[2] See Krause, Rinne, and Zimmermann (2012) and Rinne (2018) for a discussion of the advantages and disadvantages of European hiring policies that use anonymous job applications.

[3] Huff and Tingley (2015) and Levay, Freese, and Druckman (2016) compared MTurk's subject pool to that of other surveys conducted in the United States like the American National Election Studies. Importantly, Levay et al. (2016) found that differences between MTurk users and survey respondents shrunk once controls were added for age, gender, race, income, and education.

located in the United States, and have had satisfactory prior job performance on MTurk. Users on MTurk visit the platform to find short-term jobs for pay and can be more diverse than students who participate in laboratory experiments on college campuses. Therefore, MTurk provides a useful platform to study questions related to the labor market. Previous studies like Coffman et al. (2017) have also used it to study gender-based discrimination.

Table 1 describes the characteristics of this study's subjects. On average, subjects were approximately 35 years old. The majority were white, non-Hispanic, earned $1,000 to $3,000 per month, and had a bachelor's degree or higher. Most worked full-time, and roughly 40% were married. Details on the workers' characteristics by gender can be found in the Appendix.

Table 1. Subject Characteristics

[INSERT HERE]

5. Experimental Design

5.1 Experimental Design for Worker Experiment

This study was divided into two experiments, one for the workers and one for the managers. For the Worker experiment, 430 men and women located in the United States were recruited from Amazon Mechanical Turk to participate in a four-part experiment.[4] In Part One of the experiment, Workers answered questions that would help build resume-like profiles. They indicated their gender, age (as a range of values), whether they had volunteered in the past year, whether their former supervisor would consider them to be a good employee, and whether they frequently read the news. They received $0.50 for completing this part, and the full set of questions can be found in the Appendix.

After answering these questions, Workers read instructions for Part Two. They were told that they would have 2.5 minutes to answer ten questions. They would be paid either $0.10 or $0.40 for each correct answer and $0.00 for any incorrect answers. Workers were told that a third party would determine whether they would be awarded $0.10 or $0.40 per correct answer. To reduce the likelihood that Workers would condition their effort on the task based on their beliefs about their chances of earning $0.40, no other details about earnings were provided until Part Two was finished.

[4] Anyone who did not self-identify as male or female was not allowed to participate in the Worker experiment.

During Part Two, Workers were given 150 seconds to complete the stereotypically male task: a sports trivia quiz consisting of 10 multiple-choice questions.[5] After the quiz was over, subjects were presented with two nearly identical "profiles" (A and B), which were designed to be similar to resumes. Both profiles contained Workers' answers to the questions from Part One (i.e., their age, volunteer activity, former supervisor's impression of them, and upkeep with the news). The profiles also included their Partial Score, which was a noisy signal of their performance and equaled the number of questions they answered correctly from 5 randomly-chosen questions from the sports trivia quiz. The only difference between the profiles was that Profile A contained the Worker's gender and Profile B did not. An example of the profiles is shown in Figure 1.

Figure 1. Example of Worker Profile A and Profile B

[INSERT HERE]

Before Part Three began, the Manager's experiment was explained to Workers. They were told that MTurk participants would be recruited to participate in another study. These subjects, called "Managers," would start their experiment by answering the same set of sports trivia quiz questions as Workers did in Part Two. Managers would then be presented with one profile belonging to a male Worker and one profile belonging to a female Worker. Both Workers would have similar Partial Scores reported on their profiles (i.e., within one point of each other), and either both profiles would be Profile A's or both would be Profile B's.

The Managers' task was to hire one Worker, i.e. select one of the two profiles they were presented with. The Worker whose profile was chosen would be paid $0.40 for each correct sports trivia quiz answer they had submitted, and the other Worker would be paid $0.10 per correct answer. The Managers would earn $0.20 for each question answered correctly by the Worker they chose, incentivizing them to select the Worker they believed performed better on the quiz.[6]

[5] In a pilot study (n=51), 84 percent of men and 74 percent of women expected that men would answer more questions correctly on the sports trivia quiz than women. Furthermore, 75 percent of men and 70 percent of women believed that others would say that men performed better than women on the sports trivia quiz. Therefore, the sports trivia quiz was considered to be stereotypically-male. The quiz questions are available upon request.

[6] To make sure that Workers understood the Manager experiment, they completed a short comprehension quiz. These questions can be found in the Appendix.

Workers were told that they would be sorted into 1 of 3 treatment groups: Forced to Show, Forced to Hide, and Optional.[7] Depending on the treatment Workers were sorted into, some Workers had the ability to determine which version of the profiles was shown to Managers. In the Forced to Show group, Workers were required to use Profile A, which included their gender. Workers in the Forced to Hide treatment group were required to use Profile B, which excluded their gender. Lastly, subjects in the Optional group had the autonomy to determine which profiles would be shown to Managers based on their own preferences. Note that if subjects from the Optional group indicated a preference for Profile A (Profile B), then their Managers would see Profile A (Profile B) for both that subject and the subject they were matched with.

Because assignment to treatment occurred after the experiment was over, every subject was asked to indicate their preferences for Profile A and Profile B in case they were chosen for the Optional group. In Part Three, Workers were asked to consider two scenarios. In the first scenario, the default profile was Profile A, meaning that Managers would know the gender of both Workers. Workers were asked if they would be willing to pay various prices to have Managers see version B of the profiles instead. The prices ranged from \$0.00 to \$1.00 and increased in increments of \$0.10. In the second scenario, the default profile was Profile B, meaning that Managers would know that the two profiles belonged to individuals of opposite gender. However, they would not know which profile belonged to the female or the male. Workers were asked if they would be willing to pay \$0, \$0.10, …, \$1.00 to have Managers see Profile A instead.

After Workers indicated their preferences under each of these scenarios, they were told that one of their decisions would be chosen at random. If they were selected to be in the Optional group, then their choice in that randomly-chosen decision would be executed. This procedure is called multiple price list and is a common method used to elicit subjects' true preferences and willingness to pay (WTP).[8]

One decision was selected at random after the experiment was completely over. This chosen decision was whether Workers were willing to pay \$0.50 to switch from Profile A to Profile B. The sorting and matching process for determining how Workers were matched with their rivals is explained in Figure 2. If a Worker from the Optional group said he (she) would be

[7] In the experiment, these treatments were called Group 1, Group 2, and Group 3, respectively.
[8] See Kahneman, Knetsch, and Thaler (1990) and Eckel and Grossman (2008) for other examples of studies that use multiple price lists.

willing to pay $0.50 to switch profiles, then $0.50 was subtracted from his (her) final earnings. He (she) was matched with a random female (male) Worker from the Forced to Hide group who had a similar Partial Score. Their Manager would then see Profile B for both Workers. In other words, their Managers would not know which profile belonged to the male Worker and which belonged to the female Worker.

Figure 2. Sorting Process

[INSERT HERE]

If a subject from the Optional Group had said he (she) would *not* be willing to pay $0.50 to switch profiles, then nothing was subtracted from his (her) final earnings. He (she) was matched with a random female (male) from the Forced to Show group, and their Manager would see Profile A for both. In this case, their Manager would clearly be able to identify which profile belonged to the male Worker and which profile belonged to the female Worker.

Approximately 45 percent of Workers were randomly sorted into the Forced to Show group, 45 percent were in the Forced to Hide group, and 10 percent were in the Optional group. After all Workers from the Optional group were paired with an appropriate Worker from the Forced to Show or Forced to Hide group, the rest of the pairing went as follows. The remaining Workers from the Forced to Show group were paired with a random Worker from the same group, and their Managers saw version A of their profiles. Similarly, all remaining Workers from the Forced to Hide group were paired with a random Worker from the Forced to Hide group, and their Managers saw version B of their profiles.

Before Workers made their decisions about their WTP to change the profiles, they were made aware that Managers would never know anything about Part Three.[9] This meant that Managers would not know the Workers' WTP to change the profiles nor would they know that more than one version of the profiles ever existed. As a result, Workers did not have to worry that the Managers might judge them based on their responses to Part Three.

Once Part Three concluded, Workers completed a post-experimental survey in Part Four. They provided information about their exact age, race, ethnicity, highest level of education, status as an undergraduate student, employment status, monthly salary, and marital status.

They also answered five questions about their beliefs. First, they were asked how many questions they thought their rival (i.e., the Worker with whom they were matched) answered

[9] Workers also completed a second comprehension quiz, which can be found in the Appendix.

correctly. Workers were reminded of how many questions they had solved correctly themselves and reminded that their rival's Partial Score was similar to theirs before they answered. In the second and third question, Workers estimated how many questions the average male and average female Worker answered correctly. Their answers could range from 0.0 to 10.0 in increments of 0.1.

In addition, Workers were asked two questions that measured whether or not Workers believed that Managers were biased. For the first question, Workers were told that I would randomly select two Managers who had seen version A of the profiles. Each Manager had to choose between one male Worker and one female Worker and would see gender on their profiles. There were three possibilities: both of the Managers could choose the male worker, one of the Managers could choose the male worker while the other chose the female worker, or both of the Managers could choose the female worker. Workers guessed which of these three possibilities happened without seeing the profiles of the four selected Workers. This question captured what Workers expected to happen when Managers knew the gender of the Workers.

