



The Customer Is Always Right: An Experiment on Discrimination by Customers

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:38811547>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Acknowledgements

I am deeply grateful to my advisor, Professor Benjamin Enke, for giving me the confidence to pursue my research ideas and offering me invaluable guidance every step along the way. I continue to be inspired by his contagious passion, and it is because of him that I am so excited about the world of research ahead of me. I thank Professors Michael Norton and Katherine Baldiga Coffman for taking so much of their time to share their expertise with me. I feel truly honored to have had the opportunity to develop my ideas with such insightful role models and look forward to continue learning from them. I also extend great appreciation for Daniel Svirsky and Oliver Hauser. Their encouragement and thoughtful advice means the world to me. I thank Vu Chau and Matthew Lilley for their continued support and helpful feedback. Last but not least, I thank my family and friends for their love and laughter.

1. Introduction

In 1981, over 100 male job applicants sued Southwest Airlines for gender discrimination after being denied consideration for the job of flight attendant (*Wilson v. Southwest Airlines Co.*, 1981). In response, the airline asserted that this gender criterion in hiring flight attendants was essential to its business practice because the airline's unique, "youthful, feminine image" that appealed to passengers was the "largest single component" of its success during that time. While the court ultimately rejected Southwest Airlines' claims that these customer preferences were a justifiable means to discriminate—and it remains unclear whether Southwest Airlines' statements were true or not—*Wilson v. Southwest* illustrates how discriminatory customer preferences can incentivize businesses to behave in biased ways.

The classic economic model of discrimination introduced by Gary Becker (1957) identifies three main sources of marketplace discrimination: discrimination by employers, by co-workers, and by customers. Becker's model predicts that while discrimination by employers in a competitive market reduces profits and is consequently eliminated in the long run, customers can cause persisting discrimination in the marketplace. Thus, discriminatory acts by employers may be driven by customer preferences. While much of the economic literature and popular media on discrimination has focused attention on employer decisions (e.g. Goldin & Rouse, 2000; Bertrand & Mullainathan, 2004), understanding the propensity for customers to act on their discriminatory preferences in the marketplace is crucial in order to understand the true extent to which the sources of discrimination in the marketplace drive such outcomes.

Becker (1957) assumes that actors in the marketplace with discriminatory tastes will act on such preferences. However, a large body of behavioral research suggests that people are often motivated to appear consistent with social norms—such as fairness or gender equality—even when such behavior is costly (e.g. Krupka & Weber, 2013; Norton, Sommers, Apfelbaum, Pura, & Ariely 2006; Ariely, Bracha, & Meier, 2009). People are willing to go great lengths, sometimes even

disadvantaging themselves, to avoid appearing racist, selfish, or sexist. Therefore, even in a market where individuals can act on their discriminatory preferences by buying from discriminatory companies, they may avoid such behavior because they value fair outcomes or wish to appear impartial. Ultimately, it remains unclear whether or not customers' discriminatory preferences translate to market behavior, and if so, through which mechanisms this occurs.

This paper tests a potential mechanism through which customer discrimination arises in the marketplace by asking whether individuals are more likely to reveal discriminatory preferences when they can do so by supporting a discriminatory company rather than making a discriminatory action themselves. Specifically, I ask whether customers are more likely to endorse discriminatory hiring when the hiring decision is delegated to a company rather than when the customer herself must hire. To test this question, I conduct a laboratory experiment that exogenously varies whether customers act on their discriminatory preferences directly or indirectly.

In my experiment, participants are assigned either the role of "Customer," "Company," or "Worker". Customers hire five Workers out of a pool of ten to take sports quizzes, and are told that their bonus payments increase with the chosen Workers' quiz scores. The experiment uses a sports quiz because past literature suggests that there is a widely held perception of male advantage in sports knowledge (Bordalo, Coffman, Gennaioli, & Shleifer, 2016). Thus, Customers in the experiment face a tradeoff between hiring the Workers that they believe will maximize their payment and hiring a gender-balanced set of Workers. I randomly assign whether Customers make the hiring decision directly by hand-selecting Workers from a list of ten, or indirectly by selecting a Company to make the hiring decision for them. Interpreting deviation from gender-balanced hiring as gender bias, I test for hiring differences between these two conditions. Additionally, past literature suggests that males are more likely to act on male-favoring stereotypes (Norton, Vandello, & Darley, 2004). Since the present experiment tested customer behavior in the context of a male-favoring stereotype, exploratory analysis takes into account Customer gender to test if the effect of delegating the hiring decision to a company on discriminatory hiring outcomes was greater for male Customers than it was for female Customers.

I find that Customers generally revealed a preference for male Workers. Further, while the average ratio of females hired did not differ between conditions, Customers hiring indirectly revealed discriminatory preferences more frequently than those hiring directly: the distribution of the proportion of females hired by Customers became significantly skewed when a Company acting as an intermediary was incorporated in the hiring process. In the exploratory analysis I find that male Customers demonstrated a greater preference for male Workers than female Customers did. In fact, male Customers revealed significantly greater discriminatory preferences when the hiring decision was delegated to a selected Company than when the male Customer had to make the selection himself. On the other hand, female Customers did not show such an effect.

These results suggest that customers' discriminatory preferences in a market may become more extreme due to the intermediary role of companies, but that overall average outcomes may remain unchanged. The exploratory findings suggest that certain subgroups, such as males, may be more likely to discriminate when able to do so indirectly. While the experiment in this paper specifically tests gender discrimination, the results hold implications for other forms of discrimination, such as that on the basis of racial, age, and attractiveness.

The paper proceeds as follows. Section 2 presents findings from the relevant literature. Section 3 describes the experimental design. Section 4 presents the results, and Section 5 the exploratory results. Section 6 offers some experimental limitations. Section 7 discusses, and Section 8 concludes.

2. Literature Review

2.1 Moral Wiggle Room: Image Concerns and Responsibility Diffusion

Making decisions based on self-serving preferences can be costly: one can face social ramifications from others or experience internal tension. A large literature on revealed preferences suggests the powerful effects of "moral wiggle room" (Dana, Weber, & Kuang, 2007), which is defined as the presence of acceptable rationalizations for a negative behavior. Moral wiggle room

operates through several psychological mechanisms that affect people's motives to act in their self-interest. One such phenomenon is image concerns, the idea that people do not want to appear or feel unjust. Image concerns can—both consciously and subconsciously—motivate behavior that seems altruistic or cooperative but is not necessarily rooted in the desire to be so (Dana, Cain, & Dawes, 2006). Another phenomenon is responsibility diffusion, whereby individuals are less likely to take responsibility for an action or inaction when in a group rather than when alone (Darley & Latane, 1968).

A long history of psychological literature suggests that people are more likely to satisfy their questionable preferences when offered alternative excuses to act on them. For instance, Bernstein et al. (1983) told male lab participants to watch a movie in one of two rooms of their liking: an empty room, or a room with an attractive female. Nearly all of the male participants chose to watch the movie alone when the two rooms were playing the same movie, but opted for the room with the attractive female when the rooms were playing different movies, even when the room and movie pairings were randomized. That is, when given the opportunity to rationalize with a plausible, acceptable excuse (e.g. "I like this movie more"), behavior consistent with questionable preferences increased.

While sitting next to an attractive female is a relatively innocent consequence of moral wiggle room, studies have also found that moral wiggle room can increase the likelihood and magnitude of more harmful, norm-violating behavior. Several studies have tested people's tendencies to follow fairness norms using the "dictator game". In this game, a participant is assigned either the role of "dictator" or "recipient". The dictator is paired with a recipient and selects how much money of a fixed amount (for instance, \$10) to share with that recipient. A purely self-interested dictator should simply take \$10 and leave the recipient with \$0. Yet, the majority of participants share with the other recipient, on average giving twenty to thirty percent of their money (Forsythe, Horowitz, Savin, & Sefton, 2003). This result was initially interpreted as evidence that people are not entirely self-interested and in fact hold some preferences for fairness. In a similar vein, economic social preference models emphasize the utility that people

gain from equitable outcomes and conclude that people oftentimes favor such outcomes that do not necessarily maximize their own profits (e.g. Fehr & Schmidt, 1999; Loewenstein, Thompson, & Bazerman, 1989).

However, such behavior may be a result of an ulterior motive: to maintain a positive identity. Subsequent experiments testing moral wiggle room found that sharing in dictator games significantly decreased when dictators could blame an inequitable outcome on another person (Dana, Weber, & Kuang, 2006), or some external factor (Anderoni & Bernheim, 2009). Ultimately, moral wiggle room gives rise to increasingly selfish outcomes, suggesting that sharing behavior in previous contexts was not fully motivated by fairness preferences but in some part by the actor's desire to seem fair, either to herself or to another.

