Field and Lab Experiments in Beliefs

By

Kristine Koutout

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

May 14^{th}, 2021

Nashville, Tennessee

Approved:

Andrea Moro, Ph.D.

Andrew Dustan, Ph.D.

Greg Leo, Ph.D.

Myrna Wooders, Ph.D.

Catherine Eckel, Ph.D.

To my husband, Youssef,

and my parents, Mark and Ellen.

Acknowledgments

I would like to thank Andrea Moro, Andrew Dustan, Greg Leo, Myrna Wooders, Catherine Eckel, Jennifer Reinganum, and Andrew Daughety for their valuable insights and their support of my research. I am also grateful for feedback received from seminar participants at Vanderbilt University, New York University- Abu Dhabi, Western Illinois University, and the University of Toledo as well as conference participants at the 2020 American Economic Association Annual Meetings, 2019 Advances in Field Experiments Conference, 2018 and 2019 Economics Science Association- North America Conferences, 2018 and 2019 Southern Economics Association Conference, and 2018 Missouri Valley Economics Association Conference. Research in this dissertation was made possible via financial support from the National Science Foundation Research Improvement Grant, the Russell Sage Foundation Behavioral Economics Small Grant, the Vanderbilt Department of Economics Faculty/Graduate Student Collaborative Research Grant, the Vanderbilt Department of Economics Data Acquisition Grant, the Kirk Dornbush Research Assistantship, and the Kirk Dornbush Summer Research Grant. This dissertation would not have been possible if not for the support and encouragement from my family, Ebehi Iyoha, Matthew Chambers, Chris Sichko, Trang Hoang, Teyanna Munyan, Nick Holtkamp, and Jun Zhao, and my other classmates, faculty, and support staff in the Vanderbilt Economics Department. I am especially grateful to leadership at an anonymous firm for their willingness to trust and support a graduate student conducting an experiment with their recruiting practices.

Table of Contents

List of Tables

List of Figures

Figure **Page**

Introduction

This dissertation explores the role of beliefs in behavior using field and lab experiments. The first paper establishes a methodological framework for measuring beliefs (first-order) and beliefs about beliefs (second-order) about differences between two populations. The methodological framework is implemented in a lab experiment to measure first- and second-order beliefs about the differences between women and men in their performance on a math task and their choices in an abstract bargaining task. The second paper uses this methodological framework in a combination field and lab experiment to show that workers' second-order beliefs about managers' first-order beliefs regarding the relative productivity of women and men affect workers' job application decisions. The third paper tests how players' behavior in a game of school choice is affected by strategy advice. This lab experiment demonstrates how beliefs about optimal strategies can be affected by information.

Beliefs about beliefs—second-order beliefs—about the differences between populations are important to understanding differences in outcomes between those populations. To study their potential impact, the first paper in this dissertation (co-authored with Andrew Dustan and Greg Leo) develops an incentive-compatible experimental framework for eliciting first- and second-order beliefs about the differences in any measurable characteristics between any two populations. We implement the procedure to study beliefs about the performance of men and women on math and abstract bargaining tasks. In the math task, 78% of participants believe that most *men* believe men outscore women. In contrast, 34% believe that most *women* believe men outscore women. Despite these differences in second-order beliefs, we observe no such difference in first-order beliefs. The pattern of results is similar in the bargaining task. These results have important labor market implications for the persistence of gender gaps.

The second paper in this dissertation investigates how workers' job application decisions are affected by their beliefs about hiring managers' beliefs regarding the relative productivity of women and men. To this end, I combine a natural field experiment with a lab experiment. In the field experiment, I partner with a firm to solicit approximately 5,000 job applications using ads that randomize over the gender of the hiring manager and the gender associations of the product sector. I then recruit the same job-seekers to a structured online lab experiment to elicit their beliefs about hiring managers' beliefs, based on the manager's gender and product sector. Truth-telling is incentivized with the Binarized Scoring Rule, using the methodological framework I adapt from the first paper in this dissertation. I find that men are more likely to apply for a job with a manager whom they believe has beliefs that favor men more. A one standard deviation increase in beliefs about the manager's beliefs increases the probability a man applies by 30%. On the other hand, women are unresponsive to their beliefs about managers' beliefs. These results have important implications for the sorting by gender behavior driving a large part of the gender wage gap.

The third paper in this dissertation (co-authored with Andrew Dustan, Martin van der Linden, and Myrna Wooders) implements a lab experiment to study how strategy advice affects participant decisions in a school choice game. In the Deferred Acceptance (DA) mechanism, advice to choose the dominant strategy of truth-telling induces participants to do so. In the Immediate Acceptance (IA) mechanism, advice to implement one of two heuristic strategies induces participants to choose one of those strategies. Using the varying proportions of participants choosing sub-optimal strategies in our data and a new partially-ordered typology of DA strategies, we perform exploratory analyses on mechanism performance. We find that DA outperforms IA, particularly under sub-optimal play, for almost all participants. These results demonstrate that beliefs about optimal strategies can be affected through information provision, ultimately impacting the performance of matching mechanisms.

Chapter 1: Second-Order Beliefs and Gender

(This chapter was co-authored with Andrew Dustan and Greg Leo, and is published here with their permission.)

1.1 Introduction

Do women *believe* that leaders in science, technology, engineering, and math (STEM) fields believe that women are bad at doing science? Such beliefs about beliefs—*second-order beliefs*—could drive women to sort out of STEM fields, leading to the observed gender gap in employment (Beede et al., 2011). Importantly, this belief-driven sorting could occur regardless of leaders' true beliefs about women's scientific abilities. When historically persistent beliefs about the differences between men and women—*first-order beliefs*—cause disparities, they may generate second-order beliefs that perpetuate those disparities even once first-order beliefs change.

To facilitate investigating questions of this nature, we develop an incentive-compatible experimental framework for measuring first- and second-order beliefs about the difference in any measurable characteristic between any two populations. We implement this procedure in a lab experiment to elicit beliefs about characteristics that have received particular attention for their potential to partially explain gender gaps—ability and negotiation behaviors (Croson and Gneezy, 2009; Marianne, 2011). To operationalize the study of these general categories, we have chosen tasks that are commonly used in the experimental literature as abstracted versions of these two domains. Specifically, we elicit beliefs about men's and women's performance on a timed math task (Niederle and Vesterlund, 2007; Reuben et al., 2014) and choices in the ultimatum game (Eckel and Grossman, 2001; Solnick, 2001).

We find an interesting contrast between first- and second-order beliefs. There is no evidence that men's and women's first-order beliefs differ;[1] however, both men and women believe that such differences exist. In the math task, 78% of participants believe that most *men* believe that men outscore women. In contrast, only 34% believe that most *women* believe that men outscore women. Moreover, we find no evidence of significant differences between men and women in these second-order beliefs. Similarly in the bargaining task, we find that people believe that men and women hold different first-order beliefs even though we observe no such differences in the data.

In summary, even when men and women have similar first-order beliefs, second-order beliefs about men and women can vary substantially. These statistically and economically

[1] This result is consistent with other studies that find no gender differences in beliefs about men and women such as Babcock et al. (2017), Bordalo et al. (2019), Moss-Racusin et al. (2012), and Reuben et al. (2014).

significant differences in beliefs about men's and women's beliefs may imply different incentives to acquire skills or to engage in the labor market. Our results suggest that second-order beliefs are an important, yet relatively unexplored, mechanism that could perpetuate gender gaps regardless of differences in skills or first-order beliefs.

Our framework provides a template for eliciting beliefs about the differences between two populations using state-of-the-art tools for incentive-compatible belief elicitation from experimental economics. We carefully consider 1) what property of a participant's beliefs implies a meaningful difference in beliefs about two populations, 2) what is the simplest function of two population-specific distributions that implies this property, and 3) what experimental protocol most effectively elicits this function. The resulting framework is sufficiently general to be useful to applied and experimental practitioners alike who are interested in robustly eliciting first- and/or second-order beliefs about the differences in any measurable characteristic between two populations.

To incentivize participants to *truthfully* reveal their first- and second-order beliefs, our experimental framework uses the Binarized Scoring Rule (BSR) (Hossain and Okui, 2013) to determine payment.[2] The BSR defines a payment structure that makes truth-telling optimal for all expected utility maximizers, regardless of risk preferences, as well as some non-expected utility maximizers. The payment function specified by the BSR rewards participants for the accuracy of their stated beliefs about the outcome of a random draw—the more accurate their belief, the more likely they are to earn some prize.

We innovate on the implementation of the BSR by developing a procedure that does not require teaching relatively complex mathematical concepts, like a quadratic equation, in order to explain the incentives to participants. In a typical implementation of the BSR, participants are taught the mathematical equation that determines their payment (see for example Babcock et al., 2017 or Dianat et al., 2018).[3] We instead capitalize on the fact that all payment information can be communicated using sequences of probabilities.

Our implementation uses an interactive slider to elicit beliefs that allows participants to observe, for every possible stated belief, their probability of winning in every realization of the random draw.[4] Presenting this summary of the payment rule, rather than the payment rule itself, simplifies the belief elicitation procedure considerably. The simplicity decreases the amount of time required to elicit beliefs by shortening the instructions, in addition to

[2]Danz et al. (2020) have recently demonstrated that certain implementations of the binarized version of the quadratic scoring rule may be subject to a pull-to-center effect when eliciting a subjective probability (participants' stated beliefs are biased towards 50%); however, we do not elicit a probability, our implementation is substantially different, and such an effect would only make our results more conservative.

[3]Other approaches to implementing the BSR that have been presented theoretically include rank-order lists (Leo, 2020) and dice (Wilson and Vespa, 2018)

[4]Sliders are also used by Andersen et al. (2014) for eliciting a probabilistic belief.

eliminating the need for specialized mathematical knowledge, thereby increasing the range of potential applications. Simplicity in the belief elicitation procedure is particularly important in our experimental framework because we want to elicit second-order beliefs. To incentivize truthful revelation of second-order beliefs, participants must *believe* that other participants are incentivized to tell the truth about their first-order beliefs.

The belief elicitation procedure works as follows. First, we elicit first-order beliefs. In the math task, for example, participants are asked to reveal their belief about who correctly answered more math summations in a timed task—a randomly chosen man or a randomly chosen woman (and by how many summations). Participants' stated beliefs are then compared to a random draw from a sample of people who completed the math task. The BSR maps the difference between the participant's stated belief and the realized outcome to a probability that determines how likely the participant is to win the prize. Participants who prefer higher probabilities of winning the prize to lower probabilities are incentivized to truthfully reveal their beliefs.

After the first-order belief elicitation, we ask participants to reveal what they believe a random man and a random woman chose *when asked the same question they just answered*. Participants are again rewarded based on how close their stated belief is to a realized outcome drawn from a sample of first-order beliefs using the BSR. In this intuitive way, participants reveal their second-order beliefs. Again, participants who prefer higher probabilities of winning the prize to lower probabilities are incentivized to truthfully reveal their second-order beliefs.

Our framework provides a tool for studying higher-order beliefs empirically that can generate new insights into choice behavior. Beliefs have long been recognized as crucial in decision-making under uncertainty. In particular, beliefs about the actions of others are important in strategic scenarios. To develop an internal model of what actions to expect from others, we must draw on higher-order beliefs such as our beliefs about their beliefs (second-order beliefs).[5]

Higher-order beliefs about the strategic sophistication of opponents has received substantial attention in the experimental literature, especially with regard to the "level-k" model (see Crawford et al., 2013 for a survey). The level-k model predicts game behavior based on a player's level of rationality. A level-2 player, for instance, is rational and believes that other players believe they are rational—a second-order belief. Kneeland's (2015) innovative study of strategic sophistication uses a player's chosen strategies in a series of "ring games" to mea-

[5] The analysis of these belief hierarchies was simplified by the introduction of the type space representation by Harsanyi (1967). The study of type spaces, underlying belief hierarchies, and their strategic implications has more recently been formalized in the field of *epistemic game theory* (for a general reference see Perea, 2012).

sure lab participants' levels of rationality. She finds that 71% of participants make choices that rely on second- or higher-order beliefs. The most closely related paper to ours in the literature on beliefs and strategic decision-making is Manski and Neri (2013), in which the authors elicit first- and second-order beliefs about actions in a 2x2 game to study consistency between actions and beliefs.

We are the first, to our knowledge, to directly measure higher-order beliefs outside of these abstract game contexts. Despite the link between beliefs and actions in strategic scenarios, little attention has been paid to measuring second-order beliefs and their potential impact on economic outcomes of interest in real-world markets. This lack of attention may be, in part, due to the difficulty in measuring higher-order beliefs. We address this issue within the context of a particular type of second-order belief: beliefs about the differences between two populations.

Second-order beliefs about the differences between populations may be particularly important for understanding observed differences in outcomes between those populations. We introduced this paper with an example of beliefs about the beliefs of leaders in STEM fields that could partially drive the employment gap in those disciplines; however, second-order beliefs could also contribute to empirically documented differences in men's and women's outcomes in education Lundberg (2017) and wages Blau and Kahn (2017), among other outcomes of interest. Moreover, second-order beliefs about the differences between populations characterized by other dimensions, such as race/ethnicity, religious affiliation, or sexual orientation, may be important to understanding differences in outcomes between these groups.

First-order beliefs about the differences between populations have been studied in the lab using various elicitation procedures. Some of these procedures are indirect, where beliefs are inferred from actions. For instance, in Aguiar et al. (2009) participants choose whether they prefer to have a dictator allocation from a man or woman. Similarly, Castillo and Petrie (2010) and Fershtman and Gneezy (2001) infer beliefs about different races or ethnicities from contributions in a public goods game and choices in a trust game, respectively. Beliefs have also been elicited directly. Albrecht et al. (2013) use a price list to elicit beliefs about gender differences in a spatial reasoning task. Reuben et al. (2014) directly elicit expectations about men's and women's performance on a timed math task. Similarly, Schniter and Shields (2014) directly elicit expectations about the choices of young and old people in a trust game. Unlike our procedure, both Albrecht et al.'s price list and the payment functions in Reuben et al. and Schniter and Shields have incentives that are not robust to risk preferences.

We are the first, to our knowledge, to propose second-order beliefs as a potential mechanism driving gender differences in outcomes; however, several studies find evidence of mechanisms that could be consistent with second-order beliefs. For example, Alston (2019) shows that women in a lab experiment anticipate discrimination on a sports trivia task and are willing to pay to hide their gender from prospective "employers." Women may anticipate discrimination if they *believe that employers believe* women are less productive, a second-order belief, though anticipation of discrimination could also act through preferences.[6] Charness et al. (2020) similarly show in a lab experiment that men are twice as likely as women to choose to reveal their gender in a job market for a stereotypically male task. Exploring another type of beliefs-based mechanism, Babcock et al. (2017) consider how the distribution of low-promotability tasks may impede women's career progression. They find that beliefs about willingness to accept these low promotability tasks are a primary driver of their inequitable distribution in the lab. In a related thread of literature that studies beliefs about social norms, Bursztyn et al. (2020) measure and treat men's beliefs about other men's opinions about women working outside the home in Saudi Arabia.

Coffman (2014) and Bordalo et al. (2019) study an idea closely related to second-order beliefs. In a series of lab experiments, Bordalo et al. test for the effects of "self-stereotyping" on confidence and behavior. Stereotypes such as "women are bad at math" are first-order beliefs about a measurable characteristic. In order to self-stereotype, a person must have beliefs about what those stereotypes are; therefore, when stereotypes can be classified as first-order beliefs, the person uses their *second-order beliefs* to self-stereotype. In the case of Bordalo et al., these beliefs are with respect to performance on quizzes in different trivia categories, such as pop culture and sports. By eliciting each participant's beliefs about their own performance on the tasks, as well as their beliefs about the gendered nature of the trivia category, Bordalo et al. show that stereotypes contribute to gender gaps in confidence and behavior in their experiments.

At the bargaining table, beliefs about job productivity affect the parties' beliefs about the monetary payoffs of various bargaining outcomes. However, higher-order beliefs can affect the outcomes of bargaining in more direct ways. Bargaining outcomes could be affected by differences in beliefs about preferences over monetary outcomes. For instance, Eckel and Grossman (2001) find that women are more socially-oriented than men, offering nearly twice as much in anonymous dictator games. Alternatively, bargaining outcomes could be affected by higher-order beliefs about strategies, even under common-knowledge of preferences. This is the central theme of the level-k literature discussed above.[7] By studying beliefs about strategies in an abstract bargaining game, our experiment is also designed to elicit the types of second-order beliefs that could affect wage expectations more directly through the

[6]Note that beliefs about beliefs are not the same as beliefs about preferences, and we focus exclusively on the former.

[7]See Friedenberg (2019) for a model demonstrating the role that higher-order beliefs about strategies can play in affecting outcomes even under common knowledge of preferences.

bargaining process itself.

1.2 Experimental Framework

In this section, we establish a framework for eliciting first- and second-order beliefs about the differences between two populations. To begin, we precisely define the beliefs of interest. We think of subjective beliefs as existing in the mind of our participants as subjective distributions. We want to know whether participants believe that, when we take a random draw from two populations, the characteristic of interest is most likely to be larger for the person drawn from population one or two. For example, in our experiment, we ask subjects to reveal whether they believe that a randomly chosen man or a randomly chosen woman is most likely to have scored higher on a math task.

Let X_1 be the random variable measuring the characteristic of interest in population one and X_2 be the same in population two. Then, we want to learn if the participant believes that these distributions have the property $P(X_1 > X_2) \geq \frac{1}{2}$ or that $P(X_1 < X_2) \geq \frac{1}{2}$. Either condition implies that from a randomly selected pair, the most likely outcome is the person from group one (or respectively two) has a higher value in the measure of interest. These are participants' first-order beliefs. For second-order beliefs, we want to learn if a participant believes that a random draw from population one (or two) *believes* that $P(X_1 > X_2) \geq \frac{1}{2}$ or $P(X_1 < X_2) \geq \frac{1}{2}$.

The motivation for this measure of beliefs about the differences in two populations is intuitive. Consider a professor who must choose between two otherwise identical students to advise—one is male and the other female. We want to know whether the female student believes that the professor believes it is most likely that the male student is "better" in some dimension of interest.

1.2.1 The Median Difference

The property we describe, $P(X_1 > X_2) \geq \frac{1}{2}$ or $P(X_1 < X_2) \geq \frac{1}{2}$, is implied by a statement about the median of the *distribution of differences* between the populations, $X_1 - X_2$. If there exists a median strictly greater than zero:

$$P\left(X_1 - X_2 \geq Median(X_1 - X_2)\right) \geq \frac{1}{2} \Rightarrow P(X_1 - X_2 > 0) \geq \frac{1}{2} \Leftrightarrow P(X_1 > X_2) \geq \frac{1}{2}.$$

By the same argument, the existence of a median strictly below zero implies $P(X_1 < X_2) \geq \frac{1}{2}$. By eliciting the median of $X_1 - X_2$, we elicit the participant's first-order belief regarding which population (if any) is most likely to have a higher value in the measure of interest.

Now let Z_1 be the random variable measuring first-order beliefs in population one and Z_2 be the same in population two. Draws from Z_1 (Z_2) are draws of beliefs about the median of $X_1 - X_2$. Note that $Median(Z_1) > 0$ implies a belief that the probability a person from population one believes that $Median(X_1 - X_2) > 0$ is at least $\frac{1}{2}$. Since $Median(X_1 - X_2) > 0$ implies that a participant believes that $P(X_1 > X_2) \geq \frac{1}{2}$, we can interpret $Median(Z_1) > 0$ as the belief that there is at least a $\frac{1}{2}$ probability that a randomly chosen person from population one believes that $P(X_1 > X_2) \geq \frac{1}{2}$.

1.2.2 Alternative Approaches

The goal of our procedure is to elicit whether a participant has asymmetric beliefs about two populations. We choose to elicit medians because they offer precise information about the property we are interested in—whether $P(X_1 > X_2) \geq \frac{1}{2}$ or $P(X_1 < X_2) \geq \frac{1}{2}$—at the lowest cognitive and time costs to the participant. There are alternative functions of the participant's subjective belief distributions that could also elicit this information, which we consider next.

Eliciting Probabilities

One alternative approach would be to directly elicit the probabilities of interest: $P(X_1 > X_2)$ and $P(X_2 > X_1)$. These probabilities are means of binary distributions equal to 1 when the event occurs and equal to 0 otherwise, where the events are $x_1 > x_2$ or $x_2 > x_1$. As will be discussed in more detail in the next subsection on the payment structure, the BSR can elicit a mean as well as a median by using the appropriate loss function, ensuring that we could robustly elicit these probabilities. In fact, eliciting probabilities provides cardinal information about the participants' beliefs that is unobserved in our procedure. The cost of this additional information is a more complex payment structure that requires an additional task for each belief elicited.

First, we choose not to elicit probabilities because the curvature of the quadratic rule (or any other proper rule for eliciting probabilities) creates additional incentive complexity compared to the linear incentives provided in eliciting a median. This issue may be overcome by using an alternative presentation of these incentives such as the crossover methodology (Mobius et al., 2011), or preference rankings (Leo, 2020).

More importantly, we chose not to take this approach because it requires two belief elicitations for each comparison of interest to determine which event is more likely. To determine whether $P(X_1 > X_2) \geq \frac{1}{2}$ or $P(X_1 < X_2) \geq \frac{1}{2}$ using the elicitation of probabilities would require that we elicit both $P(X_1 > X_2)$ and $P(X_1 < X_2)$. Since the outcome $x_1 = x_2$

is possible, the complement of $P(X_1 > X_2)$ is $P(X_1 \leq X_2)$, *not* $P(X_1 < X_2)$. While the cardinal information may be interesting, we argue that the precise probabilities of each event are not important enough to justify the additional cognitive costs to participants from the added complexity and doubling the number of elicitations.

Eliciting Modes of a Ternary Distribution

Another approach to determining which of a set of mutually independent outcomes is most likely is simply to ask participants which event they would like to condition their payment on. That is, ask participants to choose which outcome they think is most likely: $x_1 > x_2$, $x_1 < x_2$ or $x_1 = x_2$. This procedure is proper for eliciting the mode of a ternary distribution.