Similarly, Workers were asked to guess which of the three possibilities was most likely when two other randomly-chosen Managers had seen version B of the profiles. In this case, the Managers would not know the gender of the Workers, so this question captured what Workers expected to happen when the hiring process was gender-blind.

After the experiment was completely over, two pairs of Workers from the Forced to Hide group and two pairs from the Forced to Show group were randomly chosen, and their Managers' decisions were used to determine whether workers received additional earnings for their guesses. Workers were paid $0.10 for each of the five belief-based questions that they answered correctly.

The post-experimental survey ended with four questions about the Workers' beliefs about how much discrimination women face in the world and two questions about their beliefs about how much discrimination they face in Amazon MTurk, specifically. All questions from the post-experimental survey can be found in the Appendix.

The median time spent on the Worker experiment was 15.18 minutes. Workers earned a $2 completion fee and $0.50 for completing Part One. They also had the opportunity to earn bonus payments based on their performance on the sports trivia quiz in Part Two and their guesses in Part Four. On average, Workers who fully completed the experiment earned $4.12.

5.2 Experimental Design for Manager Experiment

For the Manager Experiment, 218 MTurk users of any gender were invited to participate in the study. The Manager experiment consisted of three parts. In Part One, Managers were told that MTurk users in a previous study had been asked to complete a sports trivia quiz. Managers were asked to complete the same quiz themselves so that they could become familiar with the task that Workers had completed. They were given 2.5 minutes, just like the Workers, but were not paid for their performance, something they knew ahead of time.

In Part Two, Managers decided whom they wanted to hire. They were shown the profiles of two MTurk users called Worker 1 and Worker 2. If the profiles were version A, the Managers knew which profile belonged to the female and which profile belonged to the male. Otherwise, the Manager only knew that the profiles belonged to two Workers of the opposite sex.[10] In either case, Managers were asked to choose the Worker they wanted to earn $0.40 for each correct answer he/she submitted from the sports trivia quiz. To incentivize Managers to select the Worker with the higher performance, Managers were told that they would earn $0.20 for each correct answer submitted by the Worker they chose.

After Managers chose a Worker, they completed a post-experimental survey similar to the Workers'. They were asked to report their gender, age, race, ethnicity, highest level of education, status as an undergraduate student, employment status, monthly salary, and marital status. They also stated whether they identified as male, female, or another gender.

Managers answered three questions about their beliefs about the Workers they were assigned to, which can be found in the Appendix. They were asked how many questions they believed Worker 1 and Worker 2 answered correctly out of 10. They were also asked whether they thought Worker 1 was female or male. For Managers who saw version A of the profiles, this question checked to see how well they paid attention to and remembered the profiles. For Managers who saw version B of the profiles, this question indicated whether or not there was any information in the profiles that indirectly revealed the gender of the Worker. For each correct guess, the Manager earned a bonus payment of $0.10. Lastly, the Managers stated their beliefs about discrimination in the world and in MTurk, specifically.

[10] This was done so that Managers always had the same information about the gender composition of the Workers. There would never be a case where, for example, Managers saw version B of the profiles and believed that they were choosing between two men or two women. Every Manager knew that there was a possibility that they could hire a man and a possibility that they could hire a woman.

The median time Managers spent on the experiment was 8.13 minutes. Managers earned $2 for completing the experiment. They had the opportunity to increase their earnings depending on the sports trivia quiz score of the Worker they selected and the accuracy of their beliefs. On average, Managers earned $3.32.

6. Models for the Experiments

6.1 Model for the Worker Experiment

The framework described in section 3 can be used to model the Workers' behavior in this experiment. In the experiment, Workers $i \in \{1,2\}$applied to a job and chose $G_i \in \{M,F,N\}$ such that they are maximizing $v(\boldsymbol{E}_i, G_i) = u(1, \boldsymbol{E}_i, G_i)$ from (1). In other words, Workers have already decided to apply for the job. If Managers do not know the gender of the Workers, then $0 \leq p(\boldsymbol{E}_i, N) \leq 1$ and $p(\boldsymbol{E}_i, N) = \theta p(\boldsymbol{E}_i, F) + (1-\theta)p(\boldsymbol{E}_i, M)$ where $\theta \in (0,1)$. The experiment defines $H \equiv 0.4$ and $L \equiv 0.1$. In addition, $\boldsymbol{E}_i \equiv \{\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{x}}_{-i}, G_{-i}\}$ where $\tilde{\boldsymbol{x}}_i$ includes all of the information on version B of the Worker i's profile and $\tilde{\boldsymbol{x}}_{-i}$ includes all of the information from version B of the other Worker's profile. $x_i \subset \tilde{\boldsymbol{x}}_i$ represents Worker i's score on the sports trivia quiz.

Assume that the Worker's true gender is female. Without loss of generality, let $c \equiv 0$ from (1). Then, her expected utilities from the task can be written as follows:

(2) $E\big(v(\boldsymbol{E}_i, F)\big) = p(\boldsymbol{E}_i, F) * 0.4 * x_i + \big(1 - p(\boldsymbol{E}_i, F)\big) * 0.1 * x_i = 0.1x_i(3p(\boldsymbol{E}_i, F) + 1)$

(3) $E\big(v(\boldsymbol{E}_i, N)\big) = p(\boldsymbol{E}_i, N) * 0.4 * x_i + \big(1 - p(\boldsymbol{E}_i, N)\big) * 0.1 * x_i = 0.1x_i(3p(\boldsymbol{E}_i, N) + 1)$

Workers were first asked how much they would be willing to pay to hide their gender. The maximum amount they should be willing to pay to use version B of the profiles, $MWTP_B$, should be $MWTP_B = \max\{0, E(v(\boldsymbol{E}_i, N) - E(v(\boldsymbol{E}_i, F))\}$. If Worker i does not believe that Managers will discriminate, then $\forall \tilde{\boldsymbol{E}}_i\ \tilde{p}\big(\tilde{\boldsymbol{E}}_i, F\big) = \tilde{p}\big(\tilde{\boldsymbol{E}}_i, N\big)$ and $MWTP_B = 0$. If the Worker believes that Managers will discriminate against women, then $\forall \tilde{\boldsymbol{E}}_i\ \tilde{p}\big(\tilde{\boldsymbol{E}}_i, F\big) < \tilde{p}\big(\tilde{\boldsymbol{E}}_i, N\big)$ and $MWTP_B \geq 0$. Conditional on the Worker having a positive quiz score (i.e., $x_i > 0$), then $MWTP_B > 0$. If the Worker believes that Managers will discriminate against men, then $\forall \tilde{\boldsymbol{E}}_i\ \tilde{p}\big(\tilde{\boldsymbol{E}}_i, F\big) > \tilde{p}\big(\tilde{\boldsymbol{E}}_i, N\big)$ and $MWTP_B > 0$. In fact, the Worker would be willing to pay to reveal their gender.

In the next part of the experiment, workers were asked how much they would be willing to pay to make their gender known by using version A of the profiles, $MWTP_A$. This can be

written as $MWTP_A = \max\{0, E(v(\boldsymbol{E}_i, F) - v(\boldsymbol{E}_i, N)\}$. If the Worker does not believe that Managers discriminate, then $MWTP_A = 0$ because $\forall \widetilde{\boldsymbol{E}}_i\ \tilde{p}(\widetilde{\boldsymbol{E}}_i, F) = \tilde{p}(\widetilde{\boldsymbol{E}}_i, N)$. If the Worker believes that Managers are biased against women, then $MWTP_A = 0$ as well because $\forall \widetilde{\boldsymbol{E}}_i\ \tilde{p}(\widetilde{\boldsymbol{E}}_i, F) = \tilde{p}(\widetilde{\boldsymbol{E}}_i, N)$. If the Worker believes that Managers are biased against men, then $MWTP_A \geq 0$ because $\forall \widetilde{\boldsymbol{E}}_i\ \tilde{p}(\widetilde{\boldsymbol{E}}_i, F) > \tilde{p}(\widetilde{\boldsymbol{E}}_i, N)$. When $x_i > 0$, $MWTP_A > 0$.

Therefore, by studying Workers' choices in Part 3, I can use their reported MWTP as an indicator of whether they believed discrimination existed and whether they could correctly identify which gender would be the target of discrimination (if it existed). The magnitude of their MWTP also reveals the extent of discrimination they believe existed. If Workers made optimal choices, then their reported MWTP should be equivalent to the difference between their earnings if their gender was revealed and if their gender was hidden.