Furthermore, by manipulating the ethnicity of the recipient name in two versions of the dictator game, Danilov and Saccardo (2016) found similar implications of moral wiggle room in the context of another social norm: racial equality. In one version of the experiment, dictators chose how much money to share with a recipient with either a majority-ethnicity or minority-ethnicity name. No discriminatory outcomes emerged in this iteration. However, in another version, dictators "chose" how much money to give the recipient by rolling a die. Dictators could only give a fair allocation to their recipient by lying about the die roll, and received a self-advantageous allocation if they honestly reported the die roll. Dictators shared significantly less with the minority-ethnicity name recipient than they did with the majority-ethnicity name recipient, suggesting that once the dictator could hide discriminatory dislike for the minority-ethnicity recipient under the guise of a desire to remain moral (e.g. "I am an honest person"), discriminatory behavior in sharing outcomes increased. These studies on moral wiggle room collectively suggest that moral wiggle room can increase norm-violating behavior to reveal people's self-serving preferences.

2.2 Delegation as Moral Wiggle Room

Similarly, recent studies have found that people can give themselves moral wiggle room by delegating tasks (e.g. Hamman, Loewenstein, & Weber 2010). In many contexts, such as politics and business, people choose to hire an outside party to act as an intermediary who carries out actions or makes decisions on their behalf. Traditionally, delegation was theorized to provide an efficiency benefit; the intermediary may be more informed or better equipped to carry out the task (e.g. Bolton & Dewatripont, 2005). However, recent studies have found that delegation increased the likelihood of self-interested outcomes by diffusing accountability (Hamman, Loewenstein, & Weber 2010; Coffman, 2011; Bartling & Fischbacher, 2011; Oexl & Grossman, 2013). Hamman, Loewenstein, and Weber (2010) found that by manipulating the dictator game such that dictators had to choose one of three intermediary participants who decided the money allocation for them, dictators were much more likely to choose intermediaries who had selected unfair allocations in previous rounds. Consequently, sharing significantly decreased when dictators acted through intermediaries. In fact, dictators who delegated the sharing decision to an intermediary reported feeling less responsible for an inequitable outcome than did dictators who chose the allocation themselves, and were even seen as less blameworthy by third parties (Coffman, 2011; Bartling & Fischbacher, 2011; Oexl & Grossman, 2013).

This paper proposes that companies in a market may act as intermediaries for discriminatory customer preferences. Customers rely on companies to make a host of decisions. A restaurant customer relies on the restaurant to choose which waiters and waitresses to hire, which items to put on the dessert menu, and how to arrange the tables. As customers hold buying power in a competitive market, companies can benefit by appealing to customer desires in making these decisions. As a result, customers in a market can indirectly act on their preferences when selecting which companies to purchase from. In my study, I randomized whether customers hire workers for a male stereotyped task directly or indirectly through a company. Based on the findings that increased moral wiggle room escalates self-interested behavior, I hypothesized that customers are

more likely to act on discriminatory preferences and hire more male workers when indirectly acting through a company rather than acting directly.

3. Experimental Design

The entire study consisted of three types of participants: Workers, Companies, and Customers. Workers and Companies were recruited prior to the main experiment, which tested Customer preferences. The main experiment will be described in the following section. Recruitment of Workers and Companies will be described in detail in the succeeding sections (3.2 and 3.3). The full instructions and materials for all steps of the experiment are provided in Appendix F.

3.1 Main Experiment Design: Testing Customer Preferences

Customer participants ($N = 651$) were recruited using Amazon's Mechanical Turk (MTurk), an online platform for crowdsourcing work. Many studies have reported high levels of reliability on MTurk, confirming that classical findings using economic games were successfully replicated using MTurk participants, even with smaller financial stakes—such as up to \$1 (e.g. Amir & Rand, 2012).

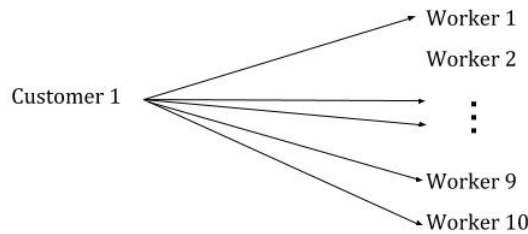
All Customers were asked to hire five Workers from a list of ten who would complete a sports quiz on their behalf. Customers' payoffs depended on how well their five Workers answered the sports questions.

Before selecting Workers, Customers were shown Workers' gender, age, preferred music genre, and birth month. These additional characteristics were presented to prevent the Customer from recognizing that the study was testing gender discrimination.

Customers were randomly assigned to one of two experimental conditions: the No-Company Condition or Company Condition (Figure 1). In the No-Company Condition, the Customer directly made the hiring decision: they were shown the list of ten Workers and asked to select the five that they wanted to hire. In the Company Condition, the Customer delegated the

hiring decision by selecting one Company out of ten to make the hiring decision on their behalf. Company payoffs, like the Customers, increased with the sports quiz scores of the hired Workers. Customers were informed of this and thus aware that Company incentives were in line with the Customers’.

No-Company Condition: Customer individually hires five Workers.



Company Condition: Customer sees every Company’s first Worker set and chooses one Company. Chosen Company’s second set of Workers are hired.

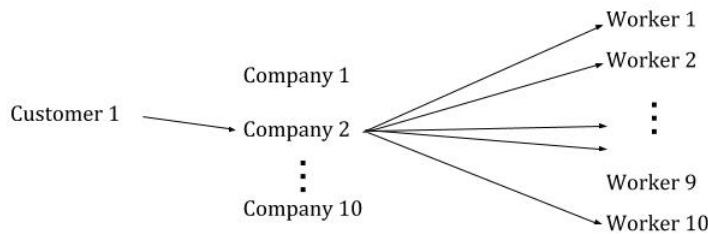


Figure 1: Experimental Conditions

Customers in the Company Condition were told that each Company had created two sets of Workers that they believed to be the best performing on the sports quiz. In order to convey to the Customers the type of Workers that each Company was likely to hire on their behalf, Customers were shown a set of five Workers that each Company selected previously. Using this information, Customers chose a Company. Then, the chosen Company’s second set of Workers was implemented and hired. A Company’s two Worker sets were allowed to be identical such that they comprised of the same Workers. As a result, Customers could expect a Company’s first and second Worker sets to be similar, if not identical.

The motivation for associating a Company with two Worker sets was to increase the propensity for Customers’ moral wiggle room. Having an uncertain second Worker set allows for the Customer to feel more that they are choosing a Company rather than a specific Worker set.

Appendix B explains the results for the relationship between a Company’s first and second Worker sets.

The ten Companies presented in the Company Condition were deliberately selected to represent one of six different Company types. Each type corresponded to a unique ratio of female Workers. This design ensured that Customers had the option of selecting from the same possible ratios of female workers as Customers in the No-Company Condition: 0, 0.2, 0.4, 0.6, 0.8, 1. Table 1 summarizes the distribution of Company type presented to Customers. While there were two of every other possible gender proportion, only one all-female Company and one all-male Company were presented: because there were five female and five male Workers, only one unique combination for each was possible. The distribution of choices in the Company Condition over-weights the extremes compared to that of the Customer Condition. This is addressed in the analysis.

Table 1: Number of Company Types in Each Company Condition Session

Company Type (Proportion of Females)	Number of Type in a Company Condition Session
0/5	1
1/5	2
2/5	2
3/5	2
4/5	2
5/5	1

At the conclusion of the experiment, participants were asked why they made the decision that they did, and to report their gender and age.

3.2 Recruiting Workers

The ten Workers that would be presented to the Companies and Customers in the subsequent steps were recruited by email from a list of undergraduate students at Harvard College. They were informed that participant information would be used as stimuli in a future experiment. They were told that in addition to the completion payment, those who were “selected” in the future

experiment would be paid an additional bonus at the conclusion of all future experimental sessions.

Worker participants were asked their age, gender, three preferred music genres, and birth month. At the end of the survey, Workers took an untimed sports trivia quiz. Once ten Workers were recruited in this fashion, each Worker was assigned a single music genre of the three that he or she indicated. This was to somewhat evenly match music genre types across gender in order to minimize effects of music on the hiring outcomes, although this is controlled for in the analyses.

Upon completion of the main experiment, Worker participants were paid a bonus payment based on the number of Customers that chose to hire him or her. Workers' sports quiz results are explained in Appendix A.