While the incentives of this procedure are clear and simple, participants with symmetric beliefs may nonetheless be incentivized to choose $x_1 > x_2$ or $x_1 < x_2$ instead of $x_1 = x_2$. Consider a continuous distribution that is identical for X_1 and X_2. Even though $X_1 = X_2$, it is sub-optimal to bet on the outcome $x_1 = x_2$ since $P(x_1 = x_2) = 0$. This also applies when X_1 and X_2 are discrete but the probability of equality is sufficiently low.

Under this payment structure, participants in our experiment who believe that men and women perform equally well on the math task would be incentivized to choose one of the non-gender-neutral outcomes simply because there are many more ways for two people to have a different math score than there are for two people to have the same math score. Therefore, we would not be able to distinguish gender-neutral participants.

In contrast, using the median procedure, a participant with symmetric beliefs is incentivized to select zero as their median belief regardless of their belief about the probability that the two randomly chosen subjects score identically. Participants with symmetric beliefs and participants whose beliefs are substantially asymmetric can always be differentiated.

Eliciting Population Medians

We elicit the median of a distribution of differences. An alternative approach would be to elicit the medians of each distribution separately and take the difference. In other words, there are two possibly relevant quantities involving medians: the median of the differences and the difference in the medians.

Eliciting the medians of X_1 and X_2 does not provide us the relevant information to assess our property of interest: whether $P(X_1 > X_2) \geq \frac{1}{2}$ or $P(X_1 < X_2) \geq \frac{1}{2}$. Specifically, $Median(X_1) > Median(X_2)$ does not imply that $Median(X_1 - X_2) > 0$. Consider the data in Table 1.1: $Median(X_1) > Median(X_2)$ since $Median(X_1) = 3$ and $Median(X_2) = 2$; however, $Median(X_1 - X_2) = -1$ implying that $P(X_2 > X_1) > \frac{1}{2}$.

Table 1.1: Example distributions illustrating that $Median(X_1) > Median(X_2)$ does not imply $P(X_1 > X_2) \geq \frac{1}{2}$

	Value		
X_1	0	3	4
X_2	1	2	5

1.2.3 Incentive Structure

When eliciting beliefs, the first priority is incentivizing truthful revelation. We begin with a payment structure that is incentive-compatible for all expected utility maximizers and some non-expected utility maximizers. The Binarized Scoring Rule (BSR), generalized by Hossain and Okui (2013), works by taking any proper scoring rule (i.e. a payment rule that reaches its maximum under truthfulness) and binarizing it, so that participants are maximizing the probability of winning the "large" prize rather than maximizing the size of the prize. This change in objective makes the payment rule incentive-compatible for all risk preferences. Using "probability currency" to induce risk-neutral behavior has a long tradition in experimental economics (Roth and Malouf, 1979; Smith, 1961), and similar binary procedures for belief elicitation are discussed by Karni (2009), Schlag et al. (2013), and Qu (2012). Schlag et al. specifically discuss binary lotteries for eliciting medians.

The probabilistic structure of the BSR outperforms other payment rules such as the popular Quadratic Scoring Rule (QSR) introduced by Brier (1950) (Hossain and Okui, 2013). The QSR incentivizes participants by varying the amount of money earned, rather than the probability of earning some fixed amount of money. That is, the closer a participant's predicted value is to the random realization, the more money they earn. This rule works for risk-neutral participants, but risk-averse participants would be incentivized to "hedge" their guess. Hossain and Okui show that participants in a lab experiment report more accurate beliefs under the BSR compared to the QSR when reporting probabilities, but the rules perform equally well in eliciting means, as theory predicts. In general, incentivized belief elicitation outperforms non-incentivized elicitation (Trautmann and van de Kuilen, 2015), particularly when there is a social cost to revealing beliefs as is the case with gendered beliefs (Babin, 2019).

The BSR proceeds as follows: participants in the experiment win either prize A or prize B, with the value of A exceeding the value of B: $U(A) > U(B)$. We are interested in the random variable X. Participants report $\theta \in \Theta$ where θ is the participant's predicted value of some function of X. A loss function $l(x, \theta)$ returns the prediction error from a random realization of X and the participant's predicted value θ. The experimenter compares the prediction error to a random draw K from a uniform distribution $U(0, \overline{K})$. If the prediction error is less than K, the participant wins prize A. Otherwise, the participant wins the lesser prize B. The form of the loss function determines which function of X participants should report. For example, the BSR would elicit the mean by binarizing the QSR loss function $(x-\theta)^2$. Other payment rules elicit the mode or quantiles, for example. The BSR procedure can be reduced to calculating the probability of winning the large prize A:

$$P(A) = 1 - \frac{l(x, \theta)}{\overline{K}}$$

As discussed in the previous subsection, we are interested in the median of participants' subjective distributions. The loss function for the median is $|x - \theta|$, so in our experiment

$$P(A) = 1 - \frac{|x - \theta|}{\overline{K}} \tag{1.1}$$

In this case, x is defined as a draw from the distribution of $X_1 - X_2$ for the first-order belief elicitation. For the second-order belief elicitation, x is defined as a draw from the distribution of Z_1 or Z_2.

The BSR is incentive-compatible for all expected-utility maximizers and some non-expected utility maximizers (Hossain and Okui, 2013). A sufficient assumption on the utility function is *monotonicity with respect to stochastic dominance*, originally defined by Machina and Schmeidler (1992). Moreover, *Theorem 4* of Hossain and Okui (2013) extends the incentive-compatibility of the BSR to account for preferences defined by prospect theory (Kahneman and Tversky, 2013a, 2013b). Although the monotonicity assumption is not satisfied by the expected utility functions in prospect theory, the incentive-compatibility of the BSR holds when the participant treats the large prize as a gain and the small prize as a loss. We have followed the advice of Hossain and Okui in setting the small prize to zero.[8]

While our procedure is incentive-compatible for many decision theories discussed in the literature, there is a possibility that the formal incentive compatibility does not extend to *some* decision process used by one of our participants. This has the potential to impair our interpretation of the elicited value as the median; however, note that $P(X > 0) \geq \frac{1}{2}$ is also implied by *any* quantile below 50% being larger than zero. In other words, $P(X > 0) \geq \frac{1}{2}$ implies that $P(X > 0) \geq \frac{1}{2} - \epsilon$ for all $\epsilon \in [0, \frac{1}{2}]$. Thus, for some hypothesized decision

[8]When losing the lottery, the subjects still leave with their show-up fee of \$5; however, we believe the subjects treat this as endowed wealth at the time of assessing the lotteries. The instructions re-enforce this by emphasizing that a loss in the lottery leads to zero gain.

theory to impair our interpretation of the ordinal information we collect, it would have to lead participants to report a quantile of their subjective belief *above* 50%.

1.2.4 Generating Samples for Incentives

In order to pay participants using the BSR, we need a sample from which to draw realizations. We pay participants for their first-order belief elicitation by sampling the measure of interest from populations one and two, X_1 and X_2. Then, we pay participants for their second-order belief elicitation by sampling from the first-order beliefs of populations one and two, Z_1 and Z_2. Therefore, we need two samples: one measuring the characteristic of interest and the other measuring first-order beliefs.[9]

The sample measuring the characteristic of interest can be generated as part of the experiment or taken from an existing data source (e.g. past experiments or administrative data). For example, the publicly available population distributions of SAT scores by gender can be sampled to incentivize elicitation of beliefs about the differences in men's and women's SAT performances. If the experimenter generates the data themselves, a single participant can be treated as a random draw from the population. Large samples are not needed— the measurement of one person from each population is sufficient.[10]

The characteristics of interest in this experiment are choices in an abstract bargaining task and scores on a timed math task. We use the Ultimatum Game as the bargaining task (see Eckel et al., 2008). In the Ultimatum Game, called "Task 1" in the experiment, Player 1 is endowed with $10 and must decide how much to offer Player 2. Player 2 decides whether to accept Player 1's offer, or to reject, in which case both participants receive nothing. We use the strategy method to elicit participants' choices as both Player 1 and Player 2. Our measure of interest is Player 2's minimum acceptable offer (MAO), the smallest amount Player 1 could propose such that the participant would accept. Appendix A shows the instructions for the strategy-style Ultimatum Game.

Any differences between men's and women's MAOs (their *willingness to accept*) can be interpreted in multiple ways. First, since any amount above $0 generates a higher payoff than rejecting, a participant interested only in maximizing earnings accepts any offer above $0. A higher MAO indicates that the participant is motivated by more than earnings and may be interested in fairness, inequality aversion, competitiveness, etc. Since the Ultimatum Game has the structure of a take-it-or-leave-it offer in negotiation, differences in MAO can

[9]Note that we cannot measure both in the same sample since we need the former to pay participants in the latter.

[10]While only one data point from each distribution is needed to incentivize belief elicitation, that data point must be truly random from the perspective of the subjects.

also be interpreted in that context. For instance, women's lower average MAO in Eckel and Grossman (2001) could be due to social norms dictating that women should be more cooperative or less demanding. This interpretation is why we call the Ultimatum Game the bargaining task.

In Task 2, the math task, participants add sets of five two-digit numbers for five minutes. Participants are paid \$0.50 for each correct sum. Appendix B shows the instructions for the math task. Previous work (Niederle and Vesterlund, 2007; Reuben et al., 2014) use timed arithmetic tasks because women and men perform equally well on them (see also Hyde et al., 1990). Despite this, people believe that men score higher than women in math tasks Reuben et al. (2014).

Unlike the sample measuring the characteristic of interest, the sample measuring first-order beliefs should be collected using the belief elicitation procedure detailed here. The measurement of second-order beliefs relies on the recursive nature of our procedure (a belief about a belief is measured in the same terms as the original belief) to help participants understand the procedure. In other words, to intuitively define second-order beliefs, we need to be able to tell participants that other participants who we are asking about *answered the same questions they just did.*

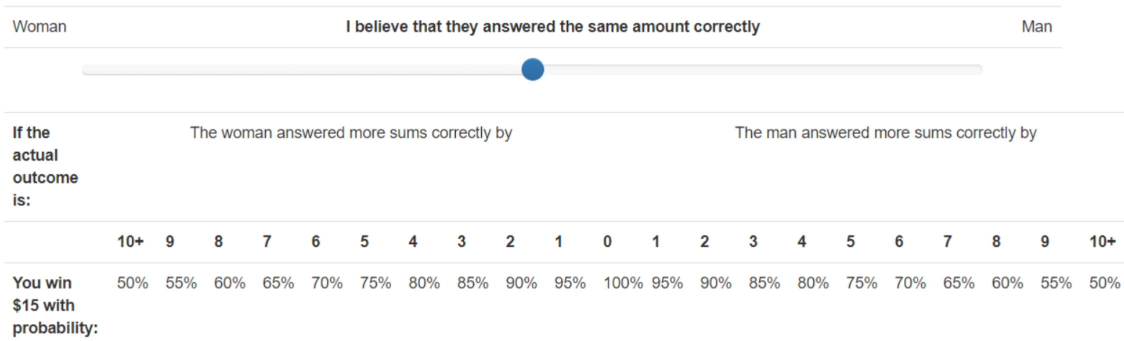
Like the sample measuring the characteristic of interest, the sample of first-order beliefs can be as small as one person from each population. For example, in this experiment, the measurement of the characteristics of interest in one man and one woman would be sufficient to elicit first-order beliefs. Likewise, the elicitation of first-order beliefs from one man and one woman would be sufficient to incentivize the elicitation of second-order beliefs. To the participant, it does not matter if the random draw used to incentivize them is from a sample of 1 or from a sample of 1,000 because the sample itself is a random draw from the population.

1.2.5 Belief Elicitation

The belief elicitation procedure begins with the first-order belief elicitations about the characteristics of interest. We elicit participants' first-order beliefs by asking them to report who they believe performed "better"[11]—a randomly drawn person from population one or a randomly drawn person from population two—and by how much. For the math task in our experiment, we ask who answered more summations correctly and, for the bargaining task, who chose the higher MAO. Participants report their beliefs by moving a slider like the one presented in Figure 1.1. The sequence of probabilities reported in the accompanying table are determined by equation (1).

[11]We put "better" in quotation marks because in tasks like the Ultimatum Game, it is unclear whether a higher or lower MAO is better. This language is not used in the experiment.

Figure 1.1: Example of slider interface used for belief elicitation



The slider's starting position is always the center, reporting that the man and woman scored equally in the task. Participants move the slider to the right if they believe the randomly selected man scored higher on the math task (or chose a higher MAO) and to the left if they believe the randomly selected woman scored higher (or chose a higher MAO). When the participant moves the slider, the table updates at each point of the support to show the associated sequence of probabilities of winning the large prize based on each possible realization of the random draw. Participants are told in the instructions that the procedure is designed such that it is optimal to report their best guess about the median.

Implementing equation (1) requires a choice for $\overline{K}$. Recall that $\overline{K}$ is the maximum on the uniform distribution from which we take a draw to compare to the evaluated loss function. That means $\overline{K}$ determines the size of the support over which participants can express their beliefs. There are trade-offs in the selection of $\overline{K}$. The larger the support, the flatter the slope on the objective function, weakening the incentive to be precise; however, a small $\overline{K}$ might truncate the choices of participants with more extreme beliefs. We choose to elicit beliefs over a 21 point support for both tasks: gender neutrality at zero and ten points on either side. This support matches the natural maximum of the Ultimatum Game, in which the largest difference is between a MAO of \$10 and \$0. Since there is no natural maximum for the math task, the choice might constrain our participants, so we label the endpoints as "10+".[12]

After eliciting participants' first-order beliefs, the belief elicitation procedure continues by informing participants that people from populations one and two answered the same questions they just did. We elicit second-order beliefs by asking participants to report what they believe a randomly drawn person from population one (and two) reported when *they* answered those questions. In our experiment, we ask participants what they believe a randomly chosen man from a previous session reported and, likewise, what a randomly

[12]We do not observe responses at the endpoints in our experiment.

chosen woman reported as her first-order belief for each characteristic. That is, we elicit four second-order beliefs— one for each gender/characteristic pair. As in the first-order belief elicitation, participants report their beliefs using a slider like Figure 1.1.

While we collect cardinal information about participants' median beliefs, the median was chosen only because it has an ordinal interpretation about underlying probabilities. The additional cardinal information may be interesting, but the cardinal results confound two factors: the magnitude of participants' beliefs about population differences and participants' beliefs about absolute levels of characteristics in the populations.

To illustrate this point, consider a participant who reports that their median belief is that a randomly selected man answers two more summations correctly than a randomly selected woman. The interpretation of those "two more summations" differs based on whether the participant believes people answer five summations total on average or twenty summations. Moreover, it is unclear how the additional quantitative results would be more informative than ordinal results. For example, knowing that people believe that men believe men outscore women on a simple math task may inform our understanding of the employment gap in STEM fields, but knowing specifically how many more math summations they are believed to outscore women by on this one particular task would not. Thus, in the Results section we focus on the ordinal information provided by the median beliefs.

1.2.6 Salience of Gender

We elect to make gender salient in our procedure, rather than try to disguise our intentions. Experimenters often obfuscate the purpose of an experiment about gender to avoid confounding factors such as an experimenter demand effect or social costs associated with revealing gendered beliefs. For example, one concern with our procedure is that *most* of the possible choices involve expressing some difference between men and women. This could create a demand effect, leading participants with neutral beliefs to express differences. On the other hand, revealing beliefs that "favor" one gender over another could impose some social cost on participants. This cost would bias results towards zero. Instead of attempting to design our experiment to neutralize these biases, we rely on our relatively strong and carefully designed monetary incentives to ensure that our results indicate true patterns in participants' beliefs.

Obfuscating gender is particularly untenable in our experiment because we want to elicit second-order beliefs. To elicit true second-order beliefs in our framework, it is vital that participants clearly understand that they are revealing their beliefs about men and women *and believe that other people clearly understood* that they were revealing their beliefs about men and women. When gender is obfuscated, this requirement becomes more burdensome since participants must also believe that other participants saw and interpreted the signal of gender in the same way they did. Even supposing that participants all interpret the signal of gender identically, obfuscating gender in both the first- and second-order belief elicitations means that participants reporting their second-order belief would need to deduce both the gender of the person in the first-order belief elicitation and the implied gender difference that person is asked about. This relatively complex task would confound the results in unclear ways.

These potentially confounding factors are relevant to any experiment on socially sensitive topics, but it is important that we consider the implications for our interpretation of second-order beliefs. In our procedure, we incentivize participants to report what they believe another person *reported* as their first-order belief and interpret that elicitation as the participant's second-order belief. Participants who believe there are social costs, experimenter demand effects, or any other biasing factors, should account for them when reporting their second-order belief. This argument relies on participants being rational enough to consider the incentives of other participants. We believe participants are sophisticated enough to account for the full range of incentives affecting other participants;[13] therefore, a conservative interpretation of our most compelling results would be "participants believe that men and women *reveal* different first-order beliefs" rather than "*have* different beliefs."

1.2.7 Implementation

We implemented this experiment at the Vanderbilt University Experimental Economics Lab (VUEEL) from November 2017 to January 2018. Participants were recruited using the ORSEE system (Greiner, 2015), with no restrictions on who could participate. No one participated in more than one session of the experiment. The belief elicitation data come from 157 participants, 80 of which are male and 77 of which are female. The sample is comprised almost exclusively of Vanderbilt undergraduate students. Table 1.2 lists the sample sizes by gender for the samples used to incentivize belief elicitations as well as the sample that generates our belief elicitation data.[14]

Participants in the belief elicitation sessions received paper copies of the instructions used to measure the characteristics of interest, but completed the experiment on laptops using the oTree software (Chen et al., 2016). All instructions were read aloud by the experimenter. After the belief elicitation, the experiment concluded with a demographic survey. Each session

[13] This belief is consistent with the results of Kneeland (2015), who finds that a large majority of subjects are at least second-order rational.

[14] Recall that we need only one draw from each population for each sample, but we collect slightly more.

Table 1.2: Sample sizes for incentive and belief elicitation samples

	Task	First-order beliefs only	Belief elicitation
Female	12	4	77
Male	10	4	80

Notes: Column headers denote the samples. We measure the characteristics of interest in the "task" sample (to be used in incentivizing first-order belief elicitations). The "first-order beliefs only" sample is used to incentivize second-order belief elicitations for participants in the full "belief elicitation" sessions, which provide the data analyzed in the experiment.

lasted approximately 30 minutes.[15] See Appendix C for screenshots of the full experiment.

Participants received $5 for participating in the experiment and could earn the "large" prize of $15 from the belief elicitations. One decision out of the six was chosen at random at the end of the experiment to determine payment and participants earned $18.09 on average, including the participation fee.

1.3 Results

We present the experimental results for the math task, summarized in Table 1.3 and Figure 1.2, followed by the bargaining task, summarized in Table 1.4 and Figure 1.3, and an intra-participant comparison of beliefs. We do not have predefined hypotheses about these belief distributions. One way to develop such hypotheses would be to use "common knowledge" arguments; however, the beliefs underlying those common knowledge arguments are precisely what we are seeking to measure. We describe the data instead.

1.3.1 Math Task

Most participants believe that there is *some* difference in men's and women's performance on the math task (86%, $SE = 2.8\%$), with 55% ($SE = 4.0\%$) believing that men outscore women. Testing for a difference in proportions, we cannot reject at conventional significance levels that men and women have the same probability of believing that men outscore women (59% for men vs. 51% for women, $p = 0.308$). Similarly, using a Wilcoxon rank-sum test, we cannot reject that men's and women's first-order belief distributions are identical ($p = 0.344$).[16] We note, however, that we cannot rule out a range of differences in first-order

[15]The time from the actual start of the experiment to when all participants completed the six belief elicitations and demographic survey was typically 15 to 20 minutes.

[16]The Wilcoxon rank-sum test for equality of first-order belief distributions uses the ternary distributions illustrated in Figure 1.2. Recall that we collect cardinal information, even though our outcome of interest is

beliefs, including both positive and negative differences. For example, the 95% confidence interval for the men-women gap in the proportion believing that men outscored women is $[-7\%, 24\%]$.

Although we lack evidence that first-order beliefs differ by gender, participants believe that such differences in first-order beliefs exist. Using the Wilcoxon signed-rank test, we reject equality of distributions of second-order beliefs about men's beliefs and women's beliefs regarding math performance ($p = 0.000$).[17] Furthermore, 78% (SE=3.3%) of participants believe that most men believe men outscore women, while only 34% (SE=3.8%) of participants believe this about women's first-order beliefs, and a test of difference in proportions rejects that they are equal ($p = 0.000$). As with the first-order beliefs, we cannot reject that men's and women's second-order belief distributions are identical, with respect to either men's ($p = 0.257$) or women's ($p = 0.137$) first-order beliefs.[18]

While this experiment lacks the statistical power to make definitive statements about whether participants' second-order beliefs are correctly calibrated, some conclusions are possible. The 95% confidence set for the median of men's first-order ternary belief distribution includes both "no difference between man and woman" and "man outscores woman," but excludes "woman outscores man." The same is true for women's first-order beliefs. Thus, only a reported second-order belief (about either a man's or a woman's reported first-order belief) of "woman outscores man" can be classified as miscalibrated. Participants are much more likely to report this miscalibrated second-order belief about women (38%, SE=3.9%) than for men (8%, SE=2.2%), a difference that is statistically significant at the 1% level. To the extent that we can detect miscalibrated second-order beliefs in the data, it seems that the difference in gender-specific second-order beliefs is driven by participants wrongly believing that most women's first-order beliefs "favor" women.

To summarize, we are unable to reject equality in first- and second-order belief distributions between men and women, but have strong evidence that participants *believe* men and women hold different first-order beliefs. In particular, most participants believe that most men believe men outscore women, while they do not believe this about women.

---

ternary. The Wilcoxon rank-sum test for equality of the first-order *cardinal* distributions gives $p = 0.652$. The cardinal distributions for all elicitations are in Appendix D.

[17]The Wilcoxon signed-rank test is used to account for intra-participant dependence. The Wilcoxon signed-rank test also rejects equality of cardinal second-order belief distributions ($p = 0.000$).

[18]We again use the Wilcoxon rank-sum test for equality of distributions, comparing the gender-specific ternary distributions. The null hypothesis of no differences in the gender-specific cardinal second-order belief distributions cannot be rejected for beliefs about men ($p = 0.180$) or about women ($p = 0.173$).