6.2 Model for the Manager Experiment

Each manager *m* chose which Worker, $i \in \{1,2\}$, she wanted to earn \$0.40 per correct answer (i.e., who they wanted to hire) and was paid based on the performance of whichever Worker she chose. Managers may have found hiring Worker *i* distasteful, and this disutility is denoted as $d(\widetilde{\boldsymbol{x}}_i, G_i)$. Their utility function U_m can then be written as follows:

(4) $U_m(i) = .2x_i - d(\widetilde{\boldsymbol{x}}_i, G_i)$

Managers did not observe x_i; they only knew $\widetilde{\boldsymbol{x}}_i$ and G_i. Managers had a prediction about each worker's score on the quiz, which will be denoted as $\hat{x}_i(\widetilde{\boldsymbol{x}}_i, G_i)$. For what follows, assume that $\forall G_1$ and $\forall G_2, d(\widetilde{\boldsymbol{x}}_1, G_1) = d(\widetilde{\boldsymbol{x}}_2, G_2)$. Consequently, there are no taste-based preferences to work with Workers of one gender over another. To maximize their utility, Managers should pick Worker 1 if $E(\hat{x}_1(\widetilde{\boldsymbol{x}}_1, G_1)) > E(\hat{x}_2(\widetilde{\boldsymbol{x}}_2, G_2))$. They should pick the female Worker 2 if $E(\hat{x}_2(\widetilde{\boldsymbol{x}}_2, G_2)) > E(\hat{x}_1(\widetilde{\boldsymbol{x}}_1, G_1))$. If $E(\hat{x}_1(\widetilde{\boldsymbol{x}}_1, G_1)) = E(\hat{x}_2(\widetilde{\boldsymbol{x}}_2, G_2))$, then Managers should be indifferent between the two Workers.

As will be discussed in a later section, male Workers did perform better on the sports trivia quiz, on average, than female Workers. Therefore, if the Profiles only contained information about the Workers' gender, then it would be economically rational for Managers to choose men more than women when gender was known and to be more likely to hire men when gender was revealed versus hidden. This would be a case for statistical discrimination (see Arrow, 1973).

Managers received information about Workers' Partial Scores, but they may still feel like gender provides useful information about Workers' Total Scores. If that is the case, then $E(\hat{x}_i(\tilde{\boldsymbol{x}}_i, F)) \neq E(\hat{x}_i(\tilde{\boldsymbol{x}}_i, N))$ or $E(\hat{x}_i(\tilde{\boldsymbol{x}}_i, M)) \neq E(\hat{x}_i(\tilde{\boldsymbol{x}}_i, N)$. In other words, Managers' expectation of a Workers' quiz score when the Workers' gender is revealed is different from their expectation when the Workers' gender is hidden.

It is also possible that there was taste-based discrimination in the experiment (see Becker, 1957), i.e. $d(\tilde{\boldsymbol{x}}_1, G_1) \neq d(\tilde{\boldsymbol{x}}_2, G_2)$. If the Manager did not want to hire a female and $G_1 = F$, then (4) implies that the Manager would hire the male Worker even if $E(\hat{x}_1(\tilde{\boldsymbol{x}}_1, F)) = E(\hat{x}_2(\tilde{\boldsymbol{x}}_2, M))$. Depending on the difference between level of distaste Managers have for females relative to males, Managers could chose to hire males even if $E(\hat{x}_1(\tilde{\boldsymbol{x}}_1, F)) > E(\hat{x}_2(\tilde{\boldsymbol{x}}_2, M))$. Note, however, that Managers can only act on their preferences for men when gender is known.

In the context of this study, I state that there is evidence of discrimination if the rate in which Managers hired females when gender was known, $\hat{p}(\tilde{\boldsymbol{E}}_i, F)$, was different from when gender was unknown, $\hat{p}(\tilde{\boldsymbol{E}}_i, N)$. This could be because of statistical discrimination and/or taste-based preferences. Either way, this would indicate that Managers were using gender to influence their hiring decision. If they were not using gender, then the Managers should be just as likely to hire the female when gender was revealed as when gender was hidden, i.e. $\hat{p}(\tilde{\boldsymbol{E}}_i, F)$=$\hat{p}(\tilde{\boldsymbol{E}}_i, N)$.

7. Results[11]

7.1 Worker Beliefs About Performance and Actual Performance

At the end of the Worker experiment, Workers were asked to guess how well they believed the average female and male Worker in the experiment performed on the sports trivia quiz. Workers guessed that females answered 4.2 questions correctly, which is significantly less than the 5.6 questions they expected men to answer correctly ($p<0.01$; two-tailed t-test). Interestingly, women were more skeptical about women's abilities than men were. Women (n=215) expected women's average quiz score to be 4.0 compared to men (n=213) who predicted women's average quiz score to be 4.4 ($p=0.05$; two-tailed t-test). There were not any significant

[11] The final sample of subjects that were used for the analysis below consist of individuals from MTurk who completed either of the experiments and properly submitted the completion code they were given at the end of the experiment. Two subjects from the Worker experiment identified as female on MTurk but as male on the survey. Therefore, the data of these Workers were excluded from the analysis related to the Workers. When analyzing the decisions made by the Manager, I excluded these Workers and the Workers and Managers they were paired with. Additionally, if a pair of Worker profiles was viewed by more than one Manager, only the responses from the Manager who completed the experiment first were used.

gender differences regarding men's expected scores (p=0.28; two-tailed t-test). In short, both men and women assumed that men would perform better on the task, on average, than women. This confirms that the task was viewed as stereotypically male.

Workers had information about the range of their rival's possible Partial Scores, so I also compared Workers' expectations of their rivals. Men guessed that their female rivals answered 5.4 questions correctly whereas women guessed that their male partners answered 5.7 questions correctly. Thus, Workers expected a larger gender gap between the average male and the average female versus between them and their own rival. Bu men were still predicted to perform better than women (p=0.09; two-tailed t-test).

Workers' belief that men outperformed women was correct. On average, men answered 6.3 questions correctly compared to the 5.4 questions women answered correctly (p<0.01; two-tailed test). Note that these values are higher than what the Workers predicted and that Workers predicted that the gender gap would be 1.4 instead of 0.9. The distributions of the estimated and true performance by gender are shown in Figures 3 through 6.

Figures 3-6. Distributions of Estimated and True Worker Performance

[INSERT HERE]

7.2 Managers' Hiring Decisions

Managers chose which Worker they wanted to earn $0.40 per correct answer (i.e., who they wanted to hire). In the context of this experiment, I define discrimination as occurring if the frequency in which females were hired is different when Managers saw gender than when they did not see gender. When gender was revealed, Managers hired females 34.2 percent of the time. When gender was not revealed, women were hired 39.4 percent of the time. This difference is not statistically significant (p=0.44; two-tailed t-test). While this appears to suggest that there is no evidence that Managers discriminated against female Workers, it does not account for the fact that Managers saw different combinations of profiles: profiles where the Partial Scores of the two Workers were equal, profiles where the male Worker's Partial Score was higher than the female's, and profiles where the female Worker's Partial Score was higher than the male's.

When Managers were presented with the profiles of two Workers with the same Partial Score, they hired the female Worker 31.8 percent of the time when gender was known. This increased significantly to 56.0 percent when gender was not known (p=.05; Mann-Whitney test).

In other words, Managers were far more likely to select the female Worker when they did not know the gender of the person they were hiring. This can be seen in Figure 7.

When Managers knew that the male Worker's Partial Score was higher than the female Worker's Partial Score, Managers hired the female 23.6 percent of the time. The likelihood of a female being hired fell to 20.3 percent when Managers did not know the gender of the Workers. However, this difference is not statistically significant (p=0.67; Mann-Whitney test). In these cases, Managers had the tendency to hire the Worker with the higher Partial Score; gender did not seem to affect their decisions.

In some cases, Managers saw a profile belonging to a female Worker who had a higher Partial Score than the male Worker. When gender was revealed, the female Worker was hired 70.0 percent of the time compared to 86.7 percent of the time when gender was hidden. While this appears to be an economically significant difference, there were only 35 cases in total that fall into this category. Therefore, there are not enough observations to say that the difference is statistically significant (p=0.25; Mann-Whitney test). In sum, there appear to be circumstances when Managers did discriminate. They seemed to pay particular attention to the Workers' gender when they received signals that the Workers' ability was equal. Otherwise, they tended to hire the person whose Partial Score indicated that they were the better performer.

Figure 7. Percent of Female Workers Hired
[INSER HERE]

Table 2 reports the estimates of linear probability models where the dependent variable is a dummy variable that equals one if the female Worker was hired. The independent variables include a dummy variable for *Gender Shown*, which equals one if the Manager saw version A of the Profiles. If the coefficient for *Gender Shown* is significant, this means that Managers hired females at different rates depending on whether or not gender was known. The model also includes controls for the every item of the Workers' profiles, except for gender. For example, *Male Volunteers* equals one if the male Worker reported that he volunteered in the past year.

Column 1 of Table 2 reports the estimates for every pairing included in the final sample. Overall, women were 4.5 percentage points less likely to be hired when Profile A was used versus Profile B. This is not statistically significant, suggesting that there was no strong case for discrimination in aggregate. Women's age did play a role in Managers' hiring decisions. Females had a better chance of being hired if they were younger. In addition, their probability of being hired increased by 32.7 percentage points if their former supervisor considered them to be a good

employee. Interestingly, women were 18.1 percentage points less likely to be hired if they stated that they frequently read the news.