3.3 Recruiting Companies

Potential Company participants ($N = 87$) were recruited using MTurk. Company participants were shown a list of the ten pre-recruited Workers, each presented with information on age, gender, music preference, and birth month as Customers saw in the main experiment. Company participants were asked to select five Workers and were paid a bonus on the same basis as Customers. However, Company participants determined their five Workers by creating two groups of five Workers, ultimately creating two Worker sets that could be but were not necessarily identical. Upon completion, one of the two Worker sets was randomly chosen to determine Company payment. Companies were thus incentivized to reflect their true beliefs about Worker ability in constructing both sets: a rational Company participant should select the same Workers for both Worker sets. Regardless of the Worker set that was randomly chosen to determine Company payoffs, Customers in the main experiment were always shown each Company's first Worker set. Company Worker set results are described in Appendix B.

Ultimately, ten of these Company participants were used in the main experiment: one Company whose first set was 1/5 female, two whose first set was 2/5 female, and so on (see Table 1 for exact distribution). Therefore, the distribution of Company types presented to Customers was

not representative of the true distribution of Company decisions.

4. Results

4.1 Summary Statistics

A total 651 Customer participants were recruited (Table 2). Of these, 323 Customers were randomly assigned to the No-Company Condition, and 328 to the Company Condition. Customer gender across the conditions was relatively balanced, although the Company Condition had relatively more females than the No-Company Condition.

Table 2: Summary Statistics

	No-Company Condition	Company Condition
# Participants (651 total)	323	328
# Females:	159	175
# Males	164	153
Mean Customer Age	38.7 (11.60)	39.0 (11.30)

Note: Standard deviation for age in parentheses.

The main outcome of interest in the experiment is the proportion of females that each Customer revealed preference towards hiring. In the No-Company Condition, this corresponded to the gender composition of the five Workers that the Customer hand-selected from the list of ten. In the Company Condition, this corresponded to the gender composition of the five Workers of the chosen Company's first set (the set that the Customers saw). Given that a Customer hired exactly five Workers, the outcome variable had the potential values of 0, 0.2, 0.4, 0.6, 0.8, and 1. A value of 0.5 corresponded to a Worker set with an equal number of males and females. While this proportion was not a possible outcome, deviation from 0.5 was interpreted as gender bias.

4.2 Customers Exhibit Discriminatory Preferences Through Companies

Result 1. *The distribution of gender composition of hired Workers was significantly more skewed for Customers who hired through a Company than that of Customers who hand-selected their Workers.*

There was a significant difference in the frequency distributions of hiring preferences between Customers in the Company and No-Company Conditions (Figure 2; $\chi^2(5) = 92.91, p < 0.001$), providing evidence that the introduction of an intermediary company significantly altered Customers' revealed preferences during hiring.¹

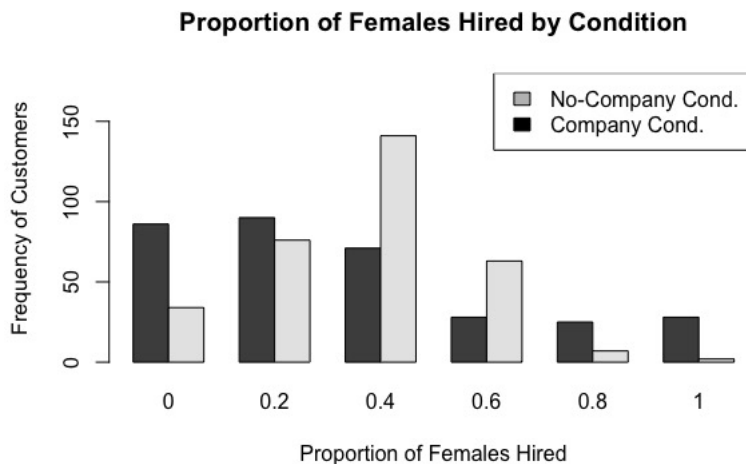


Figure 2: Histogram of Proportion of Females Hired by Customers

The distribution for the No-Company condition is not skewed ($\chi^2(2) = 0.06, p = 0.97$, Jarque-Bera), and the modal choice for Customers in this condition was to hire two females. In contrast, the distribution corresponding to the Company Condition is non-Normal ($W = 0.99, p < 0.001$, Shapiro-Wilk) and heavily skewed ($\chi^2(2) = 24.49, p < 0.01$, Jarque-Bera): the most

¹One potential concern in comparing outcome distributions between conditions is that the null distributions for each condition differs between the two due to the fact that Customers were presented with a differing number of options. Those hiring directly choose from 252 (ten choose five) options; the indirect condition offers ten choices, as noted above. Consequently, a Customer picking randomly would therefore exhibit "preferences" in either extremes more often in the Company Condition than in the Customer Condition. However, this should not cause a change in skewness. This is explained in further detail in Appendix C.

common choice was to hire one female out of five, with the second-most common choice to hire zero females. In fact, over half of Customers in the Company Condition hired zero or only one female out of five, while less than 35% of Customers in the No-Company Condition chose either of these options.²

Result 2. *A Customer hiring indirectly systematically hired fewer females than a Customer hiring directly, but the overall proportion of female Workers hired remained unchanged.*

Table 3 provides basic summary statistics on the proportion of females hired by Customers. In the No-Company Condition, Worker sets were on average 0.36 female. In the Company Condition, hired worker sets were on average 0.34 female. The means for both conditions were significantly different from 0.5 ($t = -12.628, p < 0.001$ for No-Company Condition; $t = -9.417, p < 0.001$ for Company Condition), suggesting that Customers in both conditions revealed a preference for male Workers.

Table 3: Proportion of Females Hired by Customers

	No-Company Condition	Company Condition
Mean	0.362***	0.339***
Median	0.40	0.20
Std. Deviation	0.196	0.310

Note: Comparison-of-means significance from 0.5 denoted.

*** $p < 0.001$

In the previous section, it was rejected that the proportion of females hired in the Company Condition followed a Normal distribution. As a result, a non-parametric test was used to test the treatment effect. The non-parametric Mann-Whitney Rank-Sum test results suggest that Customers who delegated the hiring decision to a Company systematically revealed a greater preference towards male Workers than did Customers who made the hiring decision themselves ($z = 3.161, p = 0.0016$). In other words, the likelihood that a Customer who hired through

²The fact that over 40% of Customers in the No-Company Condition chose two females (proportion of females 0.4) but only a little over 20% of Customers opted for this selection may suggest the role of image concerns. Hiring two females is the closest that a Customer can get to hiring a gender-balanced set of Workers while still preferring male Workers. However, this reasoning is speculative.

a Company selected fewer female Workers than did Customers who directly chose Workers is significantly greater than random chance.

However, the difference in sample means between the conditions was not significant ($t = 1.144, p = 0.253$), suggesting that Customers on average did not hire a significantly lower proportion of female Workers when hiring through a Company compared to when hiring directly.

The difference in variance in proportion of females hired between conditions begins to capture why a parametric difference-of-means test did not give significant results, despite the shift in the distribution detected by the non-parametric test. As illustrated by Figure 2, Customers in the Company Condition were more likely to hire zero or only one female than those in the No-Company Condition were, but also more likely to hire all (five) females or four females than Customers in the No-Company condition were. In other words, because Customers in the Company Condition more frequently hired gender-unbalanced Worker sets in both directions, the null difference in mean proportion of females hired between conditions is interpreted not necessarily as a null treatment effect but as a null change in the overall hiring outcome.³

4.3 Worker-level Customer Analysis

The results from the difference-in-means are supported by a logit regression model that predicts the likelihood of a Worker being hired, controlling for other Worker characteristics. For this analysis, an observation corresponded to a unique Worker-Customer pair. Since each Customer hired five of ten Workers, there were ten observations for each Customer: five corresponding Workers who were hired and five corresponding to those who were not.

The logit regression acknowledged a limitation of the experimental design that failed to be accounted for in the previous analyses: all Customers in the Company Condition were presented with the same ten Companies. Company composition of Workers was not randomized between Customers, and as a result, some Workers corresponded to a greater number of Companies than others. Consequently, a Worker's frequency of appearing in a Company is a potential confound.

³Some of these results may also be due to Customer confusion, as discussed in Section 5.1.

For instance, if all Customers wanted to hire a Worker who enjoyed rap music, but a female Worker that enjoyed rap music was only in two Companies, while a male who enjoyed rap music was in seven, then Customers may be more likely to choose the male Worker even if the preference was not based on gender. In acknowledging this, a control variable that captures the number of Companies a Worker was in was included. Additionally, an indicator variable corresponding to each Company was created to control for the types of Companies (i.e. 1/4 female) a Worker was associated with.