Table 1.3: Belief elicitation results for math task

	All	Men	Women	Difference
First-Order Beliefs				
W>M	0.312	0.287	0.338	-0.050
	(0.037)	(0.051)	(0.054)	(0.074)
W=M	0.140	0.125	0.156	-0.031
	(0.028)	(0.037)	(0.042)	(0.055)
W<M	0.548	0.588	0.506	0.081
	(0.040)	(0.055)	(0.057)	(0.079)
Second-Order Beliefs, about Men				
W>M	0.083	0.087	0.078	0.010
	(0.022)	(0.032)	(0.031)	(0.044)
W=M	0.140	0.175	0.104	0.071
	(0.028)	(0.043)	(0.035)	(0.055)
W<M	0.777	0.738	0.818	-0.081
	(0.033)	(0.050)	(0.044)	(0.066)
Second-Order Beliefs, about Women				
W>M	0.376	0.412	0.338	0.075
	(0.039)	(0.055)	(0.054)	(0.077)
W=M	0.287	0.313	0.260	0.053
	(0.036)	(0.052)	(0.050)	(0.072)
W<M	0.338	0.275	0.403	-0.128
	(0.038)	(0.050)	(0.056)	(0.075)
Observations	157	80	77	157

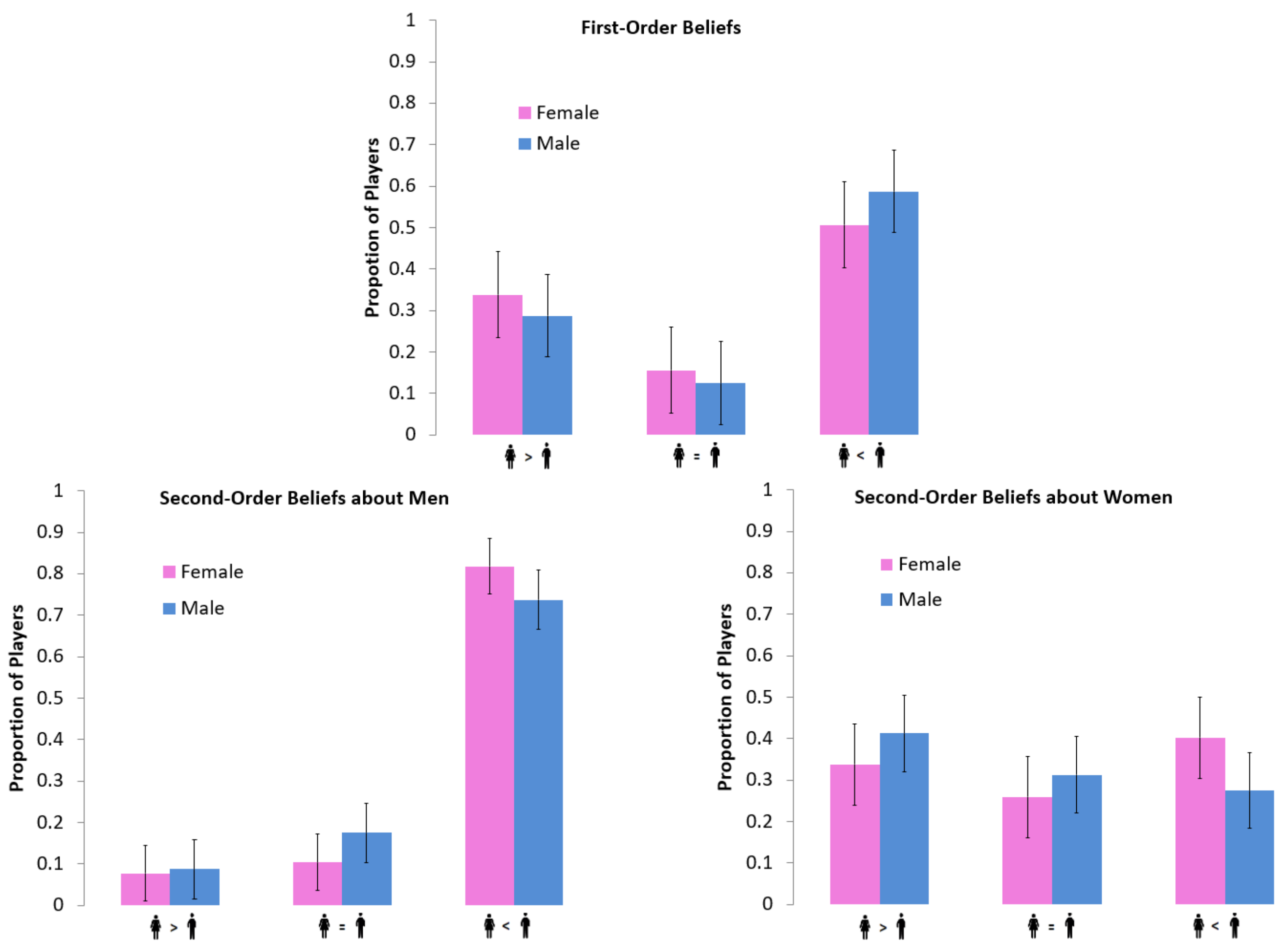
Notes: Columns (1) to (3) reference subsamples. Column (4) reports the differences between the men and women subsamples. Standard errors are reported in parentheses underneath the proportions. The rows "W>M", "W=M", and "W<M" report the proportion of participants in the math task who believe that the woman scores higher, the woman scores the same, the woman scores lower compared to the man.

Table 1.4: Belief elicitation results for bargaining task

	All	Men	Women	Difference
First-Order Beliefs				
W>M	0.089	0.113	0.065	0.048
	(0.023)	(0.036)	(0.028)	(0.045)
W=M	0.204	0.225	0.182	0.043
	(0.032)	(0.047)	(0.044)	(0.064)
W<M	0.707	0.662	0.753	-0.091
	(0.036)	(0.053)	(0.049)	(0.072)
Second-Order Beliefs, about Men				
W>M	0.166	0.188	0.143	0.045
	(0.030)	(0.044)	(0.040)	(0.059)
W=M	0.255	0.188	0.325	-0.137
	(0.035)	(0.044)	(0.054)	(0.069)
W<M	0.580	0.625	0.532	0.093
	(0.040)	(0.054)	(0.057)	(0.079)
Second-Order Beliefs, about Women				
W>M	0.089	0.113	0.065	0.048
	(0.023)	(0.036)	(0.028)	(0.045)
W=M	0.236	0.225	0.247	-0.022
	(0.034)	(0.047)	(0.049)	(0.068)
W<M	0.675	0.662	0.688	-0.026
	(0.037)	(0.053)	(0.053)	(0.075)
Observations	157	80	77	157

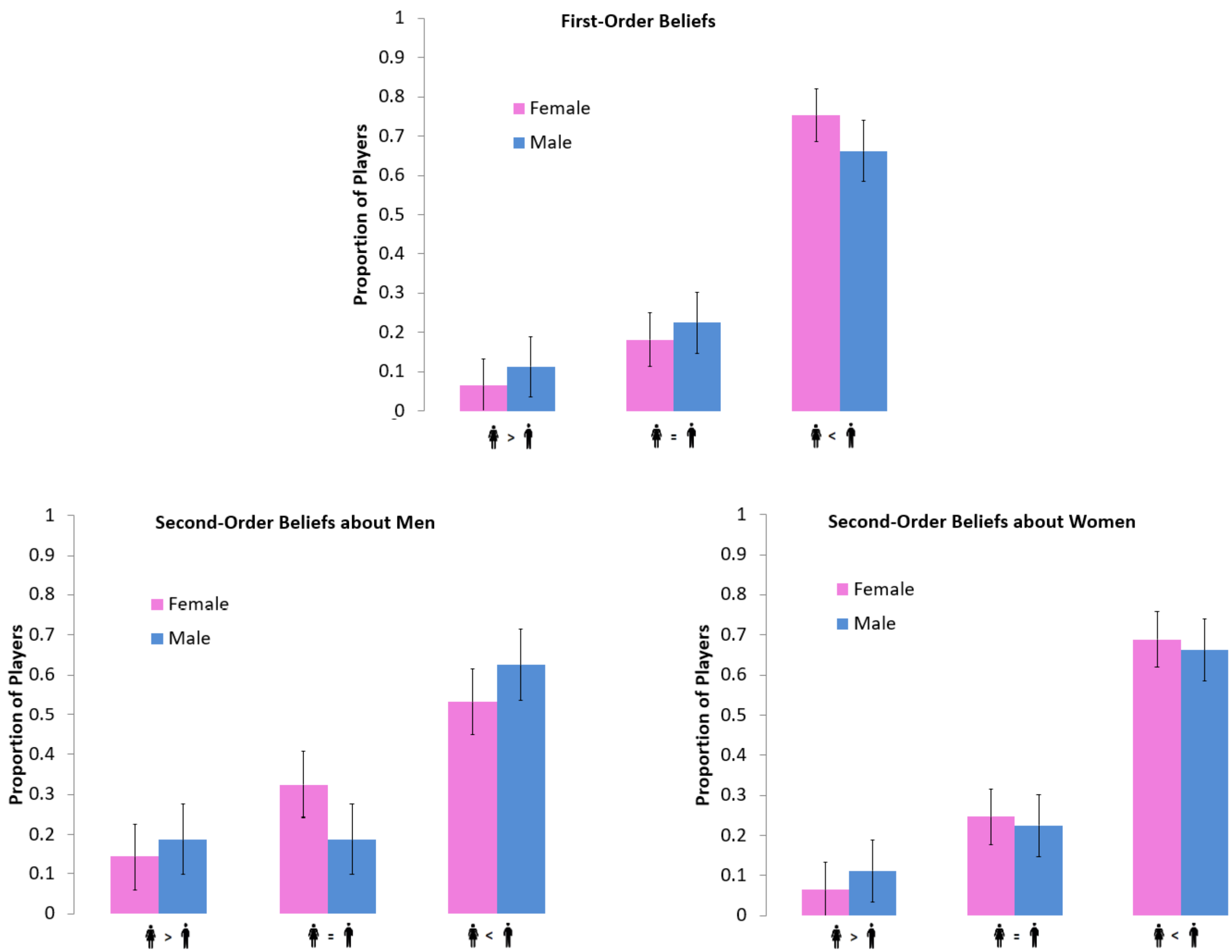
Notes: Columns (1) to (3) reference subsamples. Column (4) reports the differences between the men and women subsamples. Standard errors are reported in parentheses underneath the proportions. The rows "W>M", "W=M", and "W<M" report the proportion of participants in the bargaining task who believe that the woman chooses higher MAO, the woman chooses the same, the woman chooses lower MAO compared to the man.

Figure 1.2: Belief elicitations about the math task



Notes: From left to right on each graph, the bars represent the proportion of participants in the math task who believe that the woman scores higher, the woman scores the same, the woman scores lower compared to the man.

Figure 1.3: Belief elicitations about the bargaining task



Notes: From left to right on each graph, the bars represent the proportion of participants in the bargaining task who believe that the woman chooses higher MAO, the woman chooses the same, the woman chooses lower MAO compared to the man.

1.3.2 Bargaining Task

Most participants believe that men choose a higher MAO than women (71%, SE=3.6%). Similar to the math task, we cannot reject that the proportions of men and women believing that men report a higher MAO are equal ($p = 0.212$).[19] Nor can we reject that the distributions of men's and women's first-order beliefs are the same ($p = 0.191$).[20] Again, we cannot rule out positive or negative differences in these proportions: the 95% confidence interval for the men-women gap in the proportion believing that men had a higher MAO is $[-23\%, 5\%]$. Interestingly, the point estimates suggest that men are 9.1 percentage points (SE=7.2%) *less likely* than women to hold the "stereotypical" belief that men have a higher MAO than women, although this difference is not statistically significant at conventional levels.

Participants again believe that men and women hold different first-order beliefs. We reject equality of the distributions of second-order beliefs about men's and women's first-order beliefs about which gender proposes a higher MAO ($p = 0.027$).[21] Interestingly, 68% (SE=3.7%) of participants believe that most women believe men choose a higher MAO, which is marginally higher than the 58% (SE=3.9%) of participants believing this about men's first-order beliefs ($p = 0.072$). Again, we cannot reject that men's and women's second-order beliefs are the same about men ($p = 0.475$) or about women ($p = 0.609$).[22]

The 95% confidence sets for the medians of men's and women's first-order ternary belief distributions include only "man has higher MAO than woman," meaning that all other second-order beliefs are miscalibrated. Second-order beliefs about both genders are often miscalibrated: 42% (SE=4.0%) of second-order beliefs about men are miscalibrated, as are 32% (SE=3.7%) about women. This difference in the rate of miscalibration between genders is marginally significant ($p = 0.080$), indicating that second-order beliefs about men are less accurate than those about women in this task.

Similar to the math task, we do not have consistent evidence of gender differences in either first- or second-order beliefs about the bargaining task. Yet participants believe that men and women differ in their first-order beliefs, being more likely to believe that women believe men choose a higher MAO than they are to believe this about men.

[19]The tests used for the bargaining task analysis are the same as those used for the math task: tests for differences in proportions when comparing proportions across genders, Wilcoxon rank-sum tests when testing for differences in ternary distributions between genders, and Wilcoxon signed-rank tests when testing for within-participant differences in second-order beliefs with respect to different genders.

[20]The test for differences in the cardinal distributions does find evidence of a difference ($p = 0.020$), but given the concerns with interpreting the cardinal measures and the lack of evidence for differences between the ternary distributions, we do not interpret this finding further.

[21]We also reject equality of the cardinal second-order belief distributions ($p = 0.001$).

[22]We also fail to reject equality of the cardinal second-order belief distributions about men ($p = 0.425$) or about women ($p = 0.925$).

1.3.3 Intra-participant Beliefs

Here, we describe the extent to which participants' second-order beliefs mirror their first-order beliefs.[23] This analysis is useful in understanding whether people form second-order beliefs about others in the same population solely by considering their own beliefs. To do this, we compare a participant's reported first-order belief to their second-order belief about a person of the same gender. Table 1.5 shows that, while the majority of participants believe that other participants of their same gender believe the same as themselves (57%, $SE = 4.0\%$ for the math task and 68%, $SE = 3.7\%$ for bargaining), these proportions are far from 1 and are quite similar for men and women.

Table 1.5: Proportion of participants reporting same-gender second-order beliefs matching their own first-order beliefs

	All	Men	Women
Math	0.567	0.613	0.519
	(0.040)	(0.054)	(0.057)
Bargaining	0.682	0.675	0.688
	(0.037)	(0.052)	(0.053)
Observations	157	80	77

Notes: Comparison is with respect to ternary belief distributions. Columns refer to subsamples. Standard errors are reported in parentheses underneath the proportions.

Table 1.6 further explores the correspondence between first- and second-order beliefs, now conditioning on the first-order belief. Participants who believe that men perform better in the math task are more likely to believe that others of the same gender share that belief (70%, $SE = 5.0\%$) than those who believe the genders perform the same (41%, $SE = 10.7\%$) and those who believe that women performed better (also 41%, $SE = 7.1\%$), and these differences in proportions are statistically significant at conventional levels ($p = 0.001$ and $p = 0.012$, respectively). Similarly for the bargaining task, participants believing that men choose a higher MAO were more likely to believe that others of the same gender shared that belief (78%, $SE = 4.0\%$) than those believing that genders performed the same (50%, $SE = 9.0\%$, $p = 0.003$ for difference in proportions) or that women choose a higher MAO

[23]We have not provided results on whether second-order beliefs are accurate. If second-order beliefs were accurate, then the distribution of second-order beliefs would converge to a point-mass at the true median of first-order belief medians. Thus, the accuracy of the distribution of second-order beliefs requires a degenerate distribution of second-order beliefs. In this sense, the "accuracy" of second-order beliefs does not involve second-order and first-order beliefs being similar. Instead, we study the similarity in these distributions through the lens of intra-participant consistency.

(36%, $SE = 13.3\%$, $p = 0.001$ for difference in proportions).[24]

Table 1.6: Proportion of participants reporting same-gender second-order beliefs matching their own first-order beliefs, by first-order beliefs.

	W>M	W=M	W<M
Math	0.410	0.410	0.700
	(0.070)	(0.105)	(0.050)
Observations	49	22	86
Bargaining	0.357	0.500	0.775
	(0.128)	(0.088)	(0.039)
Observations	14	32	111

Notes: Columns specify participant's first-order belief: woman higher than man, gender-neutral, and man higher than woman. Comparison is with respect to ternary belief distributions. Standard errors are reported in parentheses underneath the proportions.

1.4 Discussion

We establish an experimental framework for measuring both first- and second-order beliefs about the difference in some measurable characteristic between two populations. The procedure is simple in that participants do not need specialized mathematical knowledge to understand the incentives. Instead, we use an interactive slider that presents all relevant payment information through sequences of probabilities.[25] Moreover, the procedure inherits the robust incentive-compatibility of the BSR for all participants who prefer higher probabilities of winning a prize to lower probabilities. Our experimental framework enables the easy adaptation of the procedure to elicit beliefs about any number of interesting characteristics and populations. We also note that the procedure can be used to elicit beliefs about non-random outcomes (the height of Mt. Everest), as well as higher-order beliefs.[26]

We implement the procedure in the lab to measure beliefs about the differences between men and women in their performance on a math task and choices in an abstract bargaining task. Our results are interesting, but intuitive. While men and women exhibit no statistically distinguishable differences in their first-order beliefs, people *believe* that such differences exist.

[24]We disaggregate the results by gender in Appendix D and note that there is a significant difference between men's and women's intra-participant beliefs in the math task; however, we do not interpret these results as the cell sizes are very small.

[25]The programming for this slider is available from the authors upon request.

[26]The recursiveness of the procedure means that eliciting third- or higher-order beliefs is limited only by the rationality of the participant.

The potential implications of such discordant beliefs in real-world markets are far-reaching. Consider a woman who believes that male managers believe men to be more productive than women in STEM fields. She may pay some economic cost to be matched with a female manager rather than a male manager, even though there may be, in fact, no difference in male and female managers' beliefs. These second-order beliefs could contribute to observed gender differences in outcomes like the employment gap in STEM, regardless of differences in first-order beliefs or skills. Beyond the labor market, these second-order beliefs may have important implications in marriage and fertility decisions, as well as human capital investment in the next generation.

Mechanisms have been proposed to explain gender differences in market outcomes that may be, in part, driven by second-order beliefs, further underlining their importance. For example, statistical discrimination models (see Fang and Moro, 2010 for a review) require that minority workers believe that employers believe they have lower human capital—a second-order belief—to establish the self-fulling prophecy. Dianat et al. (2018) recognize the necessity of workers' "second-order rationality" in their lab experiment artificially creating statistical discrimination. Glover et al. (2017) find evidence of a self-fulfilling prophecy in French grocery stores. Minority workers exert more effort than majority workers under unbiased managers, but perform worse under biased managers. The workers' second-order beliefs about managers' beliefs are essential to explaining this behavior. Our experimental framework can be used to test the underlying assumptions on beliefs in models and experiments such as these.

A number of avenues are open for future work. First, and foremost, is showing whether second-order beliefs affect market behavior. Second is understanding how second-order beliefs are formed. One promising theory applies the stereotype model in Bordalo et al. (2019). In this model, second-order beliefs are an exaggeration of first-order beliefs.

While we have focused on gender in this paper, the procedure is sufficiently general to study differences about other types of populations. The experimental framework can be used to elicit beliefs about differences by races/ethnicities, religious beliefs, sexual orientation, STEM/non-STEM workers, and political affiliation. Only small samples from the populations of interest are required to incentivize first- and second-order belief elicitation, enabling the study of beliefs about much smaller and difficult to recruit populations than was previously practical. Second-order beliefs likely play a role in how all of these populations interact with each other, so our experimental framework provides a general tool that can be adapted to study beliefs in most contexts.

Chapter 2: Gendered Beliefs and the Job Application Decision: Evidence from a Large-Scale Field and Lab Experiment

2.1 Introduction

One of the most important factors contributing to the gender wage gap is that women and men work in different occupations and industries (Blau and Kahn, 2017). This sorting by gender behavior is at least partially driven by workers' job application decisions, yet we know little about why women and men apply to different jobs.[1] I propose that workers' beliefs about hiring managers' beliefs—their *second-order beliefs*—may be an important factor contributing to workers' job application decisions. If workers believe that managers believe one demographic group is more productive in a job than another group, then workers will likely sort based on that demographic characteristic. Through anticipation of discrimination, historically persistent beliefs about productivity differences across groups—*first-order beliefs*—could have reverberating consequences through workers' second-order beliefs.

To determine how second-order beliefs affect workers' job application decisions, I combine a large-scale natural field experiment with a structured online lab experiment. In the field experiment, I solicit applications for a *real* high-paying white-collar job using ads that vary the relevant second-order belief. Then, I invite participants in the field experiment to an online lab experiment that elicits their beliefs using a carefully designed incentive-compatible procedure. By combining the observation of natural behavior in the high-stakes labor market with the rigor of belief elicitation in structured lab experiments, I identify precisely the primitives affecting workers' real-world job application decisions.

I vary the relevant second-order beliefs in the job ad by changing the hiring manager types that workers face. Manager type is characterized along two dimensions: manager gender and the gender associations of the product sector (which I refer to as *sector gender* throughout the paper). This variation captures two channels through which second-order beliefs could drive differential sorting behavior by gender.

First, workers may choose to apply for jobs based on the expected gender composition of management. Workers who believe that female managers' beliefs favor women and male managers' beliefs favor men may sort into a gender-matching job because they are more likely to face a same-gender manager, whom they believe is more likely to favor the worker's own gender.[2] By explicitly varying the gender of the hiring manager in the job ads I use

[1] See Barbulescu and Bidwell (2013); Fernandez and Campero (2017); Fernandez and Friedrich (2011); Fernandez and Mors (2008); Flory et al. (2015) for studies that show women and men make different job application decisions.

[2] I refer to managers as female/male and workers as women/men throughout the paper. The lab exper-

to solicit workers' applications, I isolate how workers' beliefs about the beliefs of different gender managers affects their job application decision separately from their expectations about the gender composition of management.