Column 2 of Table 2 displays the estimates of the same model, except only for pairings where the Partial Scores of the female and male Workers were equivalent. Females were 15.2 percentage points less likely to be hired when Managers knew gender. While this estimate is not statistically significant, it is economically significant. As discussed earlier, women whose Partial Scores were identical to their male rivals seemed to be put at a disadvantage when their gender was displayed. However, the only statistically significant factor that affected hiring decisions was the age of the female Worker.

Column 3 of Table 2 reports the estimates of the model when the only pairings under consideration are those where the male's Partial Score was higher than the female's. Women were 5.8 percentage points less likely to be hired when gender was displayed on the profiles, but this was not statistically significant. In these situations, the male Worker's age, whether or not the Worker was considered to be a good employee, and the female Worker's habit of reading the news all significantly affected the chances that the female was hired.

Lastly, column 4 of Table 2 considers pairings where the female Worker's Partial Score was higher than the male's. None of the variables had a statistically significant effect on the likelihood of the female being hired, but there were only 35 pairs in this subsample. It should be noted, however, that females were 12.3 percentage points less likely to be hired, which is economically significant.

To summarize, men were hired more, but Managers did not appear to discriminate when one examines all of their hiring decisions collectively. Yet there are indications, though not of statistical significance, that Managers may use gender to influence their hiring decision when they were asked to choose between two Workers with the same Partial Score or two Workers where the female's Partial Score was higher. Gender seemed to play no role when the Partial Score of the male Worker was higher than the female's.

Table 2. Linear Probability Model Estimates for the Likelihood of Females Being Hired
[INSERT HERE]

7.3 Workers' Beliefs About Discrimination

While the results above found mixed evidence of discrimination, there was clear evidence that many Workers anticipated that Mangers would use gender in their hiring decisions. There

were two ways in which I captured this information. The first was by eliciting Workers' willingness to pay to either include or exclude gender from the profiles using an incentive-compatible elicitation strategy. Except for when it was free to switch profiles, changing one's profile was a costly decision. Therefore, those who indicated a willingness to pay to modify their profiles must be those who believed that there was a benefit from doing so. To determine whether a self-regarding, utility-maximizing Worker should pay to change the profiles, he/she should compare the cost of switching profiles to the expected change in the likelihood of being hired and the resulting change in their earnings. Workers should be willing to make the change so long as the cost to do so is less than the additional earnings they would make from being hired.

Because the task was stereotypically male, I predicted that both men and women would believe that Managers would be biased against women, even when Managers were given a noisy signal of ability. As a result, I expected women to be willing to pay more for Profile B than Profile A because they would want the hiring process to be gender-blind. Men, on the other hand, should expect their chances of being hired to be greater when their gender was revealed to Managers. For this reason, I predicted that male Workers would be willing to pay more for Profile B than Profile A.

When it was free to switch to the gender-free profiles, 82.3 percent of the female Workers were interested compared to 60.1 percent of the male Workers. This difference is significantly different ($p<0.01$; two-tailed t-test). When the price to change the profiles increased to \$0.10, only 45.6 percent of females were willing to pay this amount. However, this percentage is still significantly higher than the 34.3 percent of male Workers who were willing to pay ($p=0.02$; two-tailed t-test). Table 3 displays the percentage of women and men who were willing to pay \$0.00, \$0.10, \$0.20, …, \$1.00 to use version B of the profiles. Women expressed significantly more interest than men when the cost of changing was between \$0.00 and \$0.30 ($p\leq0.10$; two-tailed t-tests). There is no significant difference between the percent of women and men willing to pay prices (\$P) above \$0.30. All in all, this implies that women suspected that their gender may be disadvantageous to them. However, most women were not willing to pay more than \$0.30 to remove gender from the profiles.

Table 3. Proportion of Workers Willing to Pay to Switch to Version B of the Profiles

[INSERT HERE]

When switching from version B to version A of the profiles was free, 55.9 percent of men were interested in that option. Women found it less appealing, with only 30.7 percent of them expressing a preference for Profile A ($p<0.01$; two-tailed t-test). Once switching became a costly decision, male interest dropped to 32.9 percent and female interest dropped to 13.0 percent. This difference is statistically significant ($p<0.01$; two-tailed t-test). In fact, as can be seen in Table 4, men were always more willing to pay to use version A of the profiles than women ($p<0.01$; two-tailed t-tests). This suggests that men anticipated being favored by Managers when Managers were aware of their gender.

Table 4. Proportion of Workers Willing to Pay to Switch to Version A of the Profiles

[INSERT HERE]

I defined each Worker's maximum willingness to pay (MWTP) to switch profiles as the highest price, $\$\overline{P}$, that he/she was willing to pay such that he/she was unwilling to pay any higher price.[12] Workers were recorded as having a MWTP of $0 if they said they would never want to switch or if they would pay $0.00 to switch Profiles but nothing higher. Figure 8 displays the average MWTP for male and female Workers to switch from one version of the profile to another. Men's MWTP to use Profile B was $0.13, and they were willing to forfeit 7.5 percent of their earnings from being hired in order to use profiles without gender. Their MWTP to use Profile A was $0.18, and they were willing to forfeit 9.4 percent of their earnings from being hired. The difference in MWTP is significantly different ($p=0.04$; two-tailed t-test) and suggests that men expected to benefit more when Managers knew the gender of the Workers.

Figure 8. Average MWTP by Gender and Profile Version

[INSERT HERE]

Women's MWTP to use Profile B was $0.15, and they were willing to give up 10.0 percent of their earnings from being hired just to gender removed. When it came to their interest in including gender, women's MWTP to use Profile A was only $0.05, and they were willing to spend 4.3 percent of their earnings from being hired to reveal their gender. The fact that they

[12] Many researchers who use multiple price lists are concerned about how to interpret the action of subjects who indicate a lack of interest in paying $P but a willingness to pay $P+$x ($x>0). (See Andersen et al., 2006 for a discussion on the advantages and disadvantages of multiple price lists.) When this occurred in this experiment, Workers were asked to review their answers but were not forced to make any changes. Therefore, the final answers that Workers submitted were considered as reflecting their true preferences and not as irrational behavior. If $P+$x was the highest price that Workers were willing to pay, then this value was recorded as their MWTP regardless of their WTP (or lack thereof) for prices less than $P+$x.

were willing to pay significantly more to use profiles without gender ($p<0.01$; two-tailed t-test) implies that they had a strong interest in making sure that Managers did not know their gender.

It is also of interest to compare Workers' MWTP for men versus women. While women expressed a higher interest in using Profile B when switching profiles was free, they were reluctant to pay high amounts for that opportunity. Consequently, men and women had a similar MWTP to use Profile B: \$0.13 and \$0.15, respectively ($p=0.44$; two-tailed t-test). This was not true for MWTP to use version A of the profiles. Men had a greater interest than women in using version A of the profiles even when switching was costly, and this translated to a higher average MWTP to use profiles that included gender for men than for women ($p<0.01$; two-tailed t-test).

In sum, men appeared to believe that they would have an advantage if Managers could identify their profile as the one belonging to the male Worker. This is in contrast to women, who believed that revealing their gender would have negative consequences. This suggests that both men and women expected women to become targets of discrimination for this stereotypically male task.

The experiment included a second measure of anticipated discrimination. At the end of the experiment, Workers were asked what they expected to happen in randomly chosen cases where two Managers were presented with version A of the profiles. 39.1 percent of women believed that both of the Managers would pick the male Worker. 55.4 percent believed that one Manager would pick the male Worker and one Manager would pick the female Worker. 5.6 percent believed that both of the Managers would pick the female Worker. Thus, over half of the women expected female Workers to be just as appealing to Managers as male Workers when gender was revealed, but there were still women who believed that men would be favored.

Men had similar beliefs. 34.3 percent of them believed that the male Worker would be selected both times, 55.9 percent believed that the male Worker would be selected half of the time, and 9.9 percent believed that the male Worker wouldn't be selected at all. In sum, most male Workers believed that men and women had an equal chance of being hired, but roughly 1 in 3 believed that Managers would prefer men when gender was revealed. Workers were asked to provide the rationale for their guesses, and a selected sample of their responses can be found in the Appendix. Mainly, Workers expected men to be chosen when gender was revealed because they believed that men knew more about sports than women.

Workers also made predictions about what would happen in cases where two randomly chosen Managers did not know the Workers' gender. In this case, 81.4 percent of women and 81.2 percent of men expected one male and one female Worker to be chosen. Workers' explanations for their decision suggest that they expected Managers' choices to be random. However, there were still some men and women who thought men would always be selected (14.1 percent and 14.9 percent, respectively).

It is important to note that the percentage of Workers who believed that men would always be chosen more than doubled when Workers were asked about Managers who knew the Workers' gender compared to Managers who did not. These responses confirm that many Workers believed that female Workers would be at a disadvantage if Managers could observe gender, but male Workers would be at an advantage. Furthermore, this is consistent with the findings on Workers' MWTP that women preferred to conceal their gender and men preferred to reveal theirs.