Table 4 summarizes the results. No matter the hiring style, female Workers were significantly less likely to be hired than male Workers ($p < 0.01$, Columns 1 - 3). Specifically, a male Worker was three times more likely than a female Worker to be hired by a Customer, even after controlling for other Worker characteristics ($p < 0.001$, Column 3).

Table 4: Logit Regression Predicting Predicting Worker Likelihood of Being Hired

	Dependent Variable: Worker hired		
	(1)	(2)	(3)
Female Worker=1	-1.233*** (0.0894)	-1.131*** (0.0944)	-1.109*** (0.139)
Company=1		0.102 (0.0896)	0.108 (0.0948)
Female Worker \times Company=1		-0.204 (0.179)	-0.215 (0.189)
Constant	0.617*** (0.0447)	0.566*** (0.0472)	0.599 (0.445)
Controls	No	No	Yes
Observations	6510	6510	6510
Pseudo R^2	0.065	0.066	0.103
Clusters	651	651	651

Note: Each observation corresponds to a unique Worker-Customer pair. Column 3 controls for Worker characteristics (age, music preferences, birth month), the number of times a Worker appeared in a Company, and type(s) of Company that Worker was associated with. Robust standard errors in parentheses clustered at the Customer-level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The coefficient on the interaction term for Worker gender and Company is insignificant ($p < 0.001, p = 0.256$, Columns 2, 3, respectively), which suggests that there was no main effect of a Company's intermediary role in a female Worker's likelihood of being hired. In other words,

the likelihood that a female Worker was hired did not differ significantly whether she was being hired directly or indirectly. This follows the implications of the comparison-of-means results: the average proportion of females hired was not significantly different between conditions.

5. Exploratory Analyses

Two main exploratory analyses were conducted. One acknowledged potential Customer confusion regarding the experimental instructions and another tested the potential interacting effects of Customer gender.

5.1 Potential Customer Confusion

At the end of the main experiment, all Customer participants were asked, “Please explain how you made your decision in the earlier page”. In this question, some Customers expressed confusion and a misunderstanding of instructions. Three main types of confusion were identified:

1. General confusion with instructions (i.e. “This makes no sense.”) ($N = 1$)
2. Discrepancy between the choice that participants claimed to have made and actually made (i.e. stating “I chose all males” but choosing an all-female worker set) ($N = 1$)
3. Misinterpreting the implications of the Company’s first worker set (i.e. “I chose the all-male Company because I thought their next Worker set would be all female”) ($N = 10$)

A naive third party coded “confused” Customer participants to reduce experimenter bias, but it is still important to consider the possible biases that are introduced by this removal process. Ultimately, twelve of the 651 Customers were deemed to have been confused, leaving 639 Customers. All twelve removed Customers were from the Company Condition, including seven females and five males. Table 5 summarizes the sample excluding confused Customers.

As presented in Table 6, the mean proportion of females hired by Customers in both conditions remained significantly less than 0.5 ($t = -12.628, p < 0.001$ for No-Company Condition; $t = -10.938, p < 0.001$ for Company Condition): Customers tended to favor male

Table 5: Summary Statistics: Without Confused Customers

	No-Company Condition	Company Condition
# Participants (639 total)	323	316 (-12)
# Females:	159: 164	168 (-7): 153 (-5)
# Males		

Note: Sample omits confused participants. Change from full sample of Customer participants denoted in parentheses.

Workers over female ones. Similar to the results from the full Customer sample, Customers who hired through a Company systematically revealed a greater preference towards male Workers than did Customers who made the decision themselves ($z = 4.081, p < 0.001$, Mann-Whitney). In fact, in this reduced sample, Customers who delegated the hiring decision to a Company on average hired significantly fewer female Workers than did Customers who hand-selected their Workers ($t = 2.208, p = 0.028$). These results do not lead to conclusive claims, but do suggest that the effect of the Company’s intermediary role on hiring outcomes may have been underestimated due to Customers’ confusion regarding experimental instructions.

Table 6: Summary Statistics: Proportion of Females Hired

	No-Company Condition	Company Condition
Mean	0.362***	0.318***
Median	0.40	0.20
Standard Deviation	0.196	0.295

Note: Sample omits confused participants. Comparison of means significance from 0.5 denoted. *** $p < 0.001$

5.2 Accounting for Customer Gender: Differing Distributions

Exploratory analyses tested whether the effect of a Company’s role as an intermediary on hiring outcomes differed between male Customers and female Customers, using the full sample without omitting potentially confused Customer participants.

As can be observed in Figure 3, the distribution of hiring decisions made by male

Customers in the No-Company Condition are similar to those made by female Customers in the No-Company Condition ($\chi^2(5) = 8.09, p = 0.151$), suggesting that there was no gender difference in revealed preferences when Customers directly hired. In contrast, there was a significant difference between Customer genders in the Company Condition ($\chi^2(5) = 12.97, p = 0.024$). The distribution of choices made by male Customers in the Company Condition follows a downward slope: the modal choice was zero females hired, giving support for the notion that male Customers held a strong preference for male Workers. On the other hand, the distribution of choices made by female Customers has a peak at around one female hired (a proportion of 0.2), with a right tail that increases slightly at all females hired. These results suggest that while male and female Customers behaved similarly when hand-selecting their Workers, the introduction of a Company in the hiring process resulted in a disparity in behavior between the genders. Specifically, male Customers appeared to respond stronger to the intermediary role of a Company than did female Customers.

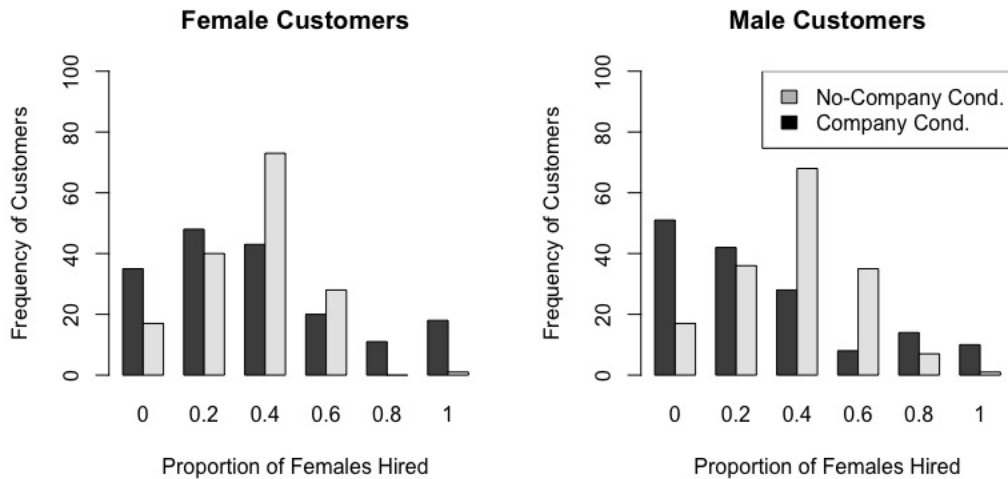


Figure 3: Histogram of Proportion of Females Hired, by Customer Gender

5.3 Accounting for Customer Gender: Statistical Tests

Table 6 summarizes the difference in treatment effect for male Customers and female Customers. For female Customers, neither the difference in mean proportion of females hired

nor the systematic tendency to hire less females between conditions was statistically significant ($t = -1.055, p = 0.292; z = 0.408, p = 0.683$, Mann-Whitney), suggesting that female Customers did not significantly alter revealed preferences whether they were hiring directly or indirectly. However, male Customers' behavior differed significantly from that of female Customers' ($t = 2.256, p = 0.025; z = 2.661, p = 0.0078$, Mann-Whitney). Males in the Company Condition systematically and on average hired a significantly lower proportion of females than did males in the No-Company Condition (Table 7; $t = -1.478, p = 0.140; z = -1.387, p = 0.167$, Mann-Whitney). This suggests, consistent with the histogram distributions (Figure 3), that male Customers were more likely to reveal discriminatory preferences when they could delegate the hiring decision to a selected Company rather than when they had to hire directly.

Table 7: Mean Proportion of Females Hired, by Customer Gender and Condition

	No-Company Condition	Company Condition	
Male Customers	0.378	0.298	$t = 2.697, p = 0.007;$ $z = 4.076, p < 0.001$
Female Customers	0.346	0.375	$t = -1.055, p = 0.292;$ $z = 0.408, p = 0.683$
	$t = -1.478, p = 0.140;$ $z = -1.387, p = 0.167$	$t = 2.256, p = 0.025;$ $z = 2.661, p = 0.0078$	

Note: Sample omits confused participants. Test statistics correspond to t-test and Mann-Whitney results, respectively.