Second, workers may make their job application decisions based on job-specific characteristics that affect managers' beliefs, and hence workers' beliefs about those beliefs. For example, if workers believe that managers of any gender believe women are better at teaching than men, and men are better at construction work, they may sort by gender into these occupations regardless of the gender composition of management. Alternatively, workers may believe that managers for different jobs hold differing beliefs about the relative productivity of women and men based on manager selection or learning. Workers could believe that men who become managers in men-dominated jobs have systematically different beliefs from men who become managers in women-dominated jobs, and likewise for women, or that beliefs are learned by managers through experience in their occupation or industry. In addition to manager gender, I vary sector gender to capture all of these potential differences in managers that could drive sorting behavior.

The experiment focuses on a single male-dominated occupation: an outside business to business (B2B) sales position at a regional wholesale distributor of disposable goods.[3] The job is highly-compensated with commissions that typically lead to six-figure incomes and generous benefits. The firm I partner with is medium-sized with several hundred employees located in seven metropolitan statistical areas (MSAs) of the U.S. Importantly, the firm distributes several categories of disposable goods, including janitorial and office products. I exploit the firms' multiple product categories in the experiment design by asking workers to choose between applying for a job in the janitorial product sector, which is traditionally male, or in the office product sector, which is traditionally more female.

The field experiment proceeds in two stages. In the first stage, I solicit the contact information of workers interested in, or with experience in, the sales occupation through ads on job boards that do not reveal the manager's gender or product sector. This first stage allows me to observe both applicants and non-applicants, as well as elicit the beliefs of both groups in the later lab experiment. In the second stage, I treat workers with an advertisement for an outside B2B sales position. Workers may apply either to the office product sector or the janitorial product sector. I vary the gender of the hiring manager in each sector and, in doing so, vary the relevant second-order beliefs.

One week after workers receive the job advertisement, I send them an invitation to

iment describes managers as female/male. The intent of the language is to indicate gender identity, not biological gender.

[3]The 2018 American Community Survey (ACS) finds that 29% of wholesale and manufacturing sales representatives are women.

participate online in a paid economics study with Vanderbilt University. The online study is structured like a traditional lab experiment, in which I adapt the procedure I developed in Dustan et al. (2020) to measure beliefs (and beliefs about beliefs) about the differences between two populations in some measurable characteristic. I design the procedure to target the beliefs relevant to the job application decisions the workers made in the field experiment.

The procedure elicits beliefs using the Binarized Scoring Rule (Hossain and Okui, 2013) to incentivize truthful revelation from all expected utility maximizers, regardless of risk preferences, as well as some non-expected utility maximizers. First, I elicit participants' first-order beliefs about the difference between women's and men's sales ability.[4] Then, I inform participants that managers responsible for hiring and supervising sales representatives reported their first-order belief in an earlier experiment. Participants then report what they believe those managers reported, which are the participants' second-order beliefs. I elicit participants' second-order beliefs by manager gender and sector gender, for a total of four second-order beliefs per participant, corresponding to each type of manager participants could have faced in the field experiment.

The main result is that men's, but not women's, second-order beliefs affect their job application decision. I obtain the result by combining the field experiment data on application behavior and the lab experiment data on beliefs for the subsample of job-seekers in the field experiment who participated in the lab experiment. Using the random assignment of the first manager type listed in the job ad, I show that men are more likely to apply for a job with a manager type whom they believe has beliefs that favor men more. The effect size is large, with a one standard deviation increase in the favorability of beliefs increasing the probability a man applies by 5.5 percentage points on a base of 17.9%. I find no such effect for women in the male-dominated occupation for which I recruit. Since women in the experiment are already selected on interest in a male-dominated occupation, I suggest that this selection could be partly based on unresponsiveness to anticipated discrimination.

In addition to the main result, I find independently policy-relevant results in the component experiments. I test three pre-specified hypotheses in the field experiment. First, men are about twice as likely to apply to the job as women. Second, workers do not systematically sort based on sector gender. Lastly, women are unresponsive to manager gender, but men are significantly more likely to apply to *female* managers.

In the lab component of the experiment, I find that participants' beliefs about managers' beliefs are consistent with sorting by gender behavior. Most participants believe that female managers in traditionally female sectors favor women and that male managers in traditionally

[4] Specifically, I ask participants who they believe performed better on a sales ability assessment used by employers to screen workers—a randomly selected woman or a randomly selected man—and by how much.

male sectors favor men. Beliefs about managers whose gender does not match the sector gender—female managers in male sectors and male managers in female sectors—are less homogeneous, suggesting that policies like quotas and affirmative action may have mitigating effects on sorting by gender behavior. Importantly, I find little difference in women's and men's beliefs.

This paper makes three contributions to the existing literature. I present the first direct evidence that second-order beliefs affect labor market behavior. While beliefs about other people's beliefs may play a role in mechanisms studied in other papers, none to my knowledge explicitly study second-order beliefs. Second, the lab experiment provides the first measurement of second-order beliefs outside of an undergraduate student population. Second-order beliefs have typically been studied in a game context using lab experiments on students and focus on beliefs about beliefs regarding strategic actions, not beliefs about beliefs regarding differences in populations. Lastly, the field experiment produces new evidence on differential job application decisions by gender and demonstrate that workers' beliefs about managers' beliefs are an important mechanism driving these job application decisions.

The second-order beliefs I study were first measured by Dustan et al. (2020). In that paper, we develop a methodological framework for eliciting first- and second-order beliefs about the differences between two populations in some measurable characteristic in the lab. We find that a sample of undergraduate students *believe* that women and men hold different beliefs, despite observing no statistically distinguishable differences in the beliefs of women and men. Prior work measuring second-order beliefs elicited beliefs about the beliefs of another player regarding game play strategies. For example, Manski and Neri (2013) elicit second-order beliefs about actions in a simple game to study consistency between actions and beliefs and Kneeland (2015) uses a ring-network game to establish second- and higher-order levels of rationality. In this paper, I investigate whether second-order beliefs about the differences between populations are relevant to real-world decisions.

The closest paper to the field experiment component of this study is Flory et al. (2015), which uses a natural field experiment to study how women's and men's job application rates respond to competitive compensation schemes. They also vary the gender associations of the job and, in a follow-up to the main experiment, consider whether expectations of discrimination drive their results by varying manager gender. Unlike Flory et al., I find no effect of sector gender on application decisions. I also detect an effect of manager gender on men, while they find none. More importantly, I build on Flory et al. by combining the natural field experiment with a lab experiment eliciting beliefs to precisely identify the role of second-order beliefs in workers' job application decisions.

While second-order beliefs have not been explicitly studied in the labor market, theoretical and empirical work exists that is suggestive of their importance. Expectations of discrimination, as Flory et al. (2015) describe it, or anticipation of discrimination is a mechanism that potentially operates on workers' second-order beliefs. Alston (2019) shows that participants anticipate discrimination in a lab experiment in which they can pay to either reveal or hide their gender on a job application for a stereotypically male task. This anticipation of discrimination could be driven by participants who believe that managers in the synthetic labor market believe that women are less productive at the stereotypically male task; however, other explanations are possible. Workers could also anticipate discrimination based on the manager's tastes, which may be unrelated to the worker's beliefs about the manager's beliefs regarding the worker's productivity.

Similar mechanisms operate in equilibrium models of statistical discrimination, where beliefs become a self-fulfilling prophecy (see Fang and Moro, 2010 for a review). In a recent empirical paper, Glover et al. (2017) show that minority workers in a French grocery store chain exert less effort under more biased managers, as measured by the Implicit Association Test (IAT). The self-fulfilling prophecy Glover et al. uncover could be based on workers' beliefs about the beliefs of managers with respect to worker productivity; however, again, other explanations are possible. Only measurements of belief can elucidate their relevance, but second-order beliefs have never been directly measured using an incentive-compatible mechanism. The beliefs I elicit demonstrate that workers' primitives are in line with the self-fulfilling prophecy mechanism.

Numerous studies have shown that workers sort by gender at the job application decision. When applying for jobs, women sort into lower-paying jobs (Barbulescu and Bidwell, 2013), lower-level job queues (Fernandez and Campero, 2017; Fernandez and Mors, 2008), more stereotypically female job titles (Fernandez and Friedrich, 2011), and less competitive compensation schemes (Flory et al., 2015). My results suggest that this sorting behavior could be due to workers' second-order beliefs.

Moreover, sorting by gender behavior is an important factor contributing to the gender wage gap. Large parts of the gender wage gap in numerous countries can be attributed to women and men working in different firms, occupations, and industries.[5] This study suggests that workers' beliefs about managers' beliefs in these different jobs may contribute to this sorting behavior underlying a large part of gender wage gaps.

The paper proceeds as follows. I describe the experiment design and implementation in Section 2.2. Then, I present my main results on workers' beliefs and the job application

[5]See Bayard et al. (2003); Blau and Kahn (2017); Carrington and Troske (1998); Groshen (1991) for evidence in the U.S., Card et al. (2016); Cardoso et al. (2016) for Portugal, Amuedo-Dorantes and De la Rica (2006) for Spain, Fortin and Huberman (2002) for Canada, Gallen et al. (2019) for Denmark, and Fafchamps et al. (2009) for Africa.

decision in Section 2.3. I present additional results from the component field experiment in Section 2.4 and the component lab experiment in Section 2.5. I conclude with a discussion of the results in Section 2.6.

2.2 Experiment Design

In this section, I detail the experiment design and implementation for determining the causal effect of workers' second-order beliefs on their job application decisions. First, I describe the field experiment in subsection 2.2.1. I then turn to the lab experiment in subsection 2.2.2, including a careful specification of the beliefs of interest. Lastly, I discuss implementation in subsection 2.2.3 and present descriptive statistics for the full sample from the field experiment and the subsample who completed the lab experiment.

2.2.1 Field Experiment

The field experiment proceeds in two stages. In the first stage, I solicit the contact information of workers with experience or interest in the sales occupation. In the second stage, I treat those workers with a job advertisement that varies the gender of the manager workers face in each sector and observe their job application decision.

The job for which I solicit applications is an outside business to business (B2B) sales representative position with a medium-sized regional wholesale distributor of multiple product sectors. The job is highly compensated, advertised as six-figure income potential, and includes generous benefits like health insurance, employer-paid life insurance, vacation days, and a 401(k) with matching contributions. The partner firm has an ongoing need for sales representatives and hires on a rolling basis. I solicit job applications from workers in all seven metropolitan statistical areas (MSAs) in which the firm is located.

This job has a number of characteristics that facilitate the study of workers' beliefs on the job application decision. First, an "outside" sales representative spends most of their time away from the office, as suggested by the title. This factor decreases the amount of interaction with the manager, minimizing the role of preferences associated with in-person exchanges (e.g. sexual harassment). Second, compensation is self-determined through commissions.[6] This factor decreases the relevance of the manager's role in determining wages, increasing the relative weight of the manager's role in the hiring decision. Both of these factors contribute to isolating the probability of being hired as the determining factor in the worker's job application decision.

[6] Initially, workers are paid a salary in a type of "trial period"; however, successful workers are transitioned to commission-based pay.

In addition, the flexibility in hours and place of work of the outside sales position increases the appeal of this job to women workers (Goldin, 2014; Wiswall and Zafar, 2018). Since wholesale sales is a male-dominated occupation, women make up a smaller proportion of the population of workers interested in this job. The flexibility factor increases my ability to recruit women workers and is noted in the job advertisement to augment its salience.

Lastly, the multiple product sectors in which the firm operates enable the study of gendered product sectors. I can vary the product sector in which the worker would potentially be employed without varying any firm-specific characteristics including: firm size, location, culture, and gender composition. The two product sectors I advertise are the office products sector, which I characterize to workers as traditionally female, and the janitorial products sector, which I characterize as traditionally male, though workers familiar with the sectors will already be aware of their gendered nature.

Stage 1: Create Pool of Workers

In the first stage of the field experiment, I solicit the contact information of workers with experience or interest in the sales occupation in order to develop a pool of workers. By establishing a pool of workers, I can observe the behavior of both applicants and non-applicants, as well as elicit the beliefs of both groups in the lab experiment. I solicit the contact information of workers in two ways. First, I place a generic advertisement on the job boards ZipRecruiter and Monster, as well as send the same generic advertisement through "Sponsored InMail" on LinkedIn. The generic advertisement does not include the firm's name or information about the relevant product sectors. Rather, it provides a list of job responsibilities, general compensation information,[7] and basic qualifications.[8] Workers are encouraged to respond to the ad to learn specifics about the company and a full job description. Appendix E contains the generic job advertisement.

Workers respond to the generic advertisement by clicking "Apply" on ZipRecruiter or Monster, or clicking "Learn More" on LinkedIn. The apply button on ZipRecruiter and Monster pre-fills workers' contact information and allows workers the option of attaching a resume, though a resume is not required. The button on LinkedIn takes workers to the website of the firm set up to manage the recruiting process, where they fill out a form with their contact information.

The second way in which I solicit contact information is through the resume databases on ZipRecruiter and Monster. Workers may choose to include their resume in the database

[7] No exact figures are provided for compensation. The wording describing the compensation was prescribed by the firm.

[8] A bachelor's degree, with experience in outsides sales preferred.

to make it searchable by employers. From these resumes, I obtain the contact information of workers to include them in the experiment. I download the resumes of workers within 50 miles of one of the seven MSAs in which the firm is located who have been active on the website in the last 10 days, and who either 1) have a history of working in a sales position, or 2) express interest in a sales position through their profile on the job board or their objective statement on the resume. These workers, along with those who provide their contact information directly as described above, constitute the "full sample" of workers.

Stage 2: Invite workers to apply for a job

In the second stage of the field experiment, I email the full sample of workers the complete job advertisement. Both the office product sector and the janitorial product sector are listed in the ad, with a different hiring manager associated with each sector. Workers who choose to apply must also choose a product sector.

Treatment varies the gender of the manager in each product sector. The manager may be either male or female and may be assigned to either the male or female sector. The combination of manager gender and sector gender characterize the four manager types, listed in Table 2.1. Workers never choose between a male and a female manager in the same sector because it would be deviate from job advertisement norms to ask workers to choose between two hiring managers for the same job.

Table 2.1: Manager Types

Type	Manager Gender	Product Sector
$\{m, j\}$	Male	Janitorial (M)
$\{f, j\}$	Female	Janitorial (M)
$\{m, o\}$	Male	Office (F)
$\{f, o\}$	Female	Office (F)

Since workers always face one manager in the office sector and one in the janitorial sector, there are four treatments, listed in Table 2.2. Treatment 1 has a male manager in both sectors. Treatment 2 has a female manager in both treatments. Treatment 3 has a "sector-matching manager" in each sector, meaning that the gender of the manager matches the gender of the sector in both sectors. Workers in this treatment face a female manager in the office sector and a male manager in the janitorial sector. Lastly, treatment 4 has a "non-sector-matching manager" in each sector, meaning the gender of the manager does *not* match the gender of the sector in both sectors. Workers in this treatment face a female manager in the janitorial sector and a male manager in the office sector.

Table 2.2: Treatments in the Field Experiment

Treatment	Office Sector (F) Manager Gender	Janitorial Sector (M) Manager Gender
T1: All Male Managers	Male $\{m, o\}$	Male $\{m, j\}$
T2: All Female Managers	Female $\{m, o\}$	Female $\{f, j\}$
T3: Sector-Matching Managers	Female $\{f, o\}$	Male $\{m, j\}$
T4: Non-Sector-Matching Managers	Male $\{f, o\}$	Female $\{f, j\}$

The job advertisement is delivered from a "do not reply" e-mail address at the recruiting firm. The ad first either 1) thanks them for their interest if they provided their contact information or 2) tells them where their resume was found and states that they may be a good match for the position. Next, the ad provides detailed job and firm information. Last, the ad lists the two product sectors in bold with information specific to each sector and the name and e-mail address of the hiring manager listed under each sector heading. The order the sectors appear in the e-mail is randomized within treatment groups. Workers are informed that they may submit application materials to only one hiring manager. Appendix F contains the treatment e-mail.

All treatment information is included in the product sector section of the job advertisement. First, I signal the gender associations of each product sector by describing the job titles for the types of customers with whom the prospective worker should be comfortable engaging. The ad describes the typical customers in the janitorial sector as "facility managers, operations, and maintenance workers."[9] Similarly, the ad describes the typical customers in the office sector as "office managers, receptionists, and administrative workers."[10] The job titles of typical customers are based on my previous experience in these sectors and were also validated by the sales managers at the firm.

I signal the gender of the manager in each product sector through the manager's first name. The top two female and male names from 1975, along with the top four last names, in the U.S. are used as pseudonyms for real hiring managers at the firm. In addition to naming the manager to whom workers should send their application, the e-mail address of the manager is in the format firstname.lastname@[*recruiting firm name*].com to ensure salience of manager gender as signaled by the manager's first name.

Workers submit a job application in two parts. First, they e-mail their resume and (optional) cover letter to the manager of their choice. Second, they complete a job application

[9]The 2018 Current Population Survey (CPS) finds that 96% of maintenance and repair workers, 68% of general and operations managers, and 66% of janitors and building workers are men.

[10]The 2018 CPS finds that 94% of secretaries and administrative assistants, 72% of first-line supervisors of office and administrative support, and 71% of office and administrative support workers are women.

form on the recruiting firm's website with education information and a brief job history. As a part of the application form, workers must provide the e-mail address of the manager to whom the application should be routed. In this way, participants must confirm in two places the manager type to which they wish to apply. Again, this process ensures the salience of manager type by forcing participants to look for and either write or copy and paste the manager's name from the job advertisement. Workers who submit application materials to more than one manager are informed that their application materials cannot be processed and advised to resubmit their application materials to only one manager.

One week after the worker receives the treatment e-mail, I send an e-mail recruiting them to an online lab experiment with Vanderbilt University. I carefully ensure there are no similarities between the treatment e-mail and the recruitment e-mail so that workers do not know that there is any connection between the field and lab experiment. Such a connection might bias the lab results since monetary compensation in a experiment cannot compete with the high-stakes of a job application. In other words, if the worker believes that there is a non-zero (even if trivial) chance that their potential employer sees their responses on the lab experiment, the worker would report what they believe the employer wants rather than their true beliefs. Rather than assuring workers (truthfully) that the employer does not see responses to the lab experiment, or even if a worker completes the lab experiment, I avoid any connection between the field and lab portions of the experiment. Appendix G contains the recruitment e-mail.

2.2.2 Lab Experiment

The goal of the lab experiment is to elicit workers' second-order beliefs about the first-order beliefs of managers they face in the labor market regarding the relative productivity of women and men workers. Before discussing *how* I elicit those beliefs, I must carefully define *what* these beliefs of interest are. Incorrect or imprecise specifications of beliefs can lead to the elicitation of beliefs that do not accurately answer the research question of interest. For example, Dustan et al. (2020) shows that eliciting the difference in the beliefs about medians of two populations (as opposed to what I elicit: the belief about the median of the differences) could lead to the conclusion that people believe there are no differences in those populations, when in fact people believe there are large differences.

Let X^w be the distribution of sales productivity in the population of women workers and X^m be the distribution of sales productivity in the population of men workers. I want to elicit whether workers believe that a manager believes that $Pr(X^w > X^m) \geq 0.5$ or that $Pr(X^w < X^m) \geq 0.5$ (or both). Now, let $X^d = X^m - X^w$. Then, I want to learn whether workers believe that a manager believes that $Pr(X^d > 0) \geq 0.5$ or that $Pr(X^d < 0) \geq 0.5$. This property can be inferred from the median of X^d, which is the median of the distribution of differences in women's and men's sales productivity.

The idea that this property captures is intuitive. If a manager must choose between an equally qualified (on paper) woman and man, does the manager believe it is more likely that the woman or the man will be more productive in the sales job? The worker's belief about the manager's belief about this property then reveals whether the worker believes they are more or less likely to be hired, given the same resume as a worker of the opposite gender.

To elicit this property of workers' subjective belief distributions, I adapt the procedure developed in Dustan et al. (2020) to measure first- and second-order beliefs about the difference in some measurable characteristic between two populations. To begin, the procedure requires two auxiliary samples: one to incentivize the first-order belief elicitation and one to incentivize the second-order belief elicitation. A small sample measuring productivity in sales is needed to elicit managers' beliefs about the relative productivity of women and men. A small sample measuring managers' first-order beliefs about relative productivity is needed to elicit workers' beliefs about managers' beliefs. These samples are used only to incentivize participants to truthfully reveal their beliefs and are not part of the results of this paper.

Auxiliary Sample Measuring Sales Productivity

The measurable characteristic of interest in this experiment is the productivity of workers in a sales job. I proxy this productivity with performance on a sales ability assessment used by employers to screen workers. Sales assessments like the one I use are generally comprised of psychological questions in which the worker is asked to respond to a series of hypothetical situations. I obtained a small sample of workers who completed the assessment as a part of a job application from a firm that has tested more than 2 million workers.[11]

Workers who choose to participate in the online lab experiment, henceforth "participants", are told that the "sales assessment is used by employers to evaluate job candidates for positions in outside sales" and given information about the types of characteristics that are evaluated (e.g. "controls emotions and handles rejection").[12] The metric I use from the sales assessment is the worker's percentile rank in the population of test-takers.[13] This mea-

[11] The firm provided the results of the sales assessment for two anonymous women and two anonymous men. For a discussion of why only one draw from a population is necessary to incentivize truth-telling without deception, see Dustan et al. (2020).

[12] I do not provide participants with the assessment itself or even sample questions. My goal is to learn about beliefs regarding productivity in sales, not beliefs about how women and men perform on a particular assessment.

[13] Note that these are not any workers from the field experiment.

sure has the benefit of being unit-free, so that the results can be interpreted without the potentially confounding issue of participants' beliefs about the support of the measure.

Auxiliary Sample Measuring Managers' First-Order Beliefs

The second sample measures the first-order beliefs of managers. Since I want to elicit the second-order beliefs directly relevant to a worker's job application decision, I further refine this population to "managers responsible for hiring and supervising outside sales representatives." I collected a small sample of first-order beliefs from managers at the partner firm using the elicitation procedure I describe next.[14] The only difference is that I do not elicit managers' second-order beliefs, just their first-order beliefs.[15]

Eliciting Participants' First- and Second-Order Beliefs

The belief elicitation proceeds as follows. I elicit first-order beliefs by asking participants to report who they believe did better on the sales ability assessment—a randomly selected man or a randomly selected woman—and by how much. After the first-order belief elicitation, participants learn that managers responsible for hiring and supervising outside sales representatives answered the same question they just did, as described in Section 2.2.2. Participants are then asked to report what they believe a randomly drawn manager from each manager-type population chose, which are their second-order beliefs. Manager types are described as "(gender) manager in a traditionally (gender) product sector." The first-order belief elicitation is incentivized using the sample measuring worker productivity, described in Section 2.2.2, and the second-order belief elicitation is incentivized using the sample measuring managers' first-order beliefs, described in Section 2.2.2.