7.4 Are Workers Overestimating the Cost (Benefit) of Their Gender?

Section 7.3 showed that both male and female Workers were unconvinced that Managers would be entirely impartial, and section 7.2 revealed that Managers did not always discriminate. In this section, I estimate how much financial advantage or disadvantage Workers received because of their gender and compare this to their MWTP to have gender included or excluded. This allows me to determine whether or not Workers were optimally able to set their MWTP.

Table 5 contains the results of several linear regressions where the dependent variable is *Task Earnings*, which represents the amount of money Workers earned from the sports trivia quiz. This amount depended on their score on the quiz and the hiring decision of the Manager. The regressions also include a dummy variable for all of the items of the Worker's profile. For example, *Reads News* equals one if the Worker indicated that he/she reads the news often. There is also a variable to capture the difference in the Workers' Partial Scores.

Column 1 of Table 5 is a simple regression to see if females earned less than men. Column 2 of Table 5 adds a dummy variable that equals one if Profile A was shown to the Manager, and the third column of Table 5 adds an interaction term between this dummy variable and a dummy variable for whether or not the Worker was female. If the coefficient on this interaction term is negative and statistically significant, it would suggest that females earned more when they hid their gender from the Managers.

Table 5. OLS Estimates of the Effect of Gender on Worker Earnings

[INSERT HERE]

As can be seen in each regression, women earned less than men, even when controlling for differences in age, volunteer history, supervisors' impressions, frequency of reading the news, and Partial Scores. The coefficient on *Female*Gender Shown* in Column 3 is negative but not statistically significant. However, the results suggest that women earned $0.10 less when their gender was shown versus hidden. Men, on the other hand, earned $0.02 more when their gender was shown versus hidden. So revealing gender had no financial advantage for men, and hiding gender had a small (but statistically insignificant) effect for women.

A self-regarding utility maximizer would set their MWTP such that the MWTP is equal to the financial advantage they would earn from using one version of the profiles over the other. While women's MWTP to switch to version A of the profiles was $0.05, women's MWTP to use version B was $0.15. Based on the coefficients from Table 5 and their statistical significance, both of these values should be $0.00 if women were responding optimally. By the same token, men should not have been willing to pay anything to change their profiles, yet their MWTP to use version A was $0.18. Again, these values should be $0.00 if men were responding optimally.

I find that both women's MWTP to switch to Profile B and men's MWTP to switch to Profile A are significantly different from $0.00 ($p<0.01$; one-sample t-tests). However, women's MWTP to use Profile B is close to the true difference in earnings for females when gender was shown versus hidden (i.e., the coefficient for *Gender Shown* plus the coefficient for *Female*Gender Shown*). On the other hand, men's MWTP for version A was significantly higher than $0.02 ($p=0.01$; one-sample t-test). In sum, I find that men, in particular, overestimated the cost of being female and the benefit of being male.

8. Conclusion

There is an under-representation of women in certain careers like engineers and computer programmers. This could be driven by gender differences in application rates or gender differences in hiring rates. However, an individual's decision to apply to a job may depend on their expected likelihood of being hired. This expectation may further rely on their beliefs about how gender biased hiring managers are. In this paper, I investigate whether anticipated discrimination can help explain why men are overrepresented in some fields. If men think they

will be favored or women think they will be targets of discrimination, then anticipated discrimination could explain why men opt in and women opt out of certain fields.

I conducted two experiments using Amazon Mechanical Turk to test whether men and women can accurately predict how much gender discrimination is present in this experimental labor market. Workers completed a stereotypically male task (a sports trivia quiz) and answered questions to create two resume-like profiles. One version of the profile included gender, and the other did not. Each Worker was matched with another Worker who had similar ability but was of the opposite sex. Workers were then asked how much they would be willing to pay to use one version of the profile over the other if Managers were required to hire only one of the two Workers. Workers' maximum willingness to pay (MWTP) to use the profiles with or without gender were indicators of how much they perceived their gender to be an asset or liability, respectively.

I find that, when changing versions of the profiles was costless, 82.3 percent of women were interested in switching to the profiles without gender relative to 60.1 percent of the men. When Workers were asked about switching to profiles with gender for free, 55.9 percent of men were interested compared to only 30.7 percent of women. When studying Workers' MWTP, I found that men's average MWTP to use profiles with gender was $0.18, which is significantly higher than the $0.13 they were willing to spend to use profiles without gender. Women's average MWTP, on the other hand, was $0.15 to use profiles without gender and $0.05 to use profiles with gender. This difference is statistically significant. All in all, these findings suggest that both men and women assumed that men would have an advantage in the hiring process when gender was revealed. This encouraged men to want to reveal their gender and women to want to conceal it.

Men were indeed hired more than women. Men were hired 65.8 percent of the time when gender was revealed and 60.6 percent of the time when gender was hidden. This difference is not statistically significant, so there was no strong evidence of discrimination when all of the pairs are considered collectively. However, there is suggestive evidence that Managers used gender in their hiring decision under certain circumstances like when the Partial Scores of the two Workers were equivalent. In that case, they were less likely to hire female Workers when gender was revealed than when gender was hidden.

Regardless, men (women) were not significantly better (worse) off financially by having the Managers know their gender. This means that male and female's MWTP to change the profiles should have been zero. Yet women's gender was perceived by both genders to be a disadvantage. Women were willing to forfeit 10.0 percent of their earnings from being hired to remove gender, and men were willing to forfeit 9.4 percent of their earnings to include gender. But men, in particular, seemed to overestimate how much of a role gender played on hiring decisions.

These findings suggest that perceptions about the gender bias of hiring managers matter and may affect men and women's decision to pursue certain careers. If men believe that they have an advantage in stereotypically male fields, they may be overconfident in how much their gender alone will get them ahead if they overestimate how biased hiring managers are. As a result, they may apply to more jobs of this type than what would be optimal if they had correct beliefs about managers' biases. The reverse could happen for women. If they believe that they are disadvantaged, they may apply to fewer jobs of this type than what is optimal because they have overestimated how biased hiring managers will be against them. Further research should investigate ways to help update job candidates' expectations of discrimination, which might affect the gender composition of jobs in stereotypically male fields like STEM.

REFERENCES

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics, 9*(4), 383-405.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics, 121*(1-2), 343-375.

Arrow, K. (1973). The theory of discrimination. *Discrimination in labor markets*, *3*(10), 3-33.

Becker, G. S. (1957). The economics of discrimination Chicago. *University of Chicago*.

Coffman, K. B., Exley, C. L., & Niederle, M. (2017). When Gender Discrimination Is Not About Gender. Harvard Business School Working Paper.

Eckel, C. C., & Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization, 68*(1), 1-17.

Goldin, C., & Rouse, C. (2000). Orchestrating impartiality: The impact of" blind" auditions on female musicians. *American Economic Review, 90*(4), 715-741.

Hoogendoorn, S., Oosterbeek, H., & Van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science, 59*(7), 1514-1528.

Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics, 2*(3), 2053168015604648.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy, 98*(6), 1325-1348.

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened resumes: Race and self-presentation in the labor market. *Administrative Science Quarterly, 61*(3), 469-502.

Kirkeboen, L. J., Leuven, E., & Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics, 131*(3), 1057-1111.

Krause, A., Rinne, U., & Zimmermann, K. F. (2012). Anonymous job applications in Europe. *IZA Journal of European Labor Studies*, *1*(1), 5.

Levay, K. E., Freese, J., & Druckman, J. N. (2016). The demographic and political composition of Mechanical Turk samples. *Sage Open, 6*(1), 2158244016636433.

Neumark, D., Bank, R. J., & Van Nort, K. D. (1996). Sex discrimination in restaurant hiring: An audit study. *The Quarterly Journal of Economics, 111*(3), 915-941.

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

Paglin, M., & Rufolo, A. M. (1990). Heterogeneous human capital, occupational choice, and male-female earnings differences. *Journal of Labor Economics, 8*(1, Part 1), 123-144.

Riach, P. A., & Rich, J. (2006). An experimental investigation of sexual discrimination in hiring in the English labor market. *Advances in Economic Analysis & Policy, 5*(2)

Rinne, U. (2018). Anonymous job applications and hiring discrimination. *IZA World of Labor*.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science (New York, N.Y.), 330*(6004), 686-688.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources, 48*(3), 545-595.

TABLES

Table 1. Subject Characteristics

	Workers	Managers
Average Age	35.92	35.16
	(10.93)	(10.19)
Percent Male	49.30%	59.15%
Percent White, Non-Hispanic	73.36%	73.71%
Mode Monthly Salary	$1,001 - $2,000	$2,001 - $3,000
Percent with Bachelor's Degree or Higher	57.48%	65.26%
Percent Employed Full-Time	64.02%	75.59%
Percent Employed Part-Time	16.82%	10.80%
Percent Married	39.25%	43.19%
Percent Never Married	5.84%	46.48%

Note: These values are for the final sample. Standard deviations are written in parentheses. Monthly income was reported in ranges of values, so the mode was reported.