The results from a linear probability model (i.e. ordinary least squares (OLS) regression) predicting the proportion of females hired by a Customer support the findings from the parametric comparison-of-means tests (Table 8). Pooling across the genders, the overall average proportion of females hired did not have a significant relationship with whether a Customer hired directly or not, even after controlling for Customer gender ($p = 0.253, p = 0.234$, Columns 1, 4 respectively).⁴

However, when the treatment effects are separated by Customer gender (Columns 2 and 3),

⁴A linear probability model is presented here over a logit one, particularly for ease of interpretability. An ordinal logistic model predicting the proportion of females hired is run as a robustness check and explained in detail in Appendix D. Further, a logit regression predicting a Worker's likelihood of being hired, controlling for Worker characteristics, finds results consistent with those of the linear probability model. The logit is explained in detail in Appendix E.

a different pattern emerges: male Customers on average revealed preference towards hiring eight percentage points fewer females when hiring through a Company rather than directly ($p = 0.007$, Column 3), while female Customers did not exhibit a significant change ($p = 0.292$, Column 2). Columns 5 and 6 test the full model, including Customers of both genders as well as the interaction with the treatment: the introduction of an intermediary Company in the hiring process reduced the proportion of females hired significantly more for male Customers than it did for female Customers (both Columns 5 and 6: $p = 0.007$).

Table 8: Linear Probability Model Predicting Proportion of Females Hired

Dependent Variable: Proportion of females hired						
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Females	Males	All	All	All
Company=1	-0.0232 (0.0203)	0.0289 (0.0274)	-0.0800** (0.0297)	-0.0241 (0.0203)	0.0289 (0.0274)	0.0287 (0.0274)
Male=1				-0.0226 (0.0203)	0.0321 (0.0217)	0.0312 (0.0222)
Company×Male=1					-0.109** (0.0404)	-0.108** (0.0402)
Customer Age						-0.000236 (0.000926)
Constant	0.362*** (0.0109)	0.346*** (0.0146)	0.378*** (0.0161)	0.374*** (0.0146)	0.346*** (0.0146)	0.356*** (0.0411)
Observations	651	334	317	651	651	651
R^2	0.002	0.003	0.023	0.004	0.015	0.015

Note: Each observation corresponds to a unique Customer. Robust standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A logit regression predicting the likelihood of a Worker being hired, controlling for Worker fixed effects, produce results consistent with the findings of the linear probability model: the treatment effect was significantly larger for male Customers than female Customers. This analysis

is explained in detail in Appendix D.

6. Limitations

There are limitations of this experiment: some that come from experimental participant confusion while others are a consequence of the experimental design.

As described in the exploratory analysis, a few participants (twelve of 651 identified) misunderstood the experimental instructions in the Company Condition. Some Customers ($N = 10$) reported choosing the Company corresponding to an all-female Worker set, assuming that this meant the Company's second Worker set would consist of more males. Such behavior might result in an underestimation of the treatment effect as it increases the number of Customers who hire females through a Company.

Additionally, imbalance of information between conditions was a potential confound of the treatment effect. Customers may have updated beliefs on Workers' sports quiz abilities using Companies' presented Worker sets. For instance, a Customer who previously believed that female Workers would perform significantly better on the sports quizzes than male Workers would may readjust beliefs after seeing that some Companies selected four male Workers, or even all five. A follow-up in which Customers in the No-Company Condition are also shown the ten Companies' selections but are not able to delegate the hiring decision to one would address this concern.

7. Discussion

A direct implication of the supply-demand relationship in a competitive market is that companies are incentivized to heed customer preferences. While the buying power of the customer can eliminate inefficient or undesired suppliers from a market, it also leaves potential for customer discrimination to induce and promote discriminatory company behavior. Further, as companies provide service and goods for the customer, companies by nature act as an intermediary for actions ultimately guided by customer preference. Thus, customers who may regularly consider themselves unbiased (or at least like to be seen that way), may feel more comfortable revealing

discriminatory preferences by supporting companies that act in line with such beliefs.

The present experiment found that Customers on average preferred hiring male Workers to female Workers for a male stereotyped task. Moreover, while the average proportion of females hired did not differ with whether a Worker hired directly or indirectly, tendency to hire a gender biased set of Workers increased when the hiring decision was delegated to a Company, as captured by a change in the distribution of proportion of females hired.⁵

The exploratory analysis found that while male Customers hiring through a Company were significantly more likely to reveal preference for male Workers than male Customers hiring directly, this effect did not hold for Females. In fact, within the subgroup of male Customers, hiring through a Company worsened hiring outcomes for female Workers, while this effect was not significant for female Customers.

Why the intermediary role of Companies affected male Customers and female Customers differently is ambiguous. One potential explanation is that males hold a stronger belief than females about a male advantage in sports knowledge but are not willing to act on them until moral wiggle room is presented. Another potential reason is that males are overall more sensitive to moral wiggle room. Or, perhaps females may not feel as discriminatory when favoring male Workers (e.g. “I am a female so clearly I am not sexist against females”) and therefore do not show as strong of a response to increased moral wiggle room for male-favoring preferences. These reasons are speculative, and further research must be done in order to truly conclude what drove the gender differences in customer behavior. One way to test these ideas is to examine customer behavior

⁵One important thing to note is that this experiment did not attempt to differentiate the effects of delegation between different sources of discriminatory preferences, such as ones driven by taste (Becker, 1957) and those by “statistical” beliefs (Arrow, 1972; Phelps, 1972). Both taste-based and statistical discrimination can fuel customers to exhibit discriminatory preferences in the market. Typically customer discrimination is tied most heavily with taste-based discrimination, as customers will develop group preferences regardless of true group productivity (e.g. Nardinelli Simon, 1990; Holzer Ihlanfeldt, 1998). It is most probable that taste-based discrimination is more sensitive to moral wiggle room than statistical discrimination is, as taste preferences are typically much more subjective and directly indicative of personal preference than predicted productivities. In the present study, customers hired workers to complete sports quizzes. Customers most likely preferred male workers to female workers because they believed that males on average knew more about sports and would perform better than females, not because they simply had a taste for hiring males over females. Consequently, the present experiment, in which customers’ discriminatory preferences were a result of differing predictive productivities for workers, the effect of moral wiggle room may have been underestimated compared to that for taste-based discrimination.

regarding a female-favoring stereotype.

In all, the results suggest that companies' role as intermediaries in a market affect customers' tendencies to discriminate. Becker's model (1957) of discrimination in the marketplace suggests that customer preferences have strong effects on employer behavior. Much of the economic literature pertaining to this topic has found evidence that this effect does indeed pervade several industries in the market (Nardinelli & Simon, 1990; Holzer & Ichniowski, 1998; Kahn & Sherer, 1988; Neumark, 1996; Combes et al., 2016; Bar & Zussman, 2016). My experiment adds to this by demonstrating a psychological mechanism through which the existence of an intermediary can make discrimination worse.

In fact, the experimental nature of this study could have underestimated the effects of a company's role as intermediary on revealed preferences. Amazon Mechanical Turk is a relatively private setting. Customers in either condition should not have been substantially affected by public-image concerns (as opposed to self-image concerns) in both conditions, for only the experimenter was ever notified of Customer decisions, and there was never any in-person contact between the Customers, Workers, and Companies, nor between the Customer and experimenter. This lack of an audience could have underestimated the treatment effect, as increasing moral wiggle room can also weaken public image concerns.

Additionally, as companies in a real market are multifaceted, a customer in a real market may experience even more moral wiggle room when acting on discriminatory preferences through a company than is captured by the experiment. Companies are defined not just by the composition of their hired staff, but also by several attributes such as product quality, pricing, and accessibility. Therefore, customers unwilling to act on their discriminatory preferences directly can support discriminatory companies and rationalize their choice through a multitude of alternate reasons. This again suggests a greater increase in customers' revealed discriminatory preference when acting indirectly through a company rather than directly.

A natural extension from this present experiment is to examine how intermediaries affect a third party's view of a discriminatory customer. With regards to monetary sharing, those who used

intermediaries were seen as significantly less responsible for an inequitable outcome than those who did not use an intermediary (Hamman et al. 2010; Coffman, 2011; Bartling & Fischbacher, 2011; Oexl & Grossman, 2013). Do third parties also view customers who indirectly acted on their discriminatory preferences through companies as less blameworthy for an unjust outcome than they do customers who acted directly on the same preferences? Another potential avenue to explore is whether there exists a threshold for company behavior such that a company discriminating beyond that threshold becomes too obvious in its biases. Blatantly discriminatory acts by Companies may eliminate moral wiggle room. Then, a company is incentivized to follow customers' prejudiced preferences but to make sure to remain subtle and within bounds.