Participants report their first-order belief using the slider in Figure 2.1 and report their second-order beliefs using the slider in Figure 2.2. The slider always begins in the center of the support, at gender neutrality. At this point on the slider, the text on the title bar with the grey background reads "I believe that they did the same" for the first-order belief slider and "I believe that the manager guessed they did the same" for the second-order beliefs slider. Participants move the slider to the left to report that (the manager guessed) the woman outperformed the man and move the slider to the right to report that the man outperformed the woman on the sales assessment. The text on the title bar of the slider updates at each point on the slider so that participants can see their guess in words.

[14] I sampled two women and two men.

[15] I cannot elicit managers' second-order beliefs at this point because I do not have a sample of first-order beliefs to incentivize truth-telling.

Figure 2.1: First-Order Belief Slider

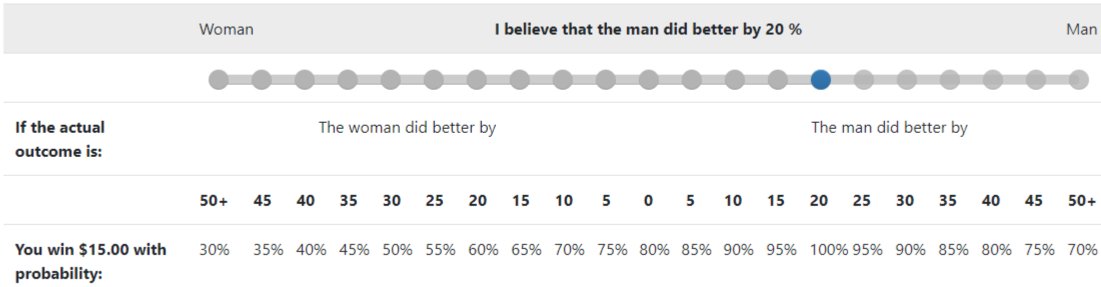
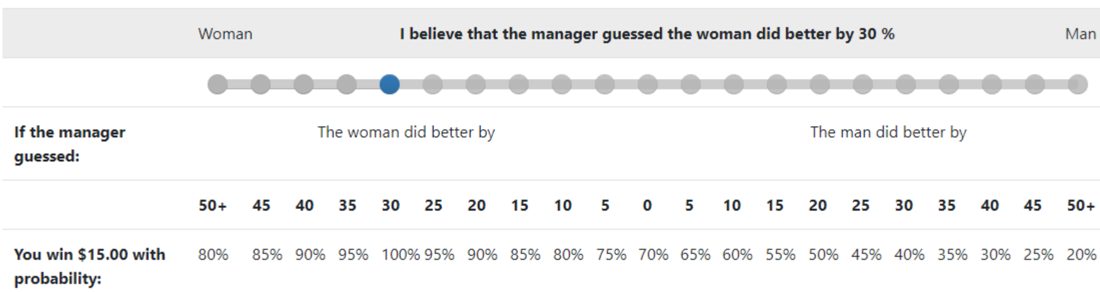


Figure 2.2: Second-Order Belief Slider



As the participant moves the slider, the last row reporting the probabilities for winning the \$15 prize also update according to the Binarized Scoring Rule (BSR).[16] The defining feature of the BSR is that participants maximize their *probability* of winning a prize, rather than maximizing the *size* of the prize. This change in objective function means that the BSR is incentive-compatible for all expected utility maximizers as well as some non-expected utility maximizers. The critical assumption is *stochastic monotonicity*, meaning that the participant must prefer higher probabilities of winning the prize to lower probabilities, which covers a wide range of preferences.

The BSR works as follows. Let x be a draw of the random variable of interest X (in my case, the difference in percentile rank on the sales ability assessment between a randomly drawn man and a randomly drawn woman), θ the participant's guess about x, and $\overline{K}$ the number of points in the support of the distribution of X. Then, the participant's probability of winning some prize is

$$\Pr(\text{Win the Prize}) = 1 - \frac{l(x,\theta)}{\overline{K}} \tag{2.1}$$

[16]It remains the state-of-the-art at present despite recent concern that it may be subject to a "pull-to-center" effect (Danz et al., 2020), meaning that participants report a belief somewhere between their true belief and the center of the support.

where $l(\cdot)$ is a loss function that maps the participant's error to a value. The form of the loss function determines what *function* of the distribution of X is elicited by the payment rule. The prize in this experiment is $15.

I choose the loss function $l(\cdot)$ to elicit the median. The median of the distribution of differences between women and men is the lowest cost, in terms of time and cognitive load, function to elicit the property $Pr(X^d > 0) \geq 0.5$ or $Pr(X^d < 0) \geq 0.5$. For example, eliciting the mode of the ternary distribution with outcomes "woman outperforms man", "man outperforms woman", and "woman and man perform the same" would require twice the number of elicitations and provide less relevant information. Eliciting the probabilities $Pr(X^d > 0) \geq 0.5$ and $Pr(X^d < 0) \geq 0.5$ directly would similarly require twice the number of elicitations to obtain the same information about the relative productivity of women and men.[17] See Dustan et al. (2020) for a detailed discussion of the appropriate function to elicit to learn about beliefs about differences between two populations. The loss function for the median is the absolute error term, so the participant's probability of winning the $15 prize in this experiment is determined by:

$$\Pr(\text{Win the Prize}) = 1 - \frac{|x - \theta|}{\overline{K}} \tag{2.2}$$

The last factor to determine the payment structure is the number of points in the support over which participants can report their belief, $\overline{K}$. The choice of $\overline{K}$ affects the incentives for the participant to be precise—the larger and denser the support, the lower the marginal cost of being one unit away from the outcome of the random draw. At the same time, small or sparse supports could limit participants' ability to express extreme beliefs or their ability to be precise if outcomes are binned. I choose $\overline{K} = 21$ to balance these factors, which corresponds to ten points of support on either side of gender neutrality.

The interactive slider communicates all information about the payment rule, without requiring the participant to understand the BSR or the equation that determines the payment probabilities.[18] When the participant chooses a point on the slider, the last row of the chart updates to reflect the probability of winning the $15 prize based on their selection, for *each possible outcome of the random draw*.[19] The chart can be read as "If the actual outcome of the

[17] Though eliciting the probabilities directly requires twice the number of elicitations, it also yields additional information that elicitation of the median does not. This additional information is not sufficiently useful, however, to warrant the higher costs of elicitation.

[18] Part of the innovation of the procedure developed by Dustan et al. (2020) was this method of implementing the BSR. The typical implementation taught participants the BSR or the underlying equation (see for example Hossain and Okui, 2013, Babcock et al., 2017, or Dianat et al., 2018).

[19] Recall that the random draw is of a woman and a man who completed the sales assessment for the

random draw is: (the woman did better by/the man did better by) x percentile points, you win \$15 with probability y" where x is the number in the first row and y is the corresponding number in the last row. For example, in Figure 2.1, the participant has a 100% chance of winning the \$15 prize if the outcome of the random draw is exactly their guess that the man did better by 20 percentile points, a 95% chance of winning if the outcome is that the man did better by 15 percentile points, etc.

2.2.3 Implementation

The field experiment was implemented in two waves. The first wave ran from February to March 2020 and was prematurely halted due to the firm's temporary hiring freeze in response to COVID-19. Recruiting for the online lab experiment continued through May 2020. The second wave ran from August to September 2020. Recruiting for the online lab experiment continued through October 2020.

Workers were recruited to the lab experiment through an e-mail invitation from Vanderbilt University to participate in a paid study. The study was advertised as a 10 to 15 minute online survey that paid \$5 for completion. Participants were told they could earn another \$15 based on their responses, as well as earn tickets into a drawing for a \$500 Amazon gift card. To begin the experiment, participants clicked on a link in the e-mail. The experiment was programmed using the Otree software (Chen et al., 2016).

Instructions were both written and recorded as audio. Participants had to press "Play" and listen to the audio track on each page before they could proceed to the next page, or stay on the page for the length of the audio file plus 25 seconds. After the initial instructions, participants completed an example with feedback to ensure they understood how to report their beliefs and how they would be paid.

Participants began by reporting their first-order belief, followed by the four second-order belief elicitations with respect to each manager type—female in female sector, female in male sector, male in female sector, and male in male sector.[20] Then, participants completed a number of other short tasks. These included, in order: unincentivized recall questions on the gender and sector of the last manager type for which they reported their second-order belief; incentivized beliefs about managers' beliefs about customer preferences; incentivized beliefs about the traditional gender associations of a number of product sectors (including office and janitorial products); incentivized risk preference elicitation; unincentivized manager type preference elicitations; unincentivized confidence in own sales ability; and a number of survey questions covering employment status, demographics, and self-reported confidence in

first-order belief, and of a manager of the relevant type for the second-order belief.

[20]The four second-order beliefs are elicited in random order.

beliefs and risk preferences. All incentivized tasks were incentivized using tickets into the drawing for the $500 Amazon gift card, with the exception of risk preferences. I elicited risk preferences using the method developed in Eckel and Grossman (2008), paying participants up to $2.80.

The measure I use from the additional short tasks in the analysis apart from the demographics are the unincentivized manager type preference elicitations. I collect preferences for manager type by asking participants 1) which gender manager they prefer, and 2) which type of product sector they prefer.[21] Participants could choose female, male, or no preference for each.

Table 2.3: Sample Sizes in the Field Experiment, by Treatment

	T1	T2	T3	T4	Total
Wave 1	538	537	535	533	2,143
Women	183	183	182	181	729
Men	336	334	332	331	1,333
Unknown	19	20	21	21	81
Wave 2	680	693	687	679	2,739
Women	268	273	274	262	1,077
Men	389	397	393	394	1,573
Unknown	23	23	20	23	89
Total	1,218	1,230	1,222	1,212	4,882

Notes: Columns reference the four treatment groups as listed in Table 2.2. Rows reference subsamples for women, men, and people of unknown gender.

One reported belief out of the five elicited in total was randomly selected for payment at the end of the survey. Participants were informed of their guess and the random draw from the relevant population it was compared to in order to determine their probability of winning the $15 prize. Participants then generated a random number between 0 and 100 using an embedded random number generator to determine if they won the prize. If their number was equal to or lower than their probability of winning, determined by equation 2.2, they won. After informing them of their total earnings, including tickets into the drawing for the $500 Amazon gift card for their responses on the additional tasks, participants were directed to another site through a link to fill out a receipt. Payment was sent to participants through PayPal within 24 hours. Screenshots of the full experiment are included in Appendix H.

[21] On the same page, I also ask participants about their preferences over workplace gender composition and provide an open comment box for them to make any notes on their preferences. The intention of the open comment box is to allow participants to reduce the cost of revealing gendered preferences by explaining them.

Table 2.4: Comparison of Demographics in Lab Sample and in 2018 American Community Survey

	Experiment	2018 ACS	Difference
Woman	0.392	0.285	0.107***
	(0.020)	(0.017)	(0.026)
White Non-Hispanic	0.677	0.892	-0.215***
	(0.020)	(0.012)	(0.022)
Black/African-American	0.201	0.067	0.133***
	(0.017)	(0.010)	(0.019)
Other Race/Ethnicity	0.105	0.041	0.064***
	(0.013)	(0.008)	(0.014)
Married	0.446	0.693	-0.247***
	(0.021)	(0.018)	(0.027)
Never Married	0.353	0.190	0.163***
	(0.020)	(0.015)	(0.025)
Other Marital Status	0.201	0.117	0.084***
	(0.017)	(0.012)	(0.020)
Age	40.966	47.373	-6.407***
	(0.570)	(0.561)	(0.806)
Number of Children	1.135	0.800	0.335***
	(0.059)	(0.040)	(0.069)
Number of Children <5yo	0.142	0.142	0.000
	(0.020)	(0.016)	(0.026)
Wage Income			
$0-$25,000	0.171	0.244	-0.074***
	(0.016)	(0.016)	(0.023)
$25,000-$50,000	0.231	0.175	0.055**
	(0.018)	(0.015)	(0.023)
$50,000-$75,000	0.258	0.175	0.082***
	(0.018)	(0.015)	(0.023)
$75,000-$100,000	0.167	0.120	0.047**
	(0.016)	(0.012)	(0.020)
$100,000-$150,000	0.091	0.167	-0.076***
	(0.012)	(0.014)	(0.019)
$150,000+	0.083	0.118	-0.035**
	(0.012)	(0.012)	(0.017)
Observations	598	684	1282

Notes: Columns reference the lab experiment sample, the 2018 American Community Survey sample of wholesale or manufacturing sales representatives in one of the seven MSAs I recruited from in the experiment, and the differences in those two samples. Standard errors are in parentheses.

A total of 4,882 workers were treated in the field experiment, with 598 of those workers choosing to participate in the lab experiment. Table 2.3 reports the number of participants in each treatment of the field experiment by wave and by gender. Workers do not report their gender in the field experiment, so gender was predicted using the first names of all children born in the U.S. from 1964 to 1998. If more than 80% of children born with a given name were female, the worker was classified as a woman. If fewer than 20% of children born with a given name were female, the worker was classified as a man. 93% of workers could be classified using these rules. Another 3% of workers were classified using their photo,[22] middle name,[23] or other clear marker of gender on their resume.[24] Workers who participated in the lab experiment self-reported their gender. Approximately 1% of workers in the field experiment were classified using this self-reported gender. The remaining 3% of workers were unable to be classified and dropped from the analysis.[25]

Since I do not collect demographic information from participants in the field experiment, I cannot compare their characteristics to the subsample who participated in the lab experiment. Instead, I compare the demographic characteristics of participants in the lab experiment to those of wholesale and manufacturing sales representatives in the 2018 American Community Survey in the seven metropolitan statistical areas (MSAs) I recruit from in Table 2.4. My sample has more women, racial diversity, people who have never been married, and younger people with slightly more children.

2.3 Worker Beliefs and the Job Application Decision

This section presents the main result of the paper, which I obtain from combining the application choice data from the field experiment with the data on beliefs from the lab experiment. The question of interest is: how do workers' beliefs about managers' beliefs affect their job application decision? To answer it, I test the following hypothesis:

Hypothesis 1. *More favorable second-order beliefs about a manager type increase the likelihood that a worker applies for a job with that manager type.*

The specification to test this hypothesis is informed by a behavioral model of job choice I outline below, and by an empirical pattern that emerges from the field experiment: 71.10% ($SE = 2.62\%$) of workers who apply choose the first manager listed in the job. This proportion is significantly different from 50% ($p = 0.000$). Figure 2.3 shows that this proportion is consistent across treatment groups and gender. There are no significant differences between treatment groups and the proportions are significantly different from 50% in all treatments ($p = 0.000$ for treatments 1, 2, and 4, and $p = 0.007$ for treatment 3). There is also no

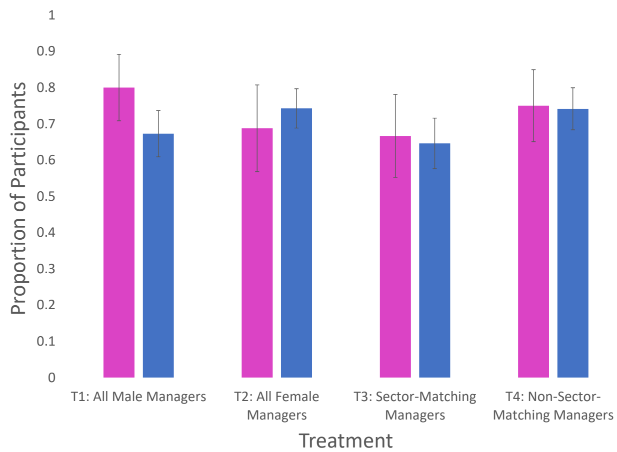
[22] Photos could be found on resumes, ZipRecruiter profiles, and LinkedIn profiles, where the link to the worker's LinkedIn profile was included on their resume.

[23] Using the same classification strategy described for first names.

[24] Examples include references to a men's basketball scholarship and an award for "Top 10 Female in Sales."

[25] Self-reported gender in the lab replaced predicted gender in the field for 11 participants whose self-reported gender in the lab experiment contradicted predicted gender from the birth certificate algorithm.

Figure 2.3: Proportion of Workers who Choose the First Manager Listed, Conditional on Applying



Notes: From left to right on each graph, the bars show the proportion of workers who applied to the first manager type listed, conditional on applying for the job, in treatment 1 with a male manager in both sectors, treatment 2 with a female manager in both sectors, treatment 3 with a sector-matching manager in both sectors, and treatment 4 with a non-sector-matching manager in both sectors. Standard error bars are robust.

significant difference ($p = 0.682$) in the probability a woman worker chooses the first manager (73%, $SE = 5.2\%$) and the probability a man worker chooses the first manager (70%, $SE = 3.0\%$).

Workers' application behavior underscores the importance of the order of options that has been recognized in a variety of fields, including psychology, marketing, and political science.[26] For example, a "primacy effect" favoring the first option has been identified on election ballots (Van Erkel and Thijssen, 2016), lists of website links (Murphy et al., 2006), and in wine tastings (Mantonakis et al., 2009). This behavioral phenomenon has been modeled as a framing effect in economics (Salant and Rubinstein, 2008) and studied in the context of revealed preference as "limited attention" (Masatlioglu et al., 2012).

[26]The order of options was randomized in the experiment for this reason.

Based on this result and the literature, I develop the following theoretical framework (its formalization is presented in Appendix I) to inform the empirical specification I use to test my main result. I model workers' job application options in a dynamic setting under the assumption that workers' (second-order) beliefs reflect their expectations about the probability of being hired. For simplicity, I assume workers' beliefs about the hiring manager enter their utility directly; therefore, *ceteris paribus*, workers prefer to apply for a job with a manager type for whom they hold higher beliefs.

I assume that workers observe the first manager type, as defined by the manager's gender and sector, listed in the job ad at no cost. After observing the first manager type, workers choose among the following three options: 1) apply to the first (observed) manager type, 2) pay a (psychic) cost to observe the second manager type (this cost could be motivated by a limited attention model), or 3) proceed to the next time period without applying. If the worker chooses option (2) by paying the cost to observe the second manager type, the worker chooses between applying to one of the two observed manager types or proceeding to the next time period without applying. Workers who proceed to the next time period choose between applying to another randomly drawn manager type or their reservation wage. This model informs the empirical specification detailed next.

2.3.1 Empirical Specification

The dependent variable of interest is the worker's decision to apply for the job $Pr(\text{Apply})$. The model suggests that the worker's second-order belief about the first manager type listed in the job ad b_i^{first} enters into the worker's decision to apply to the first manager type, to observe the second manager type, to apply to the second manager type conditional on observing it, or to proceed to the next time period without applying; therefore, it is the main independent variable of interest.[27] According to Hypothesis 1, the more favorable beliefs the worker has about the beliefs of the first manager, the more likely the worker should be to apply to the job.

In addition to the primary belief of interest, the model suggests that two other measures of beliefs also affect workers' decision. I refer to these beliefs as the "alternative" beliefs since they represent the worker's dynamic decision. The first measure of alternative beliefs captures the possibility that the worker chooses to observe the second manager type in the ad. As discussed above, workers who pay the cost of observing the second manager type choose between applying to one of the two manager types they have observed or not applying. *If* they choose to apply, the worker chooses the manager with the most favorable beliefs among

[27] Since beliefs are reported as man minus woman, I change the sign on the beliefs of women workers.

the two manager types they faced; therefore, the first measure of alternative beliefs is the most favorable belief out of the two facing manager types, which I define as b_i^{max}. That is, $b_i^{\text{max}} = b_i^{\text{first}}$ if the first manager type is also the most favorable manager of the two types the worker faced, but $b_i^{\text{max}} = b_i^{\text{second}}$ otherwise, where b_i^{second} is the worker's belief about the belief of the second manager type.

The second measure of alternative beliefs I include is the average of all second-order beliefs except for the first manager listed b_i^{avg}. This measure controls for the alternatives the worker would face if they do not apply in the experimental period, unconditional on whether they observe the second manager type. A worker who faces a manager with a relatively high belief might still choose not to apply if the alternative manager types have even more favorable beliefs. Similarly, a worker who faces a manager with relatively low beliefs might still apply if the alternative manager types have even less favorable beliefs. In essence, controlling for the average means that the coefficient on the worker's belief about the belief of the first manager captures how more favorable beliefs *relative to* the beliefs of the other manager types affect the worker's application decision.

To causally interpret the relationship between second-order beliefs and the job application decision, I must control for any factor that affects job application behavior *and also varies with the randomly assigned manager types.* Manager gender-specific preferences or sector-specific preferences may be such a factor; therefore, I control for the vector of preferences $\boldsymbol{\rho}_i$ elicited in the lab experiment in the estimating equation. There are four indicator variables corresponding to participants reporting 1) preferences for the gender of the first manager, 2) preferences for the opposite-gender of the first manager, 3) preferences for the sector gender of the first manager, and 4) preferences for the opposite-gender sector of the first manager.

In summary, I estimate the relationship between the second-order belief of interest b_i^{first} and choosing to apply for the job $Pr(\text{Apply})$ separately for women and men using the equation:

$$Pr(\text{Apply}) = \alpha \mathbf{T}_i + \beta_1 b_i^{\text{first}} + \beta_2 b_i^{\text{max}} + \beta_3 b_i^{\text{avg}} + \theta \boldsymbol{\rho}_i + \gamma \mathbf{X}_i + \epsilon_i \tag{2.3}$$

$\mathbf{T}_i$ is a vector of all treatment indicator variables and the vector of controls $\mathbf{X}_i$ includes the worker's recruitment wave, city, and job board, as well as whether or not the worker reports actively searching for a job in the lab experiment. Note there is no constant, so the vector of coefficients on $\mathbf{T}_i$ reports average application rates in each treatment, conditional on the other variables. I estimate a linear probability model in my preferred specifications. Appendix K contains estimates of a Probit model, with the same qualitative results.