Table 2. Linear Probability Model Estimates for the Likelihood of Females Being Hired

	(1) All	(2) Male=Female	(3) Male>Female	(4) Female>Male
Gender Displayed	-0.0453	-0.152	0.0582	-0.123
	(0.0671)	(0.135)	(0.0778)	(0.169)
Male's Age	0.0379	0.0264	0.0964*	0.00811
	(0.0466)	(0.0859)	(0.0553)	(0.116)
Female's Age	-0.0858*	-0.184**	0.0334	0.0711
	(0.0491)	(0.0906)	(0.0596)	(0.120)
Male Volunteers	0.0387	0.148	0.00306	0.0508
	(0.0676)	(0.134)	(0.0813)	(0.164)
Female Volunteers	0.0351	-0.00121	-0.0209	0.211
	(0.0680)	(0.146)	(0.0742)	(0.157)
Male Good Employee	-0.169	-0.0922	-0.428*	0
	(0.186)	(0.276)	(0.243)	(.)
Female Good Employee	0.327***	0.158	0.196***	0
	(0.0845)	(0.136)	(0.0744)	(.)
Male Reads News	0.109	-0.322	0.283*	-0.0128
	(0.141)	(0.193)	(0.155)	(0.269)
Female Reads News	-0.181**	-0.200	-0.139**	0.0846
	(0.0814)	(0.159)	(0.0652)	(0.237)
Constant	0.333	1.136**	-0.00743	0.183
	(0.318)	(0.454)	(0.356)	(0.588)
N	213	64	114	35

Note: Standard deviations are reported in parentheses. *p<0.01, **p<0.05, ***p<0.01

Table 3. Proportion of Workers Willing to Pay to Switch to Version B of the Profiles

Cost ($P) of Switching to Version B	Proportion of Workers Willing to Pay $P		Two-tailed t-test p-value
	Males	Females	
$0.00	0.601	0.823	<0.01
	(0.491)	(0.026)	
$0.10	0.343	0.456	0.02
	(0.476)	(0.499)	
$0.20	0.263	0.335	0.10
	(0.441)	(0.473)	
$0.30	0.160	0.223	0.09
	(0.367)	(0.417)	
$0.40	0.127	0.167	0.236
	(0.333)	(0.374)	
$0.50	0.103	0.112	0.78
	(0.305)	(0.316)	
$0.60	0.066	0.060	0.82
	(0.248)	(0.239)	
$0.70	0.056	0.047	0.65
	(0.231)	(0.221)	
$0.80	0.052	0.042	0.63
	(0.222)	(0.201)	
$0.90	0.038	0.037	0.98
	(0.191)	(0.190)	
$1.00	0.047	0.037	0.62
	(0.212)	(0.190)	

Note: Standard deviations are reported in parentheses.

Table 4. Proportion of Workers Willing to Pay to Switch to Version A of the Profiles

Cost ($P) of Switching to Version A	Proportion of Workers Willing to Pay $P		Two-tailed t-test p-value
	Males	Females	
$0.00	0.559	0.307	<0.01
	(0.498)	(0.462)	
$0.10	0.329	0.130	<0.01
	(0.471)	(0.337)	
$0.20	0.296	0.098	<0.01
	(0.457)	(0.298)	
$0.30	0.225	0.060	<0.01
	(0.419)	(0.239)	
$0.40	0.202	0.037	<0.01
	(0.402)	(0.190)	
$0.50	0.192	0.023	<0.01
	(0.395)	(0.151)	
$0.60	0.136	0.009	<0.01
	(0.344)	(0.096)	
$0.70	0.099	0.014	<0.01
	(0.299)	(0.118)	
$0.80	0.094	0.019	<0.01
	(0.292)	(0.135)	
$0.90	0.085	0.014	<0.01
	(0.279)	(0.118)	
$1.00	0.085	0.019	<0.01
	(0.279)	(0.136)	

Note: Standard deviations are reported in parentheses.

Table 5. OLS Estimates of the Effect of Gender on Worker Earnings in Dollars

	(1)	(2)	(3)
Female	-0.540***	-0.540***	-0.469***
	(0.120)	(0.120)	(0.174)
Gender Displayed		-0.0394	0.0243
		(0.106)	(0.163)
Female*Gender Displayed			-0.128
			(0.214)
Age	-0.197**	-0.197**	-0.200**
	(0.0817)	(0.0819)	(0.0823)
Volunteered in Past Year	-0.0871	-0.0855	-0.0851
	(0.108)	(0.108)	(0.109)
Good Former Employee	0.511*	0.509*	0.497*
	(0.293)	(0.291)	(0.295)
Reads News	0.393**	0.394**	0.392**
	(0.166)	(0.166)	(0.167)
Diff. in Partial Scores	0.0801***	0.0802***	0.0812***
	(0.0295)	(0.0295)	(0.0297)
Constant	1.668***	1.691***	1.677***
	(0.371)	(0.381)	(0.386)
N	426	426	426

Note: Standard deviations are reported in parentheses. *p<0.10, **p<0.05, and ***p<0.01.

FIGURES

Figure 1. Example of Worker Profile A and Profile B

Profile Version A:

Gender	Female
Age	26 - 34 years old
Volunteered in past year?	No
Considered good employee?	Yes
Reads the news often?	No
Partial Score (out of 5)	3

Profile Version B:

Age	26 - 34 years old
Volunteered in past year?	No
Considered good employee?	Yes
Reads the news often?	No
Partial Score (out of 5)	3

Figure 2. Sorting Process

	Preference to pay $0.50 for Profile B	**Matched With**	**What Manager Sees**	**Worker's Earnings from Quiz, w={0.10, 0.40}**
Female chosen for Optional	Wants gender *hidden*	Male chosen from Forced to Hide	Two profiles *without* gender	$w*correct answers - $0.50
Female chosen for Optional	Wants gender *shown*	Male chosen from Forced to Show	Two profiles *with* gender	$w*correct answers
Male chosen for Optional	Wants gender *hidden*	Female chosen from Forced to Hide	Two profiles *without* gender	$w*correct answers - $0.50
Male chosen for Optional	Wants gender *shown*	Female chosen from Forced to Show	Two profiles *with* gender	$w*correct answers