8. Conclusion

My experiment suggests that many individuals hold discriminatory preferences, but will act on them even more with the introduction of an intermediary. Customers may avoid exhibiting the true extent of their discriminatory preferences when acting on them directly but may be willing to reveal these same preferences indirectly by supporting select companies. Consequently, customer discrimination may be a more threatening cause of marketplace discrimination than predicted. However, the mean proportion of females hired did not change, suggesting that even if discriminatory Customer preferences become more frequent due to the intermediary role of Companies, the average preference revealed by Customers as a whole may remain unaffected.

The implications of the present study go beyond the context of gender discrimination and hiring outcomes. The study suggests potential consequences of other customer preferences on hiring decisions, such as those regarding race, attractiveness, or age. Moreover, while this present study focused on customer preferences towards discriminatory hiring, there are many ways in which a company can satisfy customers' discriminatory preferences, such as through the choice of branding materials. This suggests that many discriminatory outcomes in the marketplace may persist due to the notion that customers are more willing to reveal discriminatory preferences indirectly through a company rather than directly. In sum, the findings of this study suggest that an

individual's propensity to reveal discriminatory preferences is heightened due to the intermediary role of companies.

Appendix

Appendix A: Worker's Sports Quiz Results

The male Workers ($N = 5$) had a higher average sports quiz score than the female Workers ($N = 5$) did. The five male Workers scored 4, 5, 5, 5, and 7. The five female Workers scored 2, 2, 3, 4, and 4.

Appendix B: Company Worker Sets

A total of 87 potential Company participants were recruited. The proportion of females in a Company's first Worker set had a significantly positive relationship with the proportion of females in a Company's second Worker set (Figure 4; Table 9, $p < 0.001$). The modal Company choice was to select two females in both Worker sets (0.4 proportion of females). Very few Companies chose all females in one set and then all males in the other ($N = 2$).

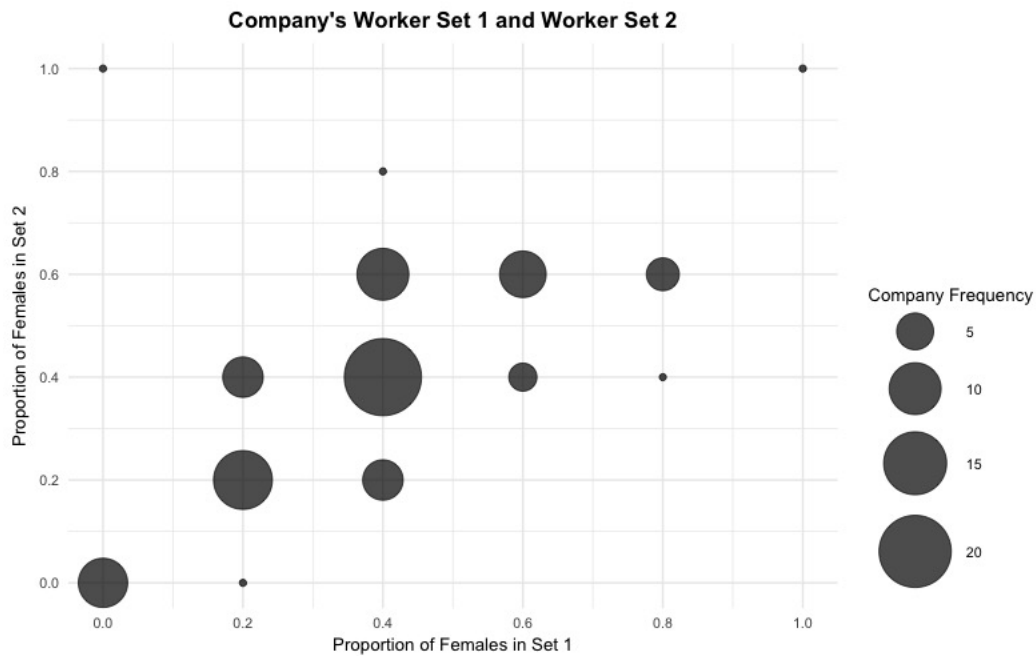


Figure 4: Relationship between Company's Worker Set 1 and Worker Set 2

Table 9: Linear Probability Model Predicting Proportion of Females in Company’s Worker Set 2

Dependent Variable: Prop. Females Set 2	
	(1)
Prop. Females Set 1	0.704*** (0.106)
Constant	0.124* (0.0482)
Observations	87
R^2	0.472

Note: Each observation corresponds to a unique Company participant. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix C: Comparing Experimental and Null Outcome Distributions

A potential concern in comparing frequency distributions between conditions is that a Customer behaving entirely randomly would reveal different behavior between conditions. Customers in the No-Company Condition had the option of choosing a Worker set of any possible five person combination of the ten workers (252) while those in the Company Condition only had ten. Consequently, the distribution of the different potential outcomes corresponding to each possible proportion of females hired significantly differs across conditions (Table 10); the null distribution assuming no gender preference for the Company Condition gives greater weight to the extreme outcomes. The implications of this issue were explored by testing whether the differences between the null and the experimental outcomes within conditions were significant, and controlling for the number of times that a Worker was presented to a Customer in regression models.

Experimental samples in both conditions showed overwhelming preference for male Workers, departing significantly from the respective null distributions (Figure 5; $\chi^2(5) = 949.87, p < 0.001$ for No-Company, $\chi^2(5) = 143.19, p < 0.001$). The extremes in the Company Condition are even more exaggerated in the experimental results compared to the null. Together, these results suggest that the experimental results are not a result of randomness.

Table 10: Null Distributions, by Condition

Proportion of Females Hired	No-Company Condition		Company Condition	
	# of Corresponding Potential Outcomes	Null CDF (approx.)	# of Corresponding Potential Outcomes	Null CDF
0	1	0.4%	1	10%
0.2	25	10%	2	30%
0.4	100	50%	2	50%
0.6	100	90%	2	70%
0.8	25	99.6%	2	90%
1	1	100%	1	100%
Total:	252		10	

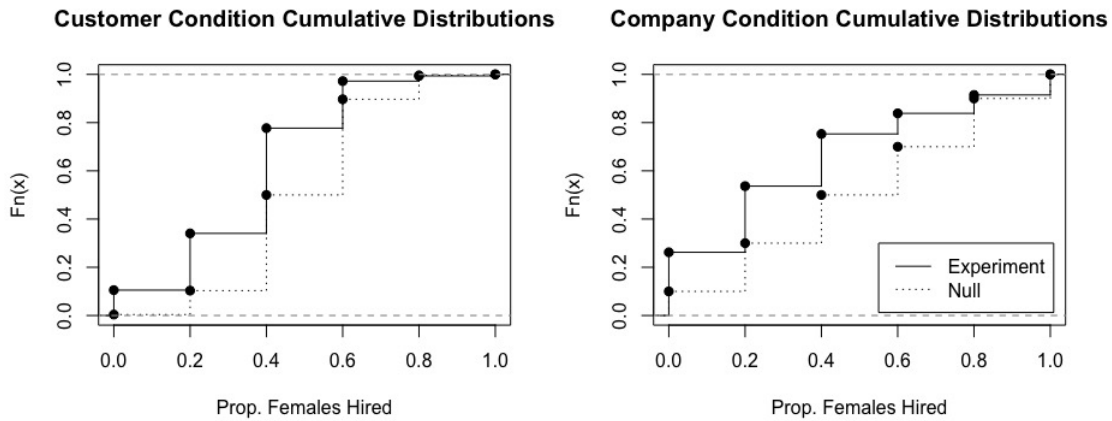


Figure 5: Experimental and Null Cumulative Distribution Functions, by Condition

Appendix D: Logit Regression Predicting Worker Likelihood of Being Hired

The logit regression predicting the likelihood of a Worker being hired within Customer genders, controlling for worker fixed effects, leads to results consistent with those of the linear probability model: the effect of the intermediary Company varied between the Customer genders (Table 11). With regards to female Customers, female Workers were not any less likely to be hired indirectly rather than directly ($p = 0.287$, $p = 0.286$ for Columns 5, 6 respectively). However, when being hired by male Customers, female Workers were significantly less likely to be hired through a Company than hired directly (both Columns 8 and 9: $p = 0.009$). Specifically, the odds that a female Worker was hired decreased by approximately one third when the male Customer

hired through a Company rather than directly.