The descriptive statistics for each of the main variables are listed in Table 2.5. The probability of applying and the preference parameters are binary variables. As discussed in Section 2.2.2, beliefs are reported about the median of the distribution of differences between women and men in percentile rank on the sales ability assessment. Positive values correspond to beliefs that favor the worker's own gender and negative values correspond to beliefs that favor the opposite-gender of the worker. So, on average, women in the sample believe that the median first manager type listed in their job ad reports that the median difference between women and men is three percentile rank points in favor of women. Men, on average, believe that the median first manager type listed in their job ad reports that the median difference is 1 percentile rank point in favor of women (because the sign is negative).

Table 2.5: Descriptive Statistics

	Women	Men
Pr(Apply)	0.076	0.179
	(0.266)	(0.384)
Belief about First Manager	3.296	-1.254
	(29.054)	(25.554)
Most Favorable Belief of Facing Managers	18.879	15.299
	(22.165)	(20.605)
Average Belief about Managers, except First	-0.045	0.204
	(14.946)	(11.956)
Prefer First Manager's Gender	0.251	0.217
	(0.435)	(0.412)
Prefer Opposite-Gender of First Manager	0.283	0.191
	(0.451)	(0.394)
Prefer First Manager's Sector	0.184	0.197
	(0.388)	(0.398)
Prefer Opposite Sector of First Manager	0.161	0.225
	(0.369)	(0.418)
Observations	223	351

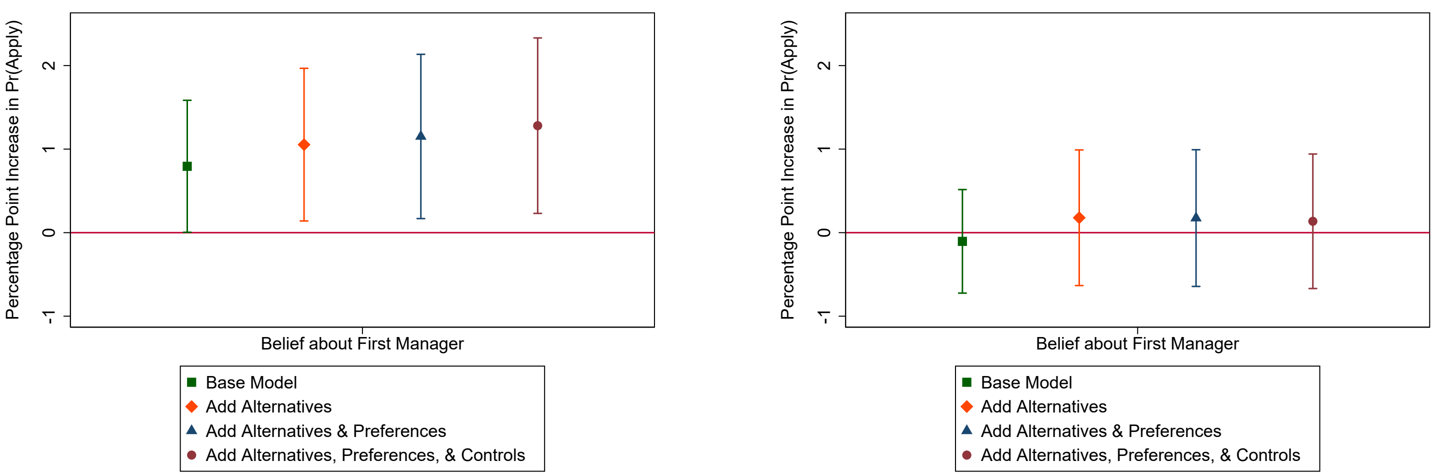
Notes: Columns reference the sample of women and the sample of men who participated in the lab experiment. Standard deviations are in parentheses.

2.3.2 Results

I find that second-order beliefs about the first manager positively and significantly affect the worker's job application decision through the random assignment of the first manager type listed in the job ad for men, but not women. Figure 3.11 reports the coefficients and their 95% confidence intervals in four specifications of equation 2.3, progressively including more explanatory variables. The *Base Model* includes only the belief about the first manager b_i^{first} and the vector of treatment indicator variables. The *Add Alternatives* specification adds $b_i^{\max}$ and b_i^{avg}. The *Add Alternatives & Preferences* specification adds the vector of preference variables $\boldsymbol{\rho}_i$. Lastly, the *Add Alternatives, Preferences, & Controls* specification adds the vector of controls $\mathbf{X}_i$.

Figure 2.4: Regression Coefficient on Belief about First Manager by Gender, Main Specification

(a) Men (b) Women



Notes: 95% confidence intervals are shown. Standard errors are robust. Coefficient is scaled to a one unit increase in the favorability of beliefs. Panel (a) reports the coefficient on the regression with men only and panel (b) reports the same for women only. In each graph, the first estimate includes only the vector of treatment variables and the belief of interest. The second estimate adds measures of alternative beliefs, the third estimate adds preferences, and the fourth estimate includes controls for the wave, city, and job board in which the worker was recruited and whether the worker reports actively searching for a job.

The coefficients on the worker's belief is positive and statistically significant at the 5% level for men in all specifications. In the full specification, a one unit increase in the worker's belief about the belief of the first manager listed in the job ad increases the probability a man applies by 1.28 percentage points ($SE = 0.53$). This effect is approximately 7% of the base application rate of 17.9% for men workers, indicating that the effect is also economically significant.

On the other hand, I find no evidence of a relationship between women's second-order belief and their job application decision. A one unit increase in beliefs changes the probability a woman worker applies between $[-0.67, 0.94]$ percentage points in the full specification. While the sample size is smaller for women (220 compared to 341 men in the full specification), the coefficient is also an order of magnitude smaller for women compared to men (0.14 versus 1.28 percentage points). As I discuss in more detail in Section 2.6, this null result could be due to the selection of women into the male-dominated occupation for which I recruit.

Result 1. *More favorable second-order beliefs about a manager type increase the likelihood that a man worker applies for a job with that manager type. I find no evidence that second-order beliefs affect the likelihood a woman worker applies for a job.*

Regression charts are contained in Appendix J. In the main specification, the vector of treatment variables is statistically significant for men in all but the full specification including controls. No other variables are statistically significant for men.

Similarly, the vector of treatment variables is statistically significant in all but the full specification with controls for women.[28] In addition, for women, preferences for the gender of the first manager are negative and statistically significant with a relatively large coefficient. Further investigation shows this effect is driven by women with male managers first, suggesting that even though women report a preference for working with a male manager, they may be less likely to apply to a male hiring manager.[29] Lastly, average beliefs is positive and statistically significant for women in the full specification with controls. This coefficient suggests that women with overall more favorable beliefs are more likely to apply.

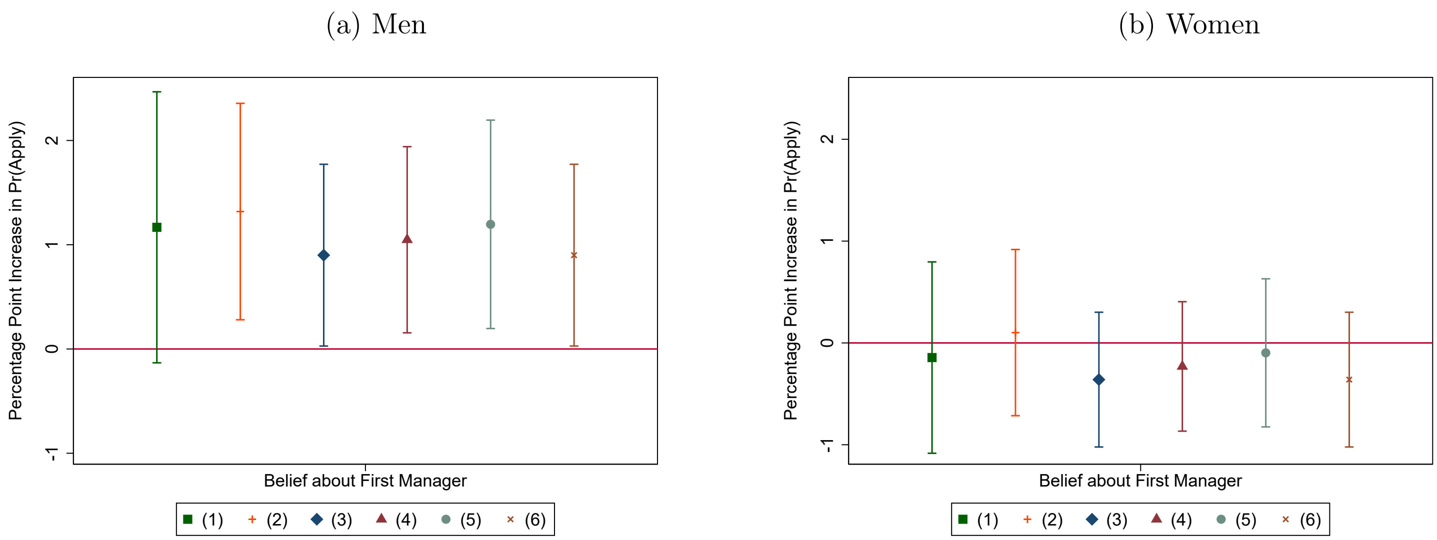
2.3.3 Robustness to Alternative Specifications

This result is robust to alternative specifications. First, I control for the first manager listed interacted with treatment. That means I include eight treatment variables, instead of four, corresponding to both the treatment group and the first manager listed, in the specification in equation 2.3. Figure 2.5 shows that the the size of the coefficient is the same for men. The estimate is noisier since I lose degrees of freedom with little gain in explanatory value, but the coefficient remains significant at the 10% level. There is still no detectable effect of beliefs on women's job application decision.

[28] With one minor exception, treatment 3 is also not statistically significant once preferences are added to the specification.

[29] Recall that the preferences I elicit in the lab experiment are coarse and unincentivized.

Figure 2.5: Regression Coefficient on Belief about First Manager by Gender, Other Specifications



Notes: 95% confidence intervals are shown. Standard errors are robust. Coefficient is scaled to a one unit increase in the favorability of beliefs. Panel (a) reports the coefficient on the regression with men only and panel (b) reports the same for women only. In each graph, specification (1) interacts treatment with first manager type. Specification (2) adds first-order beliefs. Specification (3) replaces the most favorable belief of the two facing managers with the belief about the second manager type. Specification (4) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the most favorable belief of all managers, excepting the first. Specification (5) replaces the average belief of all other managers, except the first, with the most favorable belief of the non-facing managers. Specification (6) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the belief about the second manager type and the average belief of the non-facing managers.

Second, I add first-order beliefs. If workers' beliefs about the relative productivity of women and men affect their job application decision directly, and second-order beliefs are informed by workers' first-order beliefs, the effect of second-order beliefs on the job application decision could be confounded. I find no evidence this is the case. The coefficient remains the same and statistically significant at the 5% level.

I also consider other specifications of the alternative beliefs. The main specification is determined by a theoretical model because there are many reasonable ways to control for the alternatives in the worker's dynamic application decision. I show that the results are robust to many other specifications of the alternative beliefs.

First, I replace the most favorable belief of the two facing managers with the belief about the second manager listed in the job ad. The former alternative assumes that the worker's belief about the second manager type only matters if it is more favorable than the worker's belief about the first manager type, while the latter allows the worker's belief about the second manager type to enter the worker's decision even if it is less favorable. Second, I replace both alternative beliefs with the most favorable belief out of all of the manager types that are not the first manager listed. The idea of this alternative belief is that the worker decides only between the first manager type and taking another draw, implicitly assuming the worker does not consider the second manager type. Third, I replace the worker's average belief about all manager types except the first with the most favorable belief about the non-facing manager types. This alternative treats the worker's decision as a choice between the most favorable of the two manager types the worker faces in the experiment and the most favorable of the two manager types the worker does *not* face in the experiment. Lastly, I consider a combination of the alternatives by replacing the most favorable belief about the two facing managers with the belief about the second manager *and* replacing the average of all other managers, excepting the first, with the average of just the two non-facing managers. This specification assumes the worker chooses between the first manager type, the second manager type, or taking another draw in expectation of the two manager types the worker does not face in the experiment.

Figure 3.12 shows that different ways of controlling for the alternative beliefs has little effect on the estimated coefficient on worker's beliefs about the first manager type listed for men. Moreover, the coefficient remains significant at the 5% level in all specifications. On the other hand, the estimated coefficient for women remains small and not statistically distinguishable from zero, though it does change sign to be negative in these other specifications.

2.4 Field Experiment Results

This section describes results from the component field experiment that are policy-relevant independent of the main result. In the field experiment, I observe workers' job application decisions after randomly assigning workers to a treatment that determines the gender of the manager in each of two sectors: the office products sector (female) and the janitorial products sector (male). I specify three hypotheses in the pre-analysis plan for the field experiment data, which I test in this section.

Hypothesis 2. *Men workers are more likely to apply than women workers.*

Hypothesis 3. *Workers are more likely to apply to their same-gender product sector.*

Hypothesis 4. *Workers are more likely to apply to their same-gender manager.*

I find strong evidence in favor of Hypothesis 2 that men workers are more likely to apply than women workers. Men are almost twice as likely to apply for a job (7.8%, $SE = 0.50\%$) compared to women workers (4.1%, $SE = 0.47\%$) and this difference is statistically significant ($p = 0.000$).[30] Moreover, Figure 2.6 shows that women are less likely to apply than men in every treatment group ($p = 0.0158$, $p = 0.000$, $p = 0.025$, $p = 0.010$ for treatments 1, 2, 3, and 4 respectively).

Result 2. *Men workers are more likely to apply to the job than women workers, overall and in each treatment.*

Figure 2.6: Proportion of Participants in each Treatment who Applied for the Job

Notes: From left to right on each graph, the bars show the proportion of participants who applied for the job in treatment 1 with a male manager in both sectors, treatment 2 with a female manager in both sectors, treatment 3 with a sector-matching manager in both sectors, and treatment 4 with a non-sector-matching manager in both sectors. Standard error bars are robust.

I find no evidence in favor of Hypothesis 3 that workers are more likely to apply to their same-gender sector. Figure 2.7 shows that both women (57%, $SE = 5.8\%$) and men (53%, $SE = 3.3\%$) are slightly more likely to choose the office sector than the janitorial sector, conditional on applying, but these proportions are not statistically different from 50% ($p = 0.123$ for women and $p = 0.840$ for men).[31] Unconditional on applying, I find no statistically significant differences between the probability a woman applies to the office sector (2.3%, $SE = 0.35\%$) compared to the janitorial sector (1.8%, $SE = 0.31\%$), nor the

[30] p-values are reported for one-sided tests of proportions unless otherwise noted in this subsection since the pre-specified hypotheses are one-sided.

[31] For men, I use the one-sided test on the probability they apply to the janitorial sector since that was the pre-specified hypothesis. The probability men apply to the office sector is also not statistically different from 50% using a two-sided test ($p = 0.160$).

probability a man applies to the janitorial sector (3.7%, $SE = 0.35\%$) compared to the office sector (4.2%, $SE = 0.37\%$) ($p = 0.120$ for women and $p = 0.840$ for men).

Result 3. *I find no evidence that workers are more likely to choose their same-gender sector.*

Figure 2.7: Proportion of Participants in each Treatment who Chose the Janitorial Sector, Conditional on Applying

Notes: From left to right on each graph, the bars show the proportion of participants who applied to the janitorial sector, conditional on applying for the job, in treatment 1 with a male manager in both sectors, treatment 2 with a female manager in both sectors, treatment 3 with a sector-matching manager in both sectors, and treatment 4 with a non-sector-matching manager in both sectors. Standard error bars are robust.

I cannot do a direct comparison of application rates for a female versus a male manager to test Hypothesis 4 that workers are more likely to apply to a same-gender manager because there are no treatments that hold the sector constant and vary the gender of the manager. Rather, I specify the following tests in the pre-analysis plan.

1. The probability a woman worker applies to the office sector is higher when there is a female compared to a male manager in the office sector, and a male manager in the janitorial sector.

2. The probability a man worker applies to the janitorial sector is higher when there is a male compared to a female manager in the janitorial sector, and a female manager in the office sector.

3. The probability a woman worker applies to a female manager, conditional on a female manager being available, is higher than the probability she applies to a male manager, conditional on a male manager being available.

4. The probability a man worker applies to a male manager, conditional on a male manager being available, is higher than the probability he applies to a female manager, conditional on a female manager being available.

I find no evidence in support of Hypothesis 4 for women using either test 1 or test 3. For test 1, women are no more likely to apply for a job with a female manager than a male manager in the office sector when the alternative is a male manager in the janitorial sector ($p = 0.594$).[32] In fact, the point estimate is slightly in favor of the male manager (2.7%, $SE = 0.76\%$ apply to the male manager compared to 2.4%, $SE = 0.72\%$ who apply to the female manager). Furthermore, test 3 finds that women are no more likely to apply for a job with a female manager when a female is available compared to a male manager when a male is available ($p = 0.599$) and the point estimate is again slightly in favor of male managers (2.8%, $SE = 0.44\%$ apply to a male manager, when available, compared to 2.6%, $SE = 0.45\%$ who apply to a female manager, when available).

Evidence is strongly in favor of the *opposite* conclusion from Hypothesis 4 for men workers. For test 2, men are statistically more likely to apply for a job with a female manager (4.2%, $SE = 0.75\%$) than a male manager (2.2%, $SE = 0.55\%$) in the janitorial sector when the manager in the office sector is female ($p = 0.029$ in a two-sided test). Moreover, test 4 finds that men are more likely to apply to a female manager when a female is available (76%, $SE = 3.3\%$) compared to a male manager when a male is available (59%, $SE = 3.9\%$) and this difference is statistically significant ($p = 0.018$ in a two-sided test).

Result 4. *I find no evidence that workers are more likely to apply to their same-gender manager. Men workers are more likely to apply to a female manager than a male manager.*

2.5 Lab Experiment Results

The component lab experiment provides independent evidence on workers' second-order beliefs that have policy implications beyond the main results described in Section 2.3. In the lab experiment, I elicit workers' beliefs about the beliefs of managers who are relevant to their job application decisions in the field experiment. I first present results on the primary beliefs of interest in Section 2.5.1: second-order beliefs. Then, I show the results for workers' first-order beliefs in Section 2.5.2 and their relationship to second-order beliefs, or intra-participant beliefs, in Section 2.5.3.

[32] Test 1 (and test 2 for men) was designed with the idea that workers would be most responsive to the availability of a same-gender manager in their same-gender sector. Alternative tests, such as whether women are responsive to the availability of a same-gender manager in the opposite-gender sector, are also null.

2.5.1 Evidence on Second-Order Beliefs

Average second-order beliefs, reported by manager type and by gender in Table 2.6, show that participants believe different manager types hold different beliefs.[33] I reject equality of the four averages using the sample of all participants in column (1) ($p = 0.000$ for all tests).[34] The differences in participants' beliefs about the beliefs of different manager types is further evidenced by Figure 2.8, which aggregates beliefs into a ternary outcome. I reject equality of the four ternary belief distributions ($p = 0.000$ for all tests using the Wilcoxon signed-rank test).[35,36] These differences are not reflected in differences between women's and men's second-order beliefs. Table L shows that women report slightly more extreme beliefs than men about the sector-matching managers (less that one unit on the 21-point scale); however, there are no differences in women's and men's beliefs about non-sector-matching manager types.

Table 2.6: Average Second-Order Beliefs

Sample	(1) All	(2) Women	(3) Men	(4) Difference
Female Manager in Female Sector	-17.352	-20.426	-15.440	4.986***
	(0.897)	(1.438)	(1.139)	(1.833)
Female Manager in Male Sector	0.634	-0.157	1.122	1.279
	(0.963)	(1.668)	(1.170)	(1.981)
Male Manager in Female Sector	-5.781	-5.404	-6.037	-0.633
	(1.020)	(1.722)	(1.264)	(2.098)
Male Manager in Male Sector	20.712	22.825	19.432	-3.393*
	(0.890)	(1.573)	(1.055)	(1.824)
Observations	576	223	352	575

Notes: Columns (1) to (3) reference subsamples. Column (4) reports the differences between the women and men subsamples. Standard errors are reported in parentheses underneath the means.

[33] As discussed in Section 3.3, the form of the payment function dictates that participants optimally report the median of their subjective belief distributions, so the means in Table 2.6 are the averages of the medians reported by participants.

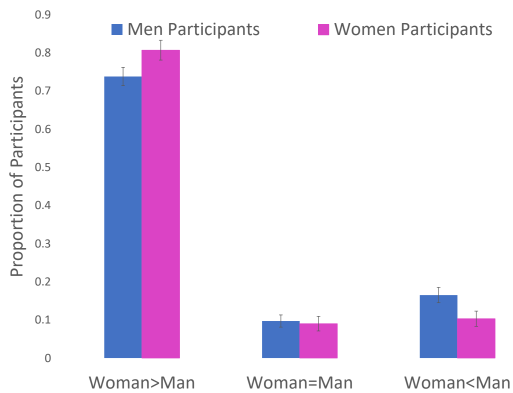
[34] All tests in this subsection are two-sided t-tests or tests of proportions, as appropriate.

[35] The Wilcoxon signed-rank test for matched pairs accounts for intra-participant dependence, meaning that the samples for each distribution tested are not independent of each other because they are composed of the same participants.

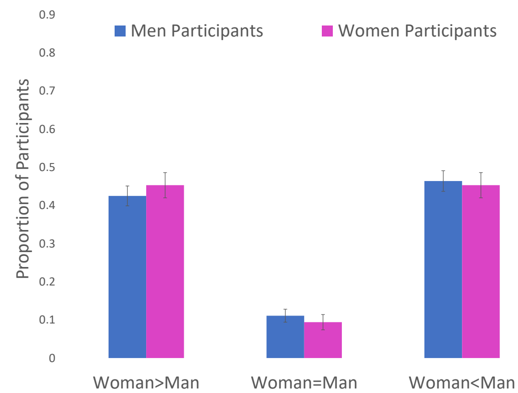
[36] I also reject the equality of the four cardinal distributions, which are reported in Appendix L (again, $p = 0.000$ for all tests).