Figure 3. Distribution of Estimated Female Worker Performance

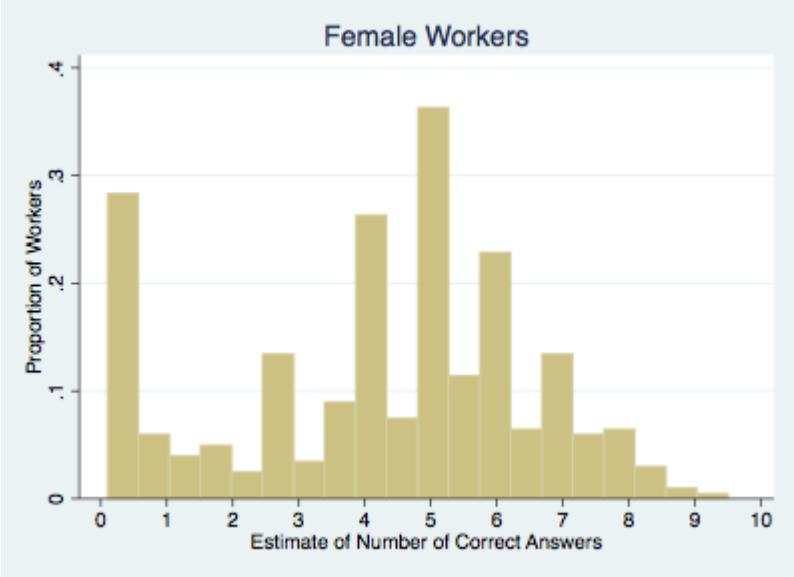


Figure 4. Distribution of Estimated Male Worker Performance

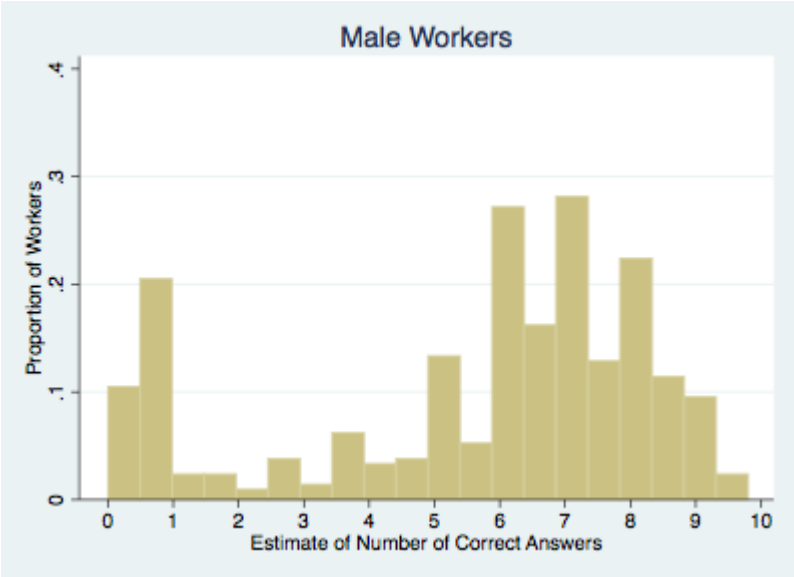


Figure 5. Distribution of True Female Worker Performance

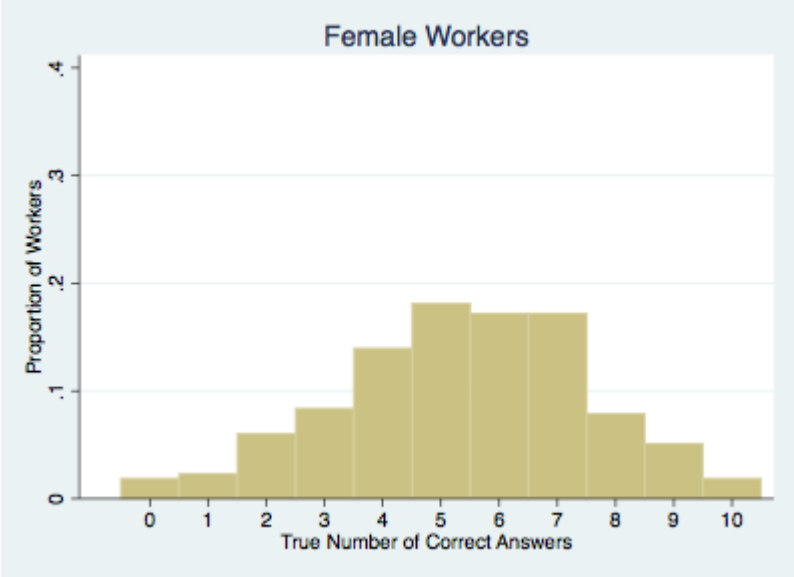


Figure 6. Distribution of True Male Worker Performance

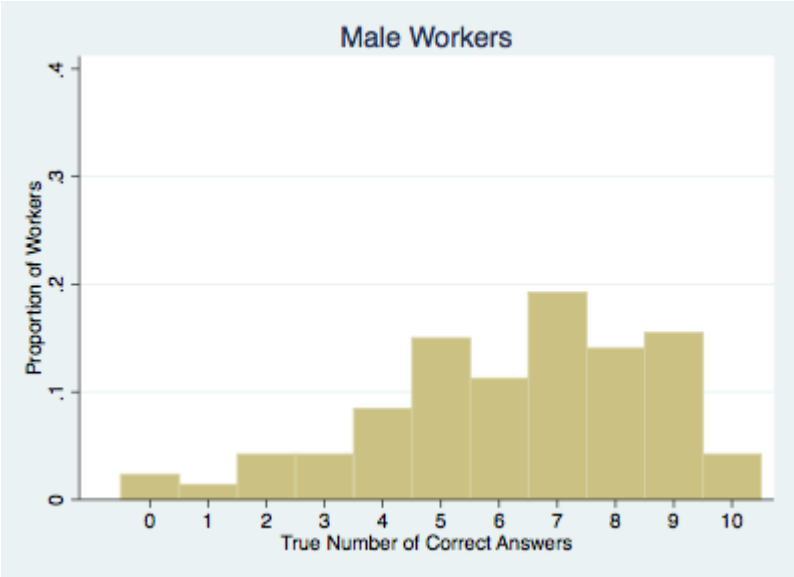


Figure 7. Percent of Female Workers Hired

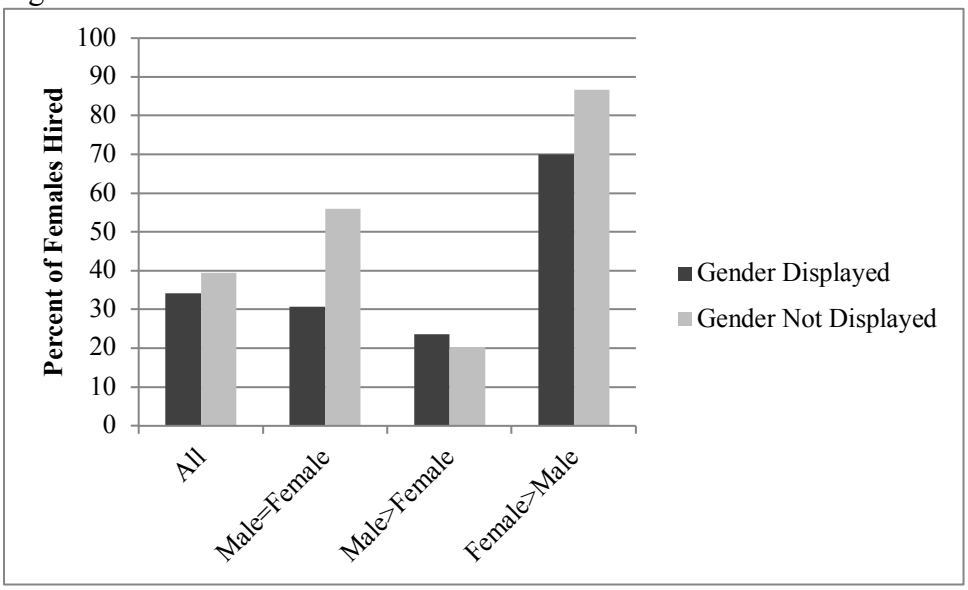
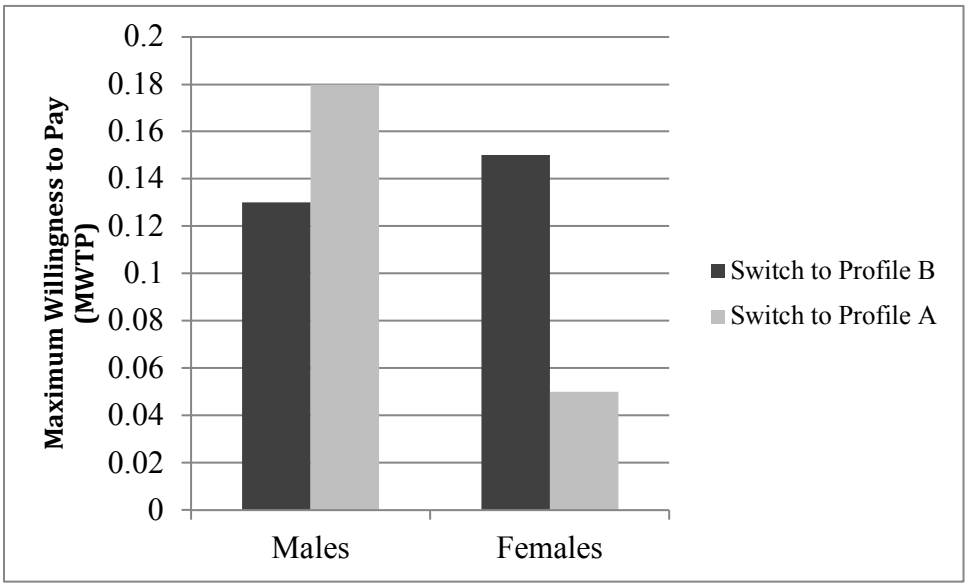


Figure 8. Average MWTP by Gender and Profile Version



APPENDIX: ADDITIONAL TABLES AND FIGURES

Table A.1 Worker Characteristics by Gender

	Males	Females	Two-tailed t-test p-value
Exact Age	34.35 (10.51)	37.47 (11.14)	p<0.01
Percent who volunteered in the past year	44.60%	52.0%	p=0.12
Percent whose supervisors would consider them a good employee	95.77%	98.60%	p=0.08
Percent who read the news often	93.43%	84.65%	p<0.01
Partial Score	3.13 (1.25)	2.76 (1.30)	p<0.01
N	213	215	

Note: Standard deviations are written in parentheses.

APPENDIX: SELECTED COMMENTS FROM WORKERS

Selected comments from male Workers who predicted that males would be chosen twice when version A profiles were used:

- I think people assume males stereotypically know more about sports. I think the managers will want to earn their bonus for correct answers and assume males are better at answering sports questions.
- I think the manager would feel the male would know more about sports.
- Think the manager might be biased

Selected comments from female Workers who predicted that males would be chosen twice when version A profiles were used:

- With money being on the line if they choose the person who got more right, I think that managers will go with the male workers based on the stereotype of men knowing more about sports.
- I think people will trust males more to answer sports questions correctly.
- men are stereotyped as more knowledgeable about sports

Selected comments from male Workers who predicted that males would be chosen once when version A profiles were used:

- because both will have a similar score
- I don't think gender is an issue here.
- to be as fair as possible

Selected comments from female Workers who predicted that males would be chosen once when version A profiles were used:

- I would hope it's kept fair and based just on scores.
- Hope that sexism isn't as bad as i think it is.
- Some men have faith in women's ability to know a little bit about sports.

Selected comments from male Workers who predicted that males would be chosen zero times when version A profiles were used:

- I think people are more sensitive to not being sexist now
- I think people tend to be more sympathetic towards women.
- I would pick females if I were them.

Selected comments from female Workers who predicted that males would be chosen zero when version A profiles were used:

- The questions are all sports related so I believe it is more fair to reward the female.
- this is just a random assumption that they picked 2 girls for this.
- Women are better

Selected comments from male Workers who predicted that males would be chosen twice when version B profiles were used:

- Since males are more likely to get a higher score, they would be picked due to their high score
- Because they will have a higher score.
- the males partial score is more likely to be higher than the females so the manager would probably pick that.