Table 11: Logit Regression Predicting Predicting Worker Likelihood of Being Hired

Dependent Variable: Worker hired						
	Female Customers			Male Customers		
	(1)	(2)	(3)	(4)	(5)	(6)
Female=1	-1.141*** (0.122)	-1.274*** (0.129)	-1.180*** (0.193)	-1.332*** (0.132)	-0.996*** (0.137)	-1.062*** (0.203)
Company=1		-0.126 (0.118)	-0.133 (0.125)		0.359** (0.137)	0.379** (0.146)
Female×Company=1		0.251 (0.236)	0.267 (0.250)		-0.718** (0.274)	-0.755** (0.290)
Constant	0.571*** (0.0608)	0.637*** (0.0643)	0.322 (0.612)	0.666*** (0.0659)	0.498*** (0.0685)	0.941 (0.658)
Controls	No	No	Yes	No	No	Yes
Observations	3340	3340	3340	3170	3170	3170
Pseudo R^2	0.056	0.057	0.099	0.076	0.081	0.117
Clusters	334	334	334	317	317	317

Note: Each observation is a unique Customer-Worker pair. Female indicator variable corresponds to Worker gender. Columns 3 and 6 control for Worker characteristics (age, music preferences, birth month), the number of times a Worker appeared in a Company, and type(s) of Company that Worker was associated with. Robust standard errors clustered at the Customer-level in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix E: Ordinal Logistic Regression Model

The linear probability model treats the outcome variable—the proportion of females hired by a Customer—as continuous. However, by construction it can only take six possible values, and therefore is not truly continuous, but ordered. The ordinal logistic model, unlike the linear regression, relaxes the assumption that the difference between each potential outcome is equal. Namely, the difference between Customers hiring a worker set of 0 females versus 20% females is not assumed to be the same as the difference between Customers hiring a worker set of 60% females versus 80%.

The Ordinal Logistic Regression results suggest that the introduction of an intermediary

had a significantly negative relationship with the proportion of females hired by a Customer, even after controlling for Customer gender (Table 14, Columns 1 and 2: $p = 0.002$). However, once an interaction term for Customer gender and Company was included (Columns 3, 4), the implications are similar to those of the OLS regression results. The overall main effect of the Company in the hiring process is insignificant ($p = 0.743$ for Column 3, $p = 0.725$ for Column 4), but had a significant interaction effect with Customer gender (both Columns 3 and 4: $p = 0.004$): the negative effect of indirect hiring on proportion of females hired was significantly stronger for male Customers than it was for female Customers.

Table 12: Ordinal Logistic Regression Predicting Proportion of Females Hired

	Dependent Variable: Proportion of females hired			
	(1)	(2)	(3)	(4)
Company=1	-0.456** (0.150)	-0.465** (0.150)	-0.0639 (0.195)	-0.0687 (0.195)
Male=1		-0.177 (0.139)	0.203 (0.152)	0.188 (0.156)
Company×Male=1			-0.856** (0.294)	-0.850** (0.293)
Customer Age				-0.00322 (0.00649)
λ_1	-1.721*** (0.115)	-1.813*** (0.132)	-1.643*** (0.134)	-1.776*** (0.305)
λ_2	-0.448*** (0.0908)	-0.536*** (0.112)	-0.351** (0.114)	-0.484 (0.297)
λ_3	0.999*** (0.0913)	0.913*** (0.111)	1.111*** (0.115)	0.978*** (0.294)
λ_4	2.066*** (0.118)	1.977*** (0.132)	2.181*** (0.139)	2.049*** (0.300)
λ_5	2.835*** (0.174)	2.747*** (0.191)	2.953*** (0.193)	2.820*** (0.331)
Observations	651	651	651	651
Pseudo R^2	0.005	0.006	0.010	0.010

Note: Each observation is a unique Customer. Robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix F: Experimental Instructions

Experimental Instructions: Main Experiment

Thanks Page: Both Customer Condition and Company Condition -----

Thank you for agreeing to participate in this survey!
Please read the instructions carefully as your decisions in this survey will affect your bonus payment.

Intro Page: Both Customer Condition and Company Condition -----

Your task for this study is to select 5 out of 10 pre-recruited college students - who we will refer to as "workers"- to take a sports trivia quiz on your behalf. **The higher the 5 workers that you hire score on the sports trivia quiz, the higher your bonus payment will be.**

The 5 workers that you hire will each separately complete a sports trivia quiz and be paid a wage.

All 10 workers were recruited before this study was launched – their personal details are real. You will be given the 4 personal characteristics that they have disclosed: age, gender, birth month, and favorite music genre. For example, a 20-year old male worker born in January who enjoys Hip Hop music will be referred to as <20, Male, Hip Hop, January> for the entirety of this study.

Company Introduction: Company Condition ONLY -----

However, you cannot hand-select your 5 workers. You will select an MTurker to make this choice for you.

You have been matched randomly with 10 other MTurk workers – who we will refer to as Mturkers – from a previous study.

These MTurkers were each presented with the 10 pre-recruited workers, and had to select 5 workers out of each set of 10 twice. So for each MTurker, their first set of 5 workers may or may not be exactly the same as their second set of 5 workers.

The MTurkers were told that the higher the workers they chose scored on the sports trivia quiz, the higher the MTurker’s bonus payment would be.

In the next page you will be presented with the **first** set of 5 workers selected by each of these 10 MTurkers. When you select an MTurker, your bonus will depend on

the performance not of this first set, but of the **second** set of workers that this MTurker selected.

Company Choose: Company Condition ONLY -----

The 10 MTurkers are presented below, each with their **first set** of workers.

Please **select 1 MTurker** to determine which MTurker's second set of workers your bonus payment will be based on. You must select exactly 1.

<p>MTurker XZ: <22, Female, Hip Hop, July> <21, Female, Pop, August> <22, Female, Jazz, September> <21, Male, Jazz, July> <20, Male, Pop, February></p>	<p>MTurker XW: <20, Male, Pop, February> <21, Male, Rap, August> <22, Male, Pop June> <21, Male, Jazz, July> <22, Male, Country, December></p>
<p>MTurker XH: <21, Female, Pop, August> <21, Female, Country, April> <20, Female, Electronic, July> <22, Female, Jazz, September> <22, Male, Pop, June></p>	<p>MTurker XT: <21, Female, Pop, August> <20, Male, Pop, February> <21, Male, Rap, August> <22, Male, Pop, June> <22, Male, Country, December></p>
<p>Mturker XQ: <20, Male, Pop, February> <22, Male, Pop, June> <22, Male, Country, December> <22, Female, Hip Hop, July> <21, Male, Jazz, July></p>	<p>MTurker XU: <20, Female, Electronic, July> <21, Male, Rap, August> <22, Female, Jazz, September> <22, Male, Country, December> <20, Male, Pop, February></p>
<p>MTurker XK: <21, Female, Pop, August> <20, Female, Electronic, July> <21, Male, Rap, August> <21, Female, Country, April> <22, Female, Hip Hop, July></p>	<p>MTurker XG: <20, Female, Electronic, July> <22, Male, Pop June> <21, Female, Pop, August> <21, Male, Rap, August> <22, Female, Hip Hop, July></p>
<p>MTurker XB: <22, Female, Hip Hop, July> <22, Male, Country, December> <22, Male, Pop, June> <20, Female, Electronic, July> <21, Male, Rap, August></p>	<p>MTurker XJ: <20, Female, Electronic, July> <22, Female, Hip Hop, July> <22, Female, Jazz, September> <21, Female, Country April> <21, Female, Pop, August></p>

Customer Choose: Customer Condition ONLY -----

Please **select 5 workers** to determine which 5 workers your bonus payment will be based on. You must select exactly 5.

<21, Female, Country April>

<20, Female, Electronic, July>

<22, Female, Jazz, September>

<22, Male, Country, December>

<22, Female, Hip Hop, July>

<21, Female, Pop, August>

<20, Male, Pop, February>

<21, Male, Rap, August>

<22, Male, Pop, June>

<21, Male, Jazz, July>

Experimental Instructions: Recruiting Companies

Intro Page -----

Your task for this study is to select 5 out of 10 pre-recruited college students - who we will refer to as "workers"- to take a sports trivia quiz on your behalf. **The higher the 5 workers that you select score on the sports trivia quiz, the higher your bonus payment will be.**

All 10 workers were recruited before this study was launched – their personal details are real. You will be given the 4 personal characteristics that they have disclosed: age, gender, birth month, and favorite music genre. For example, a 20-

year old male worker born in January who enjoys Hip Hop music will be referred to as <20, Male, Hip Hop, January> for the entirety of this study.

Payment Page -----

You will be asked to select 5 workers twice. Your first set of 5 workers can be the same or different from your second set of 5.