Figure 2.8: Ternary Distributions of Second-Order Beliefs

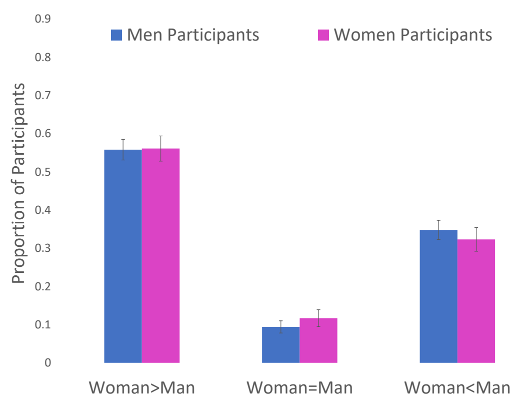
(a) Female Manager in Female Sector



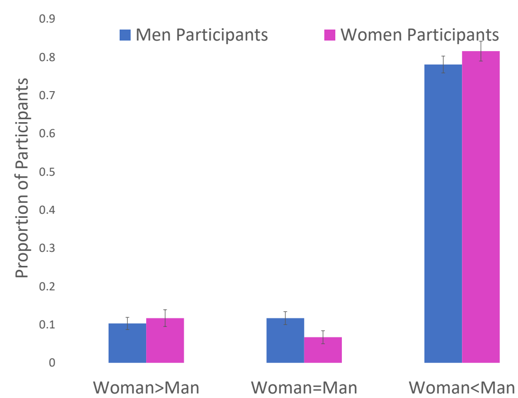
(b) Female Manager in Male Sector



(c) Male Manager in Female Sector



(d) Male Manager in Male Sector



Notes: From left to right on the graphs, the bars show the proportion of participants who report a belief that *the most likely* outcome of a random draw of the relevant manager type is that the manager reports the woman outperforms the man, the woman performs the same as the man, the man outperforms the woman on the sales assessment.

I next consider whether workers with the second-order beliefs of participants in this lab experiment would hypothetically sort by gender in the labor market if they acted exclusively on their second-order beliefs. That is, do participants believe that female managers and managers in the female sector favor women? And do they believe that male managers and managers in the male sector favor men?

Participants' beliefs are consistent with sorting by gender behavior. Most participants believe that sector-matching manager types favor the manager's own gender. 76% ($SE =$

1.8%) of participants believe that *most* female managers in female sectors believe that the most likely outcome of a random draw is that the woman outperforms the man, while 79% ($SE = 1.7\%$) believe that *most* male managers in male sectors believe the opposite, that the most likely outcome of a random draw is that the man outperforms the woman.[37]

Non-sector-matching managers mitigate sorting by gender behavior. Only 56% ($SE = 2.1\%$) of participants believe that most male managers in female sectors believe that the most likely outcome of a random draw is that the woman outperforms the man. Similarly, just 46% ($SE = 2.1\%$) of participants believe that most female managers in male sectors believe that the most likely outcome is that the man outperforms the woman.

2.5.2 Evidence on First-Order Beliefs

I find that participants' first-order belief distribution is centered around zero. Table 2.7 reports that, on average, participants reveal a belief that the median percentile rank difference between a man and a woman is approximately one percentile point in favor of the woman (-0.77, $SE = 0.91$). I cannot reject the null hypothesis that the average first-order belief is zero ($p = 0.398$).

The ternary first-order belief distribution shown in Figure 2.9 and reported in Table 2.7 shows that the gender neutrality suggested by the central tendency of the distribution conceals considerable heterogeneity in beliefs.[38]. 85% ($SE = 1.5\%$) of participants believe there is some difference in women's and men's sales ability; however, they do not agree on which gender outperforms the other. Participants are more likely to believe that the most likely outcome of a random draw is the woman outperforms the man (46%, $SE = 2.1\%$) than to believe the opposite, that the most likely outcome is the man outperforms the woman (39%, $SE = 2.0\%$), and this difference is statistically significant ($p = 0.007$).

I find no statistically distinguishable differences in the first-order beliefs reported by women and men participants. Table 2.7 shows that there are no differences in the average cardinal beliefs nor ternary belief distribution outcomes. I also cannot reject at conventional significance levels that women's and men's first-order cardinal ($p = 0.297$) or ternary ($p = 0.873$) belief distributions are identical using the Wilcoxon rank-sum test.

[37] This interpretation is derived from the use of the median. A participant who reports a negative median believes there is at least a 50% probability that the manager reports a belief that the woman outperforms the man. Similarly, a participant who reports a positive median believes there is at least a 50% probability that the manager reports a belief that the man outperforms the woman.

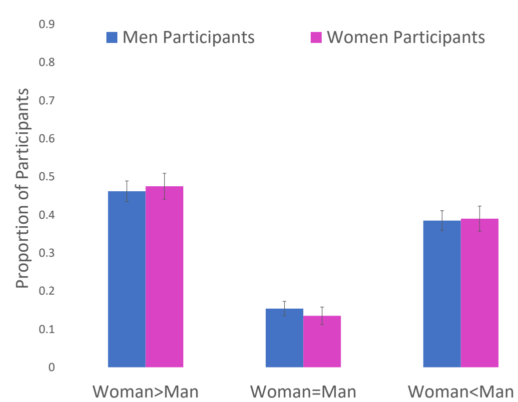
[38] Cardinal belief distributions in Appendix L also demonstrate this heterogeneity

Table 2.7: First-Order Beliefs

	All	Women	Men	Difference
Average Belief	-0.773	-2.152	0.099	2.252
	(0.914)	(1.584)	(1.108)	(1.878)
Ternary Beliefs				
W>M	0.465	0.475	0.460	-0.015
	(0.021)	(0.034)	(0.027)	(0.043)
W=M	0.148	0.135	0.153	0.019
	(0.015)	(0.023)	(0.019)	(0.030)
W<M	0.387	0.390	0.386	-0.004
	(0.020)	(0.033)	(0.026)	(0.042)
Observations	576	223	352	575

Notes: Columns (1) to (3) reference subsamples. Standard errors are reported in parentheses underneath the means. Column (4) reports the differences between the women and men subsamples. The rows labeled "W>M", "W=M", and "W<M" report the proportion of participants who believe that the most likely outcome of a random draw is the woman outperforms the man, the woman performs the same as the man, and the man outperforms the woman respectively.

Figure 2.9: Ternary Distribution of First-Order Beliefs



Notes: From left to right on the graph, the bars show the proportion of participants who believe that the most likely outcome of a random draw is the woman outperforms the man, the woman performs the same as the man, and the man outperforms the woman on the sales assessment.

2.5.3 Intra-Participant Beliefs

To evaluate how participants' first-order beliefs compare to their second-order beliefs about managers' beliefs, I calculate intra-participant beliefs by taking the difference between the participant's first-order belief and their second-order belief about the relevant manager type. For women, their first-order beliefs are compared to their second-order beliefs about female managers in each sector while, for men, their first-order beliefs are compared to their second-order beliefs about male managers in each sector. The intuition for focusing on a comparison with same-gender manager types is to compare participants' first-order beliefs with their second-order beliefs about managers "like them." Since differences are taken as first-order belief minus second-order belief (and beliefs are reported as man minus woman), positive figures indicate that the participant believes the manager's belief favors men more than the participant's own belief.

Table 2.8 shows that participants believe that they share similar first-order beliefs with managers of their same-gender in the *opposite-gender* sector, but that their same-gender managers' beliefs in the *same-gender* sector favor the opposite gender much more than participants do themselves. Women believe that the beliefs of female managers in the female sector favor men 18.27 percentile points ($SE = 2.06$) more than they themselves do ($p = 0.000$). On the other hand, I find no statistically distinguishable difference between women's first-order beliefs and their beliefs about female managers in the *male* sector ($p = 0.342$). A similar pattern emerges for men. Men believe that the beliefs of male managers in the male sector favor women 19.33 percentile points ($SE = 1.42$) more than they do ($p = 0.000$), but believe that the beliefs of male managers in the female sector favor women only marginally more than themselves (2.00, $SE = 2.09$; $p = 0.000$).

Table 2.8: Intra-Participant Cardinal Beliefs

	All	Women	Men	Differences
Female Sector	10.843	18.274	6.136	-12.137
	(1.318)	(2.063)	(1.665)	(2.660)
Male Sector	-12.609	-1.996	-19.332	-17.337
	(1.242)	(2.094)	(1.424)	(2.445)
Observations	575	223	352	575

Notes: Columns (1) to (3) reference subsamples. Standard errors are reported in parentheses underneath the means. Column (4) reports the differences between the women and men subsamples.

2.6 Discussion

In this paper, I combine a large-scale natural field experiment with a structured online lab experiment to show that workers' beliefs about managers' beliefs affect job application behavior, but I find this is true only for men workers, not women workers. This result raises the question of why women in the experiment are unresponsive to their second-order beliefs, particularly since they *believe* that most managers in the male sector, and male managers more generally, favor men in their beliefs. One explanation is that women who choose male-dominated occupations, like the sales job for which I recruit workers, may be selected based on the fact that they do not respond to their second-order beliefs. These women may have preferences such that choosing a non-conforming job or succeeding in adverse conditions overrides expectations about being hired. This experiment is not designed to study the selection of women into the occupation I recruit for, but future work should consider beliefs in a female-dominated occupation to determine if the opposite pattern arises there.

Despite the null results for women workers, the sorting patterns of men workers based on their second-order beliefs have important implications for the gender wage gap. Men believe that most male managers, and managers in male sectors, favor men. Sorting based on these beliefs increases the supply of workers to male occupations and industries relative to female occupations and industries, potentially crowding out women. Moreover, selection out of female occupations and industries is itself important if this sorting behavior leads to skill mismatch, where men who would otherwise choose female-dominated jobs choose not to because of the higher likelihood of female managers.

In addition to the main result showing that men's beliefs about managers' beliefs affects their job application decision, I observe interesting behavior in the field experiment. First, men are almost twice as likely to apply as women, even though they are selected into the field experiment in the same way. This result could mean that women have a higher cost of applying or lower expectations of being hired and/or wages to overcome that cost. Policy aimed at recruiting women to the applicant pool may be an effective tool to increase representation in men-dominated jobs.

Second, I find no evidence that women or men sort based on the sector gender. The beliefs I elicit in the lab experiment are explicitly characterized with respect to "traditionally female" or "traditionally male" sectors; however, the signal of sector gender in the field experiment may not be completely effective. It is also possible that, once sorted into a men-dominated (or women-dominated) occupation, workers are less likely to further sort within that occupation.

Third, I find that women are unresponsive to manager gender, but men are more likely to apply to their *opposite-gender* manager. The null results for women with respect to manager gender is in line with the results of Flory et al. (2015) finding that workers selected into a men-dominated occupation are unresponsive to manager gender. This result is also consistent with the unresponsiveness of women to their second-order beliefs. On the other hand, men's higher likelihood for applying to female managers is unexplained by the experiment. The fact that men are more likely to apply to female managers does not conflict with the main result that men act on their second-order beliefs, but it does indicate that there is another factor contributing to the job application decision.

The lab experiment also provides new insights. First, I find that workers' beliefs are consistent with sorting by gender behavior, a result that provides support for other literatures studying mechanisms that require workers to believe that different types of managers hold different beliefs. Moreover, the fact that workers overwhelmingly believe that sector-matching managers favor their own gender, but have less homogeneous beliefs about non-sector-matching managers, suggests a role for policies like quotas and affirmative action. Managers whose gender is not traditionally associated with a sector may attract more workers of their gender.

Second, I find few differences in women's and men's beliefs. Women report slightly more extreme second-order beliefs about sector-matching managers, but the size of the difference is quite small. Moreover, I find no differences in either second-order beliefs about non-sector matching managers or in first-order beliefs. This latter result may seem surprising, but it is consistent with the results in Dustan et al. (2020) and other studies. Women and men generally have the same beliefs about the differences between women and men, suggesting that factors influencing first-order beliefs are not gender-specific.

Third, women and men believe that their beliefs are reflected by managers of their same gender in the *opposite-gender* product sectors. For both women and men, participants believe that managers of their same-gender in gender-matching sectors have more extreme beliefs about the participant's gender. This result underscores the extreme beliefs that participants believe that sector-matching managers hold.

In summary, I provide evidence on an important new mechanism affecting workers' job application decision. I study this mechanism within the microcosm of one occupation; however, my results suggest that second-order beliefs could be driving broader sorting patterns in the labor market. Workers with the beliefs I elicit would be more likely to choose occupations and industries that match their own gender, contributing to the sorting by gender behavior underlying a large part of the gender wage gap.

Chapter 3: Mechanism Performance under Strategy Advice and Sub-Optimal Play: A School Choice Experiment

(This chapter was co-authored with Andrew Dustan, Martin van der Linden, and Myrna Wooders, and is published here with their permission.)

3.1 Introduction

Policymakers increasingly turn to school choice mechanisms to break the link between a student's residence and their public school assignment. Yet, the choice between the two most popular mechanisms remains fraught for the public as well as economists. The theoretical and experimental literatures have weighed in on both sides, but one issue that repeatedly arises is that mechanism performance largely depends on participants' strategy choices. People do not always play a theoretically optimal strategy for a number of reasons. One reason is that determining the optimal strategy may be costly. Real-world school choice markets address this issue through strategy advice. The school system, newspapers, and even internet blogs provide strategy advice intended to inform parents' strategy choices. Despite the prevalence of strategy advice in school choice, little work considers how structured advice affects participants' strategy choices. We address this gap in the literature with a laboratory experiment. We then exploit the induced variation in participant strategy from the strategy advice to address the ongoing debate about relative mechanism performance under sub-optimal strategies.

We replicate the seminal school choice laboratory experiment of Chen and Sönmez (2006), then add a strategy advice treatment, to study how strategic behavior responds to the advice. Under both the strategy-proof Deferred Acceptance (DA) mechanism and the manipulable Immediate Acceptance (IA) mechanism, our strategy advice substantially increases the proportion of participants who choose a recommended strategy. Next, we do exploratory analyses on the relative performance of the DA and IA mechanisms. In line with other studies (e.g. Chen and Sönmez, 2006, Calsamiglia et al., 2010), DA is both more efficient and more stable than IA in our experiment. This difference in performance is exacerbated in the advice treatments of our experiment.

Our new approach to evaluating relative mechanism performance considers individual welfare. We ask the following question: would a given participant in our experiment have ex-ante preferred to participate in a DA or IA session? To answer this question, we use recombinations of a tie breaker and the strategies of other participants to estimate the *expected* payoffs to participants who chose a recommended strategy in an IA session based on their observed strategy. Next, we again use this recombinant estimation technique to estimate the participant's expected payoff in a DA session. Because the participant did not actually participate in a DA session, this latter estimation requires the determination of a counter-factual strategy for the participant in DA. We address this requirement by developing a new typology of DA strategies that partially dominance-orders all possible strategies.

Comparing a participant's expected payoff from their observed strategy in IA to that same participant's expected payoffs from playing a series of increasingly irrational strategies in DA, we find that almost any participant who followed our strategy advice would have preferred to take part in a DA session rather than an IA session. On the other hand, some participants who do not follow the advice to implement one of two heuristic strategies fare better in IA sessions. In particular, the most disadvantaged participants, determined by their district school's preference ranking, who do not follow our advice choose strategies that generate higher payoffs than playing the dominant strategy in DA.[1] In summary, we show that the DA mechanism is not only more efficient in terms of overall welfare, but is welfare-enhancing for almost any individual, particularly those who play heuristic strategies like the ones recommended in real-world markets. The exceptions to this conclusion are disadvantaged participants who play more sophisticated strategies than the simple heuristics, who constitute a relatively small proportion in our sample.

3.1.1 The Experiment

To test how participants respond to strategy advice, we design our experiment as a variant of the school choice laboratory experiment in Chen and Sönmez (2006). Chen and Sönmez test the truth-telling rates and efficiency of three school choice mechanisms: DA, IA, and Top Trading Cycles (TTC). We restrict our attention to DA and IA, then add a strategy advice treatment for each mechanism. Chen and Sönmez's rich experimental design allows us to study participants' strategy choices in a relatively elaborate environment intended to replicate the size and complexity of real-world school choice problems.

The strategy advice that we give is inspired by strategy advice observed in the field. For example, the Minneapolis Central Placement and Assessment Center (CPAC) recommends that parents not tell the truth in the IA assignment mechanism used in that city. Instead, CPAC recommends that parents list their most preferred school first, but advises them to rank their district school second or third, even if this school is not their true second or third choice. News sources also contribute advice. C. Mas in The Seattle Press (1998) suggests that parents may want to list their district school first in the IA mechanism to avoid being placed at a less preferred school. We call the former recommendation the "Risky" strategy, defined as ranking one's most preferred school first and district school second. We call the

[1]Participants with the lowest district school rank would be unwise to play our recommended strategies, even though they are commonly observed in the field.

latter recommendation the "Safe" strategy, defined as ranking one's district school first. In addition to being recommended in real world markets, the two simple heuristic strategies are observed spontaneously in lab experiments (e.g. Chen and Sönmez, 2006), even though they are unlikely to be optimal for many students.

Strategy advice in DA is more straightforward because there is only one correct recommendation: tell the truth by submitting preferences in order from most- to least-preferred school.[2] School districts that adopt DA, like the New York City Department of Education, distribute manuals that include strategy advice to tell the truth when ranking schools.

We find that, in both IA and DA, our strategy advice increases the proportion of participants who choose a recommended strategy. In DA, advice to play the dominant strategy of truth-telling increases the proportion of participants who truthfully reveal their preferences from 31% to 50%. Our strategy advice in IA, inspired by strategy advice from the field, increases the probability that a participant chooses one of the two simple heuristic strategies from 51% to 71%. Overall, we provide strong evidence that strategy advice has an economically large and statistically significant effect on participants' strategy choices in the lab.

We provide the first evidence on how strategy advice affects behavior in a non-strategy-proof mechanism and new evidence on strategy-proof mechanisms in a more complex environment. Independently of our study, Ding and Schotter (2019) study strategy advice that is devised and shared by participants in the same two mechanisms as our experiment. In their experiment, participants in one generation pass down a recommended ordering over three schools to the next generation. In contrast, our strategy advice is consistent across all participants, in line with the type of advice that may be distributed by school districts for example. This difference appears to matter since Ding and Schotter find an increasing rate of truth-telling in the IA mechanism with strategy advice over generations. In our experiment, truth-telling rates in IA drop significantly in the strategy advice treatment. Similarly for the DA mechanism, while strategy advice derived by participants in Ding and Schotter decreases the truth-telling rates, structured strategy advice in our experiment leads to significantly higher rates of truth-telling. The contrasting results of our studies suggests that further research is needed since participants in real world school choice mechanisms likely receive both structured advice like ours and advice from their network, as in Ding and Schotter.

[2]Truth-telling is the weakly dominant strategy in the DA mechanism because it is strategy-proof. DA fails to be strategy-proof in some circumstances. Specifically, when the number of schools students can report is less than the total number of schools the student prefers to their outside option or when there are fewer available seats at a student's district school than there are students assigned to that school, truth-telling may not be the dominant strategy.

Zhu (2015) also considers strategy advice that is passed from participant to participant in the DA mechanism. In her lab experiment, Zhu has participants in one session play the mechanism repeatedly, then asks them to provide strategy advice to participants of the same type (preferences and priorities) in future sessions. The strategy advice takes the form of a recommended ranking over the three schools and, due to the learning that occurs over 15 rounds, is mostly correct. She finds that advice increases truth-telling rates, but not at a statistically significant level. We show that strategy advice increases truth-telling in DA at a statistically significant level in an environment with greater complexity (i.e. more players types and choices). In addition, our advice is given by a person in authority (the experimenter) as opposed to a fellow participant.

Two papers by Guillen and co-authors study strategy advice on the TTC mechanism that is structured, like our own, to be the same for all participants. First, in an introductory microeconomics course at the University of Sydney, Guillen and Hakimov (2018) test the effect of different information environments on participant strategy. In their experiment, the authors explain the strategy-proofness property of TTC in treatment one, explain just the mechanism's algorithm in treatment two, and explain both the property and algorithm in treatment three. They find that telling participants the mechanism is strategy-proof increases the likelihood that students truthfully reveal their top preference out of three possible topics for their term paper. Compared to our experiment, the environment in Guillen and Hakimov has fewer choices, perfectly correlated preferences, and the observation of truth-telling limited to the top-ranked choice, but we find comparable results for the DA mechanism.

Guillen and Hing (2014) evaluate the effects of correct versus incorrect third-party advice on strategic behavior in TTC. The correct advice advocates ranking schools according to the participant's true preferences while the incorrect advice suggests ranking the participant's district school first. Whether correct or incorrect, strategy advice from a third-party decreases the likelihood that participants truthfully reveal their preferences compared to a baseline treatment of no advice. Again, our strategy advice comes from an authority figure (the experimenter) and positively affects the rate of truth-telling, suggesting a need for further research into the effect of the source of advice.

Two other papers indirectly consider strategy advice. Ding and Schotter (2017) study chat between participants in a laboratory experiment, where the content of the chat could be considered strategy advice from other participants (but could also be signaling information about other participants' preferences, priorities, and strategies). The "chat" treatment induces more participants to change their strategy after the chat, particularly when chatting with participants who have the same preferences and priorities as them. Chat also increases the number of stable outcomes in both the DA and IA mechanisms. Individual payoffs increase only when participants communicate with others that are unlike them. Braun et al. (2014) incorporate strategy "coaching" in their lab experiment. Although they do not study strategy advice, the authors show that strategy coaching increases the likelihood of truth-telling in a strategy-proof mechanism with a small control group. In the mechanism that rewards preference manipulation, they find a lower proportion of participants play the naïve strategy of truth-telling. Our strategy advice is structured the same for each participant, as opposed to strategy coaching, and their experiment incorporates learning over rounds, while we study a one-shot game that mimics real world school choice applications.

3.1.2 Mechanism Performance

We use our experimental data to do exploratory analyses on the relative performance of DA and IA under varying levels of sub-optimal behavior. These analyses contribute to the literature launched by the theoretical paper Abdulkadiroğlu and Sönmez (2003) and laboratory experiment Chen and Sönmez (2006) pitting DA against IA. Their papers catalyzed changes in a number of school systems, including the well-known conversion from IA to DA by the Boston school system in 2005 (Abdulkadiroğlu et al., 2006). Foremost of the school board's reasons for switching mechanisms were the strategy-proofness and stability of DA. Strategy-proofness theoretically eliminates the cost of strategizing and levels the playing field for students with heterogenous costs to strategizing. The stability of DA is appealing because school boards believe the elimination of justified envy gives the appearance of a more "fair" outcome (Abdulkadiroğlu et al., 2006). The trade-off for the strategy-proofness and stability of DA, however, is IA's theoretically possible greater efficiency.