Selected comments from female Workers who predicted that males would be chosen twice when version B profiles were used:

- Because I think it's more likely for the males to have a slightly higher partial score.
- Their scores will be higher so they will choose them.
- I think the males probably did better on the questions so they have a higher partial score possibly

Selected comments from male Workers who predicted that males would be chosen once when version B profiles were used:

- It will be up to chance which gender is chosen.
- I think it's random chance if the manager does not have access to the gender info.
- Random since gender is removed

Selected comments from female Workers who predicted that males would be chosen once when version B profiles were used:

- Without the gender bias, there is a 50/50 chance on who they will choose. It all comes down to which worker had the most correct answers, and I'm willing to bet that my fellow female workers didn't do too shabby.
- It would probably be a 50-50 chance.
- I think with genders removed it will more likely end up a split because the managers won't see the genders.

Selected comments from male Workers who predicted that males would be chosen zero times when version B profiles were used:

- female get a more sympathetic ear and males are supposed to just accept what they are given.
- Females tend to volunteer to causes more than males. The managers will look at volunteering as more compassionate.
- The manager will probably choose those who donate to charity which I believe to be more likely women.

Selected comments from female Workers who predicted that males would be chosen zero when version B profiles were used:

- Just a toss up -- I feel lucky and that 2 females will win.
- Better scores will be picked.
- I trust in there judgments

APPENDIX: EXPERIMENTAL MATERIALS FROM WORKER EXPERIMENT

PART 1 INSTRUCTIONS

Please answer the questions below. Your responses will be used to create a Profile.

You will be given a bonus payment of $0.50 for completing Part 1.

What is your gender?

a. Male
b. Female
c. Other (Please explain.) _____

What is your age?

a. 18 – 25 years old
b. 26 – 34 years old
c. 35 years old or older

Have you volunteered in the past 12 months?

a. Yes
b. No

Would your former supervisor say that you were a good employee?

a. Yes
b. No
c. N/A (Never had a supervisor.)

Do you read the news (online or from a newspaper) often?

a. Yes
b. No

COMPREHENSION QUIZ #1

To make sure you understand what happens in Study #2, please answer the following questions.

1. True or False: The Manager is an adult MTurk worker located in the United States.

a. True ***Correct Answer
b. False

2. True or False: The Manager will see your Profile and the Profile of a [male/female] MTurk worker who participated in this study.

a. True ***Correct Answer
b. False

3. True or False: If the Manager decides that Worker 2 will be paid $0.40 for each correct answer from Part 2, then you will be paid $0.10 for each correct answer.

a. True ***Correct Answer
b. False

4. True or False: The Manager gets a bonus payment based on how many questions the Worker they chose answered correctly.
   a. True ***Correct Answer
   b. False

5. True or False: The Manager will know your Total Score and your Partial Score.
   a. True
   b. False ***Correct Answer

COMPREHENSION QUIZ #2

Please answer the questions below.

Remember, Version A Profiles include:
- Gender
- Age
- Volunteered in past year?
- Considered good employee?
- Reads the news often?
- Partial Score (out of 5)

And Version B Profiles include:
- Age
- Volunteered in past year?
- Considered good employee?
- Reads the news often?
- Partial Score (out of 5)

1. True or False: The Manager knows that there are 2 versions of the Profiles (Version A and Version B).
   a. True
   b. False ***Correct Answer

2. True or False: The Manager will know how much money you were willing to pay to avoid using the Version A Profiles (gender included).
   a. True
   b. False ***Correct Answer

3. True or False: It's possible for the Manager to see your Version A Profile (gender included) and Worker 2's Version B Profile (gender removed).
   a. True
   b. False ***Correct Answer

4. True or False: If the Manager sees the Version B Profiles (gender removed), then they will know with certainty that your Profile is the one that belongs to a female.
   a. True
   b. False ***Correct Answer

POST-EXPERIMENTAL WORKER SURVEY (PART FOUR)

PART 4 INSTRUCTIONS
Please follow the instructions for the questions below.

Please answer the 5 questions below to the best of your ability. You will earn $0.10 for each correct guess.

1. How many questions do you think Worker 2 answered correctly?
Remember: You answered [NUMBER] questions correctly, and you will be partnered with a [male/female] from this study who had a similar Partial Score as you.

2. We will randomly pick 2 pairs of Workers whose Version A Profiles (gender included) will be shown to the Managers. This will give us 2 Managers, 2 male Workers, and 2 female Workers. Each Manager will have to choose between the male and female Worker and will know their gender.

Which Workers do you think will be chosen to be paid $0.40? [***order randomized***]

a. 2 male Workers
b. 2 female Workers
c. 1 male Worker and 1 female Worker

Why did you choose this option?

3. We will randomly pick 2 pairs of Workers whose Version B Profiles (gender removed) will be shown to the Managers. This will give us 2 Managers, 2 male Workers, and 2 female Workers. Each Manager will have to choose between the male and female Worker but won't know which Worker is male and which one is female.

Which Workers do you think will be chosen to be paid $0.40? [***order randomized***]

d. 2 male Workers
e. 2 female Workers
f. 1 male Worker and 1 female Worker

Why did you choose this option?

4. How many questions do you think male Workers in this study answered correctly on average?

5. How many questions do you think female Workers in this study answered correctly on average?

Please tell us more about yourself.

1. What is your age?

2. How do you identify? (Check all that apply.)

- ❑ White
- ❑ Hispanic, Latino, or Spanish origin
- ❑ Black or African American
- ❑ Asian (ex. Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, etc.)
- ❑ American Indian or Alaska Native
- ❑ Middle Eastern or North African (ex. Lebanese, Iranian, Egyptian, Syrian, Moroccan, Algerian, etc.)
- ❑ Native Hawaiian or Other Pacific Islander
- ❑ Some other race, ethnicity, or origin (Please explain.) _____

3. What is the highest level of school you have completed or the highest degree you have received?

a. Some high school
b. High school graduate or other equivalent (high school diploma, GED, etc.)
c. Some college but no degree
d. Bachelor's degree (ex. BA, AB, BS)
e. Master's degree (ex. MA, MS, MBA)
f. Associate's degree
g. Professional school degree (ex. MD or JD)
h. Doctorate degree (ex. PhD)
i. Other (Please explain.) _____

4. Are you currently an undergraduate student?

a. Yes
b. No

5. What is your current employment status (excluding your MTurk activity)?

a. Unemployed
b. Employed part-time
c. Employed full time
d. Other (Please explain.) _____

6. What is your monthly salary? (If you can't remember or don't know the exact amount, please give your best guess.)
[***dropdown list***]

7. What is your current marital status?[13]

a. Now married
b. Widowed
c. Divorced
d. Separated
e. Never married

How strongly do you agree with the statements below?

[13] The wording for this question came from the U.S. Census.

1. On average, women are more educated than men.[14]
   a. Strongly Disagree
   b. Disagree
   c. Neither Disagree or Agree
   d. Agree
   e. Strongly Agree

2. On average, men face more discrimination than women when applying for jobs.
   a. Strongly Disagree
   b. Disagree
   c. Neither Disagree or Agree
   d. Agree
   e. Strongly Agree

3. On average, women are more likely to be paid less than men for doing the same job.
   a. Strongly Disagree
   b. Disagree
   c. Neither Disagree or Agree
   d. Agree
   e. Strongly Agree

4. Discrimination in the workplace has decreased over the past 10 years.
   a. Strongly Disagree
   b. Disagree
   c. Neither Disagree or Agree
   d. Agree
   e. Strongly Agree

Lastly, please answer these two questions.

1. When you first joined Mechanical Turk, which statement would you have agreed with the most?
   a. HITs on Mechanical Turk do not favor one gender over the other.
   b. Male workers have an advantage on Mechanical Turk.
   c. Female workers have an advantage on Mechanical Turk.

2. Prior to completing this particular HIT, which statement would you have agreed with the most?
   a. HITs on Mechanical Turk do not favor one gender over the other.
   b. Male workers have an advantage on Mechanical Turk.
   c. Female workers have an advantage on Mechanical Turk.

[14] Questions 1 through 4 are based on questions used in Pew Research Center surveys.

APPENDIX: EXPERIMENTAL MATERIALS FROM MANAGER EXPERIMENT

POST-EXPERIMENTAL MANAGER SURVEY (PART THREE)

PART 3 INSTRUCTIONS
Please follow the instructions for the questions below.

Please answer the 3 questions below to the best of your ability. You will earn $0.10 for each correct guess/answer.

1. How many questions do you think Worker 1 answered correctly (out of 10)?
Why did you guess this number?

2. How many questions do you think Worker 2 answered correctly (out of 10)?
Why did you guess this number?

3. Was Worker 1 female or male? (If you don't remember or don't know, take your best guess.)
   a. Female
   b. Male

Note: The rest of the questions that Managers answer are nearly identical to those from the Post-Experimental Worker Survey, starting with "What is your age?"