One of the two sets of workers will be randomly chosen at the completion of the study and implemented. So, the 5 workers that you chose in that randomly chosen round will ultimately determine your bonus payment.

Choose Workers 1 -----

Please click on 5 workers to create your first group. If the first group (rather than the second) is randomly chosen to determine payment, these five workers' scores will be used to calculate your bonus payment.

<21, Female, Country April>

<20, Female, Electronic, July>

<22, Female, Jazz, September>

<22, Male, Country, December>

<22, Female, Hip Hop, July>

<21, Female, Pop, August>

<20, Male, Pop, February>

<21, Male, Rap, August>

<22, Male, Pop, June>

<21, Male, Jazz, July>

You must select exactly 5 workers.

Choose Workers 2 -----

Please click on 5 workers to create your second group. If the second group (rather than the first) is randomly chosen to determine payment, these five workers' scores will be used to calculate your bonus payment.

You must select exactly 5 workers.

- <21, Female, Country April>
- <20, Female, Electronic, July>
- <22, Female, Jazz, September>
- <22, Male, Country, December>
- <22, Female, Hip Hop, July>
- <21, Female, Pop, August>
- <20, Male, Pop, February>
- <21, Male, Rap, August>
- <22, Male, Pop, June>
- <21, Male, Jazz, July>

Experimental Instructions: Recruiting Workers

The purpose of this survey is to collect information to use in a next-step experiment. Your payment in this survey will not depend on any of your answers to these questions.

By continuing to partake in this survey, you are **agreeing to allow your age, gender, and top 3 hobbies to be disclosed to participants of the next-step experiment**. All other information asked of you in this survey will not be disclosed.

There is, however, a chance that you to earn a bonus payment. You will be recontacted and paid again if you are eligible to receive a bonus.

Demographics -----

Which gender do you identify with?

- Male
- Female

What is your age?

What are your 3 most favorite hobbies?

- Photography
- Dancing
- Running
- Swimming
- Cooking
- Fishing
- Gardening
- Drawing
- Music
- Woodworking
- Baking
- Scrapbooking
- Singing
- Writing
- Bicycling
- Skiing
- Boxing
- Painting
- Sewing
- Gaming

Sports Intro -----

Thank you for answering the previous questions. Now, you will be presented with **10 sports questions**. Please take the time to answer all of them to your best efforts.

Please note that your performance on these questions will NOT affect neither your payment nor your potential bonus payment.

Sports Quiz -----

1. What tennis player was listed as the world's highest-earning female athlete at the start of 2006?

- Anna Kournikova
- Serena Williams
- Maria Sharapova
- Venus Williams
- Lindsay Davenport

2. What National Hockey League player scored the most career points as of 2017?

- Wayne Gretzky
- Gordie Howe
- Mario Lemieux
- Bobby Orr
- Steve Yzerman

3. How many weeks encompass the three races of The Triple Crown of horse racing?

- 3
- 10
- 5
- 52
- 12

4. What Pittsburgh Steelers quarterback held a national schoolboy record in the javelin?

- Terry Bradshaw
- Kordell Stewart
- Hienz Ward
- Boomer Esiason
- Phil Simms

5. What distance is covered by the Olympic sprinter dubbed "the fastest man in the world"?

- 10m
- 40 yards
- 50 m
- 100 m
- 500 m

6. What player was the first to win five straight Wimbledon tennis titles?

- Arthur Ashe
- Andre Agassi
- Bjorn Borg
- Roger Federer

-
- Andy Roddick
7. Who ran the first four-minute mile?
- Roger Bannister
 - James Kwambai
 - Roger Ramjet
 - Roger Moore
 - Steve Prefontaine
8. In polo, what is a period of play called?
- Quarter
 - Chukka
 - Half
 - Set
 - Round
9. Which of these NHL teams is from New Jersey?
- Ducks
 - Flames
 - Devils
 - Jets
 - Rangers
10. Who was the first NBA draft pick in 2010?
- Greg Oden
 - John Wall
 - Derrick Rose
 - LeBron James
 - Kyrie Irving

References

- Amir, O. Rand, D. G. (2012). Economic games on the internet: The effect of \$1 stakes. *PloS one*, 7(2), e31461.
- Andreoni, J., Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5), 1607-1636.
- Ariely, D., Bracha, A., Meier, S. (2009). Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, 99(1), 544-555.
- Arrow, K. J. (1972). Models of job discrimination. *Racial discrimination in economic life*, 83.
- Bartling, B., Fischbacher, U. (2011). Shifting the blame: On delegation and responsibility. *The Review of Economic Studies*, 79(1), 67-87.
- Becker, Gary S. (1957). *The Economics of Discrimination*. Chicago: The University of Chicago Press.
- Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., Topalova, P. (2009). Powerful women: does exposure reduce bias?. *The Quarterly Journal of Economics*, 124(4), 1497-1540.
- Bernstein, W., Stephenson, B., Snyder, M., Wicklund, R. (1983). Causal ambiguity and heterosexual affiliation. *Journal of Personality and Social Psychology*, 19, 78–92.
- Bertrand, M., Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4), 991-1013.
- Bolton, P., Dewatripont, M. (2005). *Contract theory*. MIT press.
- Bordalo, P., Coffman, K. B., Gennaioli, N., Shleifer, A. (2016). *Beliefs about gender* (No. w22972). National Bureau of Economic Research.
- Buhrmester, M., Kwang, T., Gosling, S. D. (2011). Amazon’s Mechanical Turk: A new source of inexpensive, yet high-quality, data?. *Perspectives on Psychological Science*, 6(1), 3-5.
- Coffman, L. C. (2011). Intermediation reduces punishment (and reward). *American Economic Journal: Microeconomics*, 3(4), 77-106.
- Combes, P. P., Decreuse, B., Laouenan, M., Trannoy, A. (2016). Customer discrimination and employment outcomes: theory and evidence from the french labor market. *Journal of Labor Economics*, 34(1), 107-160.
- Dana, J., Cain, D. M., Dawes, R. M. (2006). What you don’t know won’t hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2), 193-201.
- Dana, J., Weber, R. A., Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1), 67-80.
- Danilov, A., Saccardo, S. (2016). Disguised Discrimination. In *Working Paper*.

-
- Darley, J. M., Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4p1), 377.
- Eagly, A. H., Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European review of Social Psychology*, 5(1), 1-35.
- Fehr, E., Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817-868.
- Forsythe, R., Horowitz, J. L., Savin, N. E., Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3), 347-369.
- Goldin, C., Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4), 715-741.
- Hamman, J. R., Loewenstein, G., Weber, R. A. (2010). Self-interest through delegation: An additional rationale for the principal-agent relationship. *American Economic Review*, 100(4), 1826-46.
- Holzer, H. J., Ihlantfeldt, K. R. (1998). Customer discrimination and employment outcomes for minority workers. *The Quarterly Journal of Economics*, 113(3), 835-867.
- Horton J, Rand D, Zeckhauser R (2010) The online laboratory: Conducting experiments in a real labor market. Technical report, National Bureau of Economic Research.
- Kahn, L. M., Sherer, P. D. (1988). Racial differences in professional basketball players' compensation. *Journal of Labor Economics*, 6(1), 40-61.
- Krupka, E. L., Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary?. *Journal of the European Economic Association*, 11(3), 495- 524.
- Laffont, J. J., Martimort, D. (2009). *The theory of incentives: the principal-agent model*. Princeton university press.
- Loewenstein, G. F., Thompson, L., Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, 57(3), 426.
- Mason W, Suri S (2010) Conducting behavioral research on amazons mechanical turk. *Behavior Research Methods*. pp 1–23.
- Nardinelli, C., Simon, C. (1990). Customer racial discrimination in the market for memorabilia: The case of baseball. *The Quarterly Journal of Economics*, 105(3), 575-595.
- 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.
- Norton, M. I., Sommers, S. R., Apfelbaum, E. P., Pura, N., Ariely, D. (2006). Color blindness and interracial interaction: Playing the political correctness game. *Psychological Science*, 17(11), 949-953.
- Norton, M. I., Vandello, J. A., Darley, J. M. (2004). Casuistry and social category bias. *Journal of Personality and Social Psychology*, 87(6), 817.
- Oexl, R., Grossman, Z. J. (2013). Shifting the blame to a powerless intermediary. *Experimental Economics*, 16(3), 306-312.

Phelps, E. S. (1972). The Statistical Theory of Discrimination. *American Economic Review*, 62(4), 659-661.

Yuracko, K. A. (2004). Private Nurses and Playboy Bunnies: Explaining Permissible Sex Discrimination. *Cal. L. Rev.*, 92, 147.