DA is not Pareto efficient, but theoretically it generates the most efficient stable solution. DA's efficiency and stability, however, rely on participants choosing the dominant strategy of truth-telling. Lab experiments find truth-telling rates far from 100% (e.g. Chen and Sönmez, 2006, Klijn et al., 2013, Basteck and Mantovani, 2018). Empirical studies also find preference manipulation in DA (e.g. Hassidim et al., 2016), confirming that the sub-optimal rates of truth-telling are not an artifact of the lab. Rees-Jones (2018) studies the high-stakes residency match that implements the DA mechanism. Among medical students with prevalent (and correct) strategy advice in a system that has had years to establish the strategy-proofness of DA, Rees-Jones still finds students that fail to play the dominant strategy of truth-telling. We also find high proportions of participants choosing sub-optimal strategies in our experiment, even in the advice treatment when participants are advised to choose the dominant strategy of truth-telling.

Theoretical work argues that IA's greater efficiency may outweigh DA's stability and strategy-proofness, particularly when students have limited information about each others' preferences and about schools' priorities (Abdulkadiroğlu et al., 2011; Miralles, 2009; Troyan, 2012). Miralles (2009) and Abdulkadiroğlu et al. (2011) assert that IA may be more efficient than DA in practice because IA incorporates information about cardinal preferences in its algorithm. The strategic manipulability of IA allows participants to signal preference intensity, while DA only accounts for ordinal preferences. Troyan (2012) shows that theoretically IA ex-ante Pareto dominates all strategy-proof mechanisms. The ex-ante argument hinges critically on no information being known— participants do not know their preferences, schools do not have known priorities, and the number of students at each priority level is unknown.

Arguments for the greater efficiency of IA rely on either all participants truth-telling or all playing strategies that correspond to the same sophisticated Bayesian-Nash equilibrium; however, experiments such as Featherstone and Niederle (2016) show that participants typically fail to play a non-truth-telling Bayesian-Nash equilibrium in IA. Bayesian-Nash equilibria are particularly unlikely if participants rely on simple heuristic strategies, which naturally occur in experiments like Chen and Sönmez (2006). While a simple heuristic strategy may be optimal for some people, in an environment with any sort of complexity, such strategies are generally sub-optimal and are unlikely to lead to a Bayesian-Nash equilibrium. It is important to note, however, that we cannot say for certain that the simple heuristic strategies we recommend are sub-optimal.[3] Instead, we base our analysis only on the assumption that the induced variation in strategies led to varying proportions of sub-optimal strategies without distinguishing whether the advice or no advice treatment has the higher proportion.

Since experimental and empirical studies find that participants are choosing sub-optimal strategies, we consider how those strategies affect the relative performance of the DA and IA mechanisms. Concerning the observed preference misrepresentation in the strategy-proof medical residency match, Rees-Jones (2018) writes:

> "...the persistence of suboptimal behavior in this setting, even at low rates, suggests the requisite levels of intelligence, information, and incentivization needed to ensure full compliance may never be achieved in practice. Some strategic misunderstanding may be unavoidable in these settings, necessitating attention to the comparative performance of mechanisms in the presence of suboptimal behavior..."

We find that, consistent with the experimental literature (Calsamiglia et al., 2010; Chen and Sönmez, 2006), DA outperforms IA in both our advice and no advice treatments. Our

[3] An "optimal" strategy in IA is based on a participants' beliefs about the choices of other participants, which we do not observe.

strategy advice has little effect on the efficiency of DA; however, the advice to play a heuristic strategy in IA decreases efficiency. Likewise, stability increases under strategy advice in DA, but decreases in IA. Our experimental results show that sub-optimal play decreases the performance of DA, but DA still outperforms IA.

We next develop a new dimension to evaluating relative mechanism performance by considering the welfare of individual participants. Studying individual welfare allows us to observe heterogeneity in mechanism performance with respect to where a student's district school ranks in their preferences. Specifically, we compare the welfare differences between participating in the IA versus the DA mechanism by the preference ranking of a participant's district school. Participants with district schools that lead to a higher payoff are at a relative advantage in our experiment, much as students who live in wealthy neighborhoods with good schools in real-world school choice markets have an advantage.

To conduct this analysis, we evaluate individual participants' welfare in a counter-factual exercise using recombinant estimations of expected payoffs. For each participant in an IA treatment who chose one of the recommended heuristic strategies, we estimate the expected (rather than observed) payoff to their observed strategy and the expected payoff from participating in a DA treatment instead. We do not, however, know the counter-factual strategy a participant would have chosen in a DA treatment. Assuming the participant would choose the dominant strategy exaggerates the payoffs to DA (particularly considering we observe that less than half the participants play the dominant strategy).

We address this issue by developing a new typology of DA strategies that allows us to partially order sub-optimal strategies in DA by dominance. Sub-optimal strategies in DA are currently classified by specific biases (see Hakimov and Kübler, 2019 for a review). While these biases are informative for understanding sub-optimal behavior, they do not allow us to rank strategies according to their dominance. The typology of DA strategies we develop addresses this issue by deconstructing DA strategies into three strategic aspects. A participant who understands all three strategic aspects of DA chooses a dominant strategy.[4] Participants who understand more — but not necessarily all — of these strategic aspects play "more rational" strategies compared to participants who understand fewer of them.

Using this typology, we calculate the expected payoff for each participant in the IA sessions of, instead, participating in a DA session using *increasingly irrational* strategies as counter-factual strategies. We find that almost any participant choosing one of the recommended heuristic strategies in the IA treatment would have preferred to participate in the DA treatment. Participants can choose strategies that are "far" from the dominant strategy

[4]We say "a" as opposed to "the" here to emphasize that there are multiple dominant strategies since schools ranked below a participant's district school are irrelevant.

in DA and still be better off than by playing a heuristic strategy in IA. The results are mixed for participants who do not choose one of the recommended strategies in IA. Participants who are at a particular disadvantage in the school choice game (defined by their district school's ranking in their preferences) would be considerably better off playing any strategy in DA if they are the type to follow strategy advice, while disadvantaged participants who choose a non-heuristic strategy fare better in the IA mechanism compared to even the dominant strategy in DA.

3.1.3 Organization

The remainder of this paper is organized as follows. Section 3.2 presents the formal school choice problem, the two mechanisms, and their theoretical properties. Section 3.3 describes our experimental design and treatments to test the effects of strategy advice on participants' strategic decisions. We discuss the lab results in Section 3.4. In Section 3.5, we comprehensively evaluate the relative performance of the two mechanisms under sub-optimal strategy. First, we evaluate the relative efficiency and stability of IA and DA under varying levels of sub-optimal play. Then, we introduce our new typology of DA strategies that allows us to evaluate the welfare of participants who choose one of the recommended heuristic strategies. Finally, we conduct the counter-factual analyses of individual welfare. Section six concludes with a discussion of our results.

3.2 The Theoretical School Choice Problem

We set up the school choice problem in this section to discuss the relevant properties of mechanisms that assign students to schools in a one-sided matching game. These properties are important to understand because they inform our mechanism performance analysis as well as the strategy advice we give to participants. We then detail the algorithms for the two mechanisms we study in this paper, DA and IA, and their properties.

The school choice problem is a one-sided matching game in which a set of N students are matched to M schools with limited capacities q_m for $m \in M$. Students have a strict preference ordering P_i over schools and schools have weak priority levels F_m over students. Priorities are coarser than preferences, allowing for indifferences, so a fair lottery is used to break ties between students. These priority levels are fundamentally different from student preferences because schools are objects to be consumed by the students. In real-world settings, priorities are set by education boards or state laws. For example, priority is often given to a student whose sibling already attends the school or to a student who lives within walking distance of the school.

A solution to the school choice problem is a matching μ that assigns students to schools such that no school has more students than its capacity q_m. An assignment mechanism ϕ is a function that inputs each school's capacity q_m, students' *reported* preferences Q_i[5], and schools' priorities F_m, then outputs a matching μ. We denote the reported preferences of students other than i as Q_{-i}. Then, the outcome of mechanism ϕ when the reported preferences are (Q_i, Q_{-i}) is $\phi(Q_i, Q_{-i})$, with $\phi_i(Q_i, Q_{-i})$ denoting the school that student i matches to under $\phi(Q_i, Q_{-i})$.

An assignment mechanism may have a number of desirable properties. One property is strategy-proofness. A mechanism is **strategy-proof** if truthful revelation of preferences is the weakly dominant strategy for each participant. Strategy-proofness is popular in real-world applications of school choice for a number of reasons. First, it reduces the cost of participating in the school choice problem for all participants by eliminating the need to search for the optimal strategy. Second, it does not punish naive participants who truthfully reveal their preferences. Third, it levels the playing field between those with differential costs to strategizing. Since school choice is often implemented to give students from a disadvantaged background the opportunity to go to better schools, this equity quality is particularly important to policymakers.

Assignment mechanisms may also be stable. In the school choice context, stability is generally limited to the elimination of justifiable envy.[6] A matching μ is **stable** if *no* student i prefers school m over the school i is assigned to while having a higher priority at m than a student who currently fills a slot at m. If the matching is not stable, then student i is *justifiably envious* of the student taking a slot at the school student i prefers when that school also prioritizes student i. A mechanism is stable if it always produces a stable matching. Justifiable envy was cited by the Boston school board as a potential source of lawsuits (Abdulkadiroğlu et al., 2006). The elimination of such a risk and the perception of "fairness" make stability a desirable property for policymakers.

A third potential property of assignment mechanisms is Pareto efficiency. A matching μ is **Pareto efficient** if there does *not* exist a matching ν that all students weakly prefer to μ and at least one student strictly prefers ν to μ. When any equilibrium outcome under a mechanism is Pareto efficient, that mechanism is Pareto efficient. A Pareto efficient mechanism is desirable because each student is weakly better off in a Pareto efficient matching than in any other matching. Policymakers desire efficiency because it increases the aggregate gains of all students. Note that Pareto efficiency is often defined with respect to reported

[5]Note this is different from a student's true preferences P_i.

[6]More generally, stability is defined as the combination of (i) the elimination of justified envy, (ii) non-wastefulness and (iii) individual rationality. The two latter requirements are automatically satisfied in our experiment.

preferences as opposed to true preferences, which is how a Pareto efficient mechanism may be less efficient in practice than a mechanism that is not considered Pareto efficient.

3.2.1 Deferred Acceptance

The DA mechanism is strategy-proof and stable. DA produces a matching that Pareto dominates all other stable matchings, but it is not a Pareto efficient mechanism. The DA mechanism solves the school choice problem by implementing the actions dictated by the students' reported preferences, the schools' exogenously-determined priorities, and the tiebreaker. The mechanism proceeds as follows:

- Each student submits an application to their first ranked school (according to their *reported* preferences).
- Each school rejects the lowest priority students in excess of its capacity and holds the remaining student applications with higher priority.
- Students rejected in the first round apply to their second ranked school.
- Each school considers the new applications together with the applications of the students on hold from the first round. The school rejects the lowest priority students in excess of its capacity and holds the remaining student applications with higher priority.
  ⋮
- Each student rejected in the previous round applies to their next ranked school.
- Each school considers the new applications together with the applications of the students on hold from the last round. The school rejects the lowest priority students in excess of its capacity and holds the remaining student applications with higher priority.

The algorithm terminates when no students are rejected in a round.

3.2.2 Immediate Acceptance

The IA mechanism is neither strategy-proof nor stable. It produces a Pareto efficient matching when all students truthfully reveal their preferences or when they play another Bayesian-Nash equilibrium. The IA mechanism solves the school choice problem by implementing the actions dictated by the students' reported preferences, the schools' exogenously-determined priorities, and the tiebreaker. The mechanism proceeds as follows:

- Each student submits an application to their first ranked school (according to their *reported* preferences).

- Each school accepts the highest priority students until its available seats are filled. All other students are rejected.

- Each school who still has seats remaining from the first round accepts applications from students who rank the school second.

- Each school accepts students with the highest priority in the second round until its available seats are filled. All other students are rejected.
  ⋮

- In the i^{th} round, each school who still has seats remaining from the $i-1$ round accepts applications from students who rank the school i^{th}.

- Each school accepts students with highest priority in the i^{th} round until its available seats are filled. All other students are rejected.

The algorithm terminates when all students have seats at schools.

3.3 Experiment Design

We replicate the school choice game from Chen and Sönmez (2006) to test the effect of strategy advice on participants' strategy choices in the DA and IA mechanisms. The setup is relatively complex compared to other laboratory experiments and aims to mimic applications of the mechanism. We implement a 2 $\times$ 2 experimental design: an advice treatment and a no advice treatment for each mechanism. The no advice treatment is an exact replication of Chen and Sönmez's experiment and the advice treatment is the replication plus advice.

3.3.1 The Game

Students compete for one of 36 slots at seven schools in a one-shot game. Based on the students' reported preferences over schools, the schools' priorities over students, and the tiebreaker, the mechanism assigns one student to each slot. No students are unmatched. Student preferences are determined by the monetary payoff to the participant. The payoffs range from \$19 for assignment to the student's first-ranked school to \$5 for assignment to their last-ranked school. We use the designed environment from Chen and Sönmez,[7] but increase the payoffs by \$3. Table 3.1 shows a sample payoff matrix.

Table 3.1: Sample Payoff Matrix

Slot Received at School	A	B	C	D	E	F	G
Payoff (in \$)	12	19	5	16	14	8	10

Schools have two priority levels: in-district and not in-district. Students are assigned to a school's district according to Table 3.2. Schools A and B each have three slots and schools C, D, E, F, and G each have six slots, matching the number of students in each district. Since there are only two priority levels, ties are broken by a random draw of bingo balls after students submit their selections.

Table 3.2: District School Assignments

Students	1-3	4-6	7-12	13-18	19-24	25-30	30-36
District School	A	B	C	D	E	F	G

Students are informed of their payoffs, each school's quota and priorities, and the mechanism's algorithm in the instructions. The only information students do not have are the preferences of other students and the outcome of the tiebreaker.

3.3.2 Implementation

We implemented the experiment at Georgia State University's Experimental Economics Laboratory in November 2016. We recruited participants for eight sessions that took place over two days. There were two sessions per treatment and exactly 36 participants participated in each session. Table 3.3 shows the number of participants in each experimental treatment.

Table 3.3: Number of Participants by Treatment

	DA	IA
No Advice	72	72
Advice	72	72

We used the instructions verbatim from Chen and Sönmez (2006) (provided in Appendices N and O). The experimenter distributed paper copies of the instructions. Participants filled out paper answer sheets, which the experimenter collected and entered into a Python solver

[7]The designed environment, as opposed to the random environment, correlates preferences to school proximity (lexicographically) and school quality (proxied by smaller quotas).

while participants completed a survey about the experiment on lab computers (provided in Appendix P). In one section of the post-experiment survey, participants answered four questions about a logic puzzle. Participants could earn $0.50 for each correct answer for up to $2 total. Participants were paid $20.28 on average at the end of the experiment in cash, including a $5 show up fee.

3.3.3 The Advice

Participants in advice treatment sessions received strategy advice as an additional page at the end of the instructions packet (Appendices Q and R). We frame the advice as the way to "obtain the highest possible payout" and include an illustration of the advice using the example from the instructions. The advice to participants in the DA mechanism is to tell the truth. "In other words, you should rank the schools in the order of your payoffs, from high to low."[8]

We advise participants in the IA mechanism that “you will **not necessarily** obtain the highest possible payoff” by telling the truth (emphasis included in the advice). The advice recommends one of two strategies. The first strategy suggests listing their district school first. For the remainder of this paper, we call this strategy the “Safe” strategy. The second possible strategy is characterized as riskier and recommends listing their top preference first and district school second. We call this strategy the “Risky” strategy. Note that almost any strategy can be justified by a participant’s preferences and beliefs, so these are only two of many strategies that are undominated. We chose these two strategies for their prominence in real-world applications of assignment mechanisms and in the experimental literature.

3.4 Strategy Advice Results

Table 3.4 lists the results of our experiment by treatment, where "NA" is the no advice treatment and "A" is the advice treatment. To interpret these results, first note that "Plays Truthfully" has a slightly different meaning for the DA versus IA mechanisms. In DA, playing truthfully measures how many participants truthfully report their preferences by ordering schools according to their monetary payoff *through their district school.* In other words, playing truthfully means submitting a ranking of schools from highest payoff to lowest payoff, through the participant’s district school. Participants are guaranteed admission to

[8]The example in our DA advice has a typo in the last paragraph: “student 3 had applied to Huron” should say Ontario instead. Thank you to an anonymous reviewer for pointing it out. There were no questions about the advice during the experimental sessions and no negative comments about the advice in the post-experiment survey. We also solicit feedback on the advice in the post-experiment survey and find that 81% of participants said the advice was “clear” or “very clear” and 83% said the advice was “good” or “very good.”

their district school in DA, [9] so any rankings below the district school are irrelevant to the assignment mechanism. In IA, participants are not guaranteed admission to their district school unless they rank it first; therefore, playing truthfully means truthfully reporting all seven schools.[10,11]

Table 3.4 also lists the percentage of participants in IA treatment sessions that play the Safe strategy and the Risky strategy. As described above, participants rank their district school first in the Safe strategy, guaranteeing their assignment. Participants list their top-ranked school first and district school second in the Risky Strategy. Last, we report the number of participants who say in the post-experiment survey that they tried to follow the advice.[12] The remainder of this section discusses first the DA results and then the IA results.

Table 3.4: Strategy Advice Results by Treatment Group

	IA-NA	IA-A	DA-NA	DA-A
Plays Truthfully	26%	8%	31%	50%
IA Safe Strategy	43%	53%		
IA Risky Strategy	17%	31%		
Tries to Follow Advice		83%		89%
Observations	72	72	72	72

3.4.1 Deferred Acceptance

The weakly dominant strategy in DA is to truthfully reveal preferences. Our strategy advice explains that ordering their schools from highest to lowest payoff leads to the highest possible payoff and provides an example to show why that is the case.

Hypothesis 5. *Participants are more likely to choose the dominant strategy of truth-telling when it is recommended by strategy advice in the DA mechanism.*

[9] A student is high priority at their district school. No matter where they rank their district school, they are admitted if they apply. There are the same number of slots at each school as there are district residents, so even if all high priority students apply to the district school, each are admitted. If a student ranks their district school first, they are matched to it automatically. If a student ranks their district school third, they are matched to that school if both their first and second choices reject them.

[10] The only exception is to students whose district school is also their most preferred. In this case, participants need only truthfully report their district school in the top ranking and all of them do. We classify them as following both strategies.

[11] The percentage of DA strategies that truthfully reveal all seven schools are 17% for the no advice treatment (or about 55% of the 31% listed as "Plays Truthfully") and 43% for the advice treatment (or about 86% of the 50% listed as "Plays Truthfully").

[12] Participants chose "Yes" or "No" in response to the question "Did you try to follow the additional advice you were given?"

We find that participants in the DA advice treatment sessions are 19 percentage points more likely to truthfully reveal their preferences than in the DA no advice treatment (50% compared to 31%, respectively). A one-sided test of proportions rejects the null hypothesis that there is no difference in proportions at the 1% significance level (p-value= 0.0087 in a one-sided test).

Result 5. *The truth-telling rate is significantly higher in the advice treatment than in the no advice treatment in the DA mechanism.*

This treatment effect of advice is large compared to the low baseline rate of truthful revelation. Fewer than one third of participants truthfully reveal their preferences in the no advice treatment, despite the fact that we do not impose truth-telling beyond the district school. Even when we advise participants that they should tell the truth, only 50% do so. This low figure is not due to a lack of trust in the advice. 89% of participants report trying to follow the advice in the post-experiment survey. The problem appears to be in either understanding or implementing the advice. Of those who say they tried to follow the advice, only 56% successfully do so.[13]

While the treatment effect is large, the proportion of people who truthfully reveal their preferences is still low relative to the truth-telling rates in other school choice laboratory experiments. Exactly half of participants truthfully reveal their preferences in the advice treatment, which is still 22 percentage points lower than in the experiment we replicate where there is no advice (Chen and Sönmez, 2006). Other lab experiments find rates ranging from 44% (Klijn et al., 2013) to 79% (Basteck and Mantovani, 2018). This result suggests that baseline truth-telling rates are likely heterogeneous across populations.

3.4.2 Immediate Acceptance

There is no dominant strategy in IA. We offer strategy advice that warns participants against telling the truth and suggest one of two heuristic strategies as alternatives. There are two ways in which participants can follow the strategy advice. First, participants can implement one of the two strategies we recommend.

Hypothesis 6. *Participants are more likely to choose a heuristic strategy when it is recommended by strategy advice in the IA mechanism.*

The second way that participants can follow the advice is to *not* truthfully reveal their preferences.

[13]See Section 3.5.1 for a breakdown of the strategies implemented by participants who did not choose the dominant strategy.

Hypothesis 7. *Participants are less likely to truthfully reveal their preferences when advised against it by strategy advice in the IA mechanism.*

In the IA advice treatment, we find that 10 percentage points more participants (over a base of 43%) choose the Safe strategy and almost double the proportion of participants (14 percentage points on a base of 17%) choose the Risky strategy. A one-sided test of proportions confirms a statistically significant increase in the percentage of Risky strategies at the 5% level (p-value= 0.0249 in a one-sided test), but the difference is not significant at traditional levels for the Safe strategy (p-value= 0.1215 in a one-sided test). The total number of participants who choose one of the heuristic strategies increases 27 percentage points, from 51% to 78%, in the advice treatment.

Result 6. *The proportion of participants who choose a heuristic strategy is significantly higher in the advice treatment than in the no advice treatment. More of the increase is due to inducing participants to choose the Risky strategy than the Safe strategy.*

We instruct participants that it may not be in their best interest to tell the truth and almost all participants act in accordance with this advice. Excepting the strategy profiles in which their district school is also their top-ranked school[14], only two participants truthfully reveal their preferences in the IA Advice sessions. The decrease in truth-telling rates from 26% to 8% is statistically significant at the 1% level (p-value= 0.0001 in a one-sided test).

Result 7. *The truth-telling rate is significantly lower in the advice treatment than in the no advice treatment of the IA mechanism.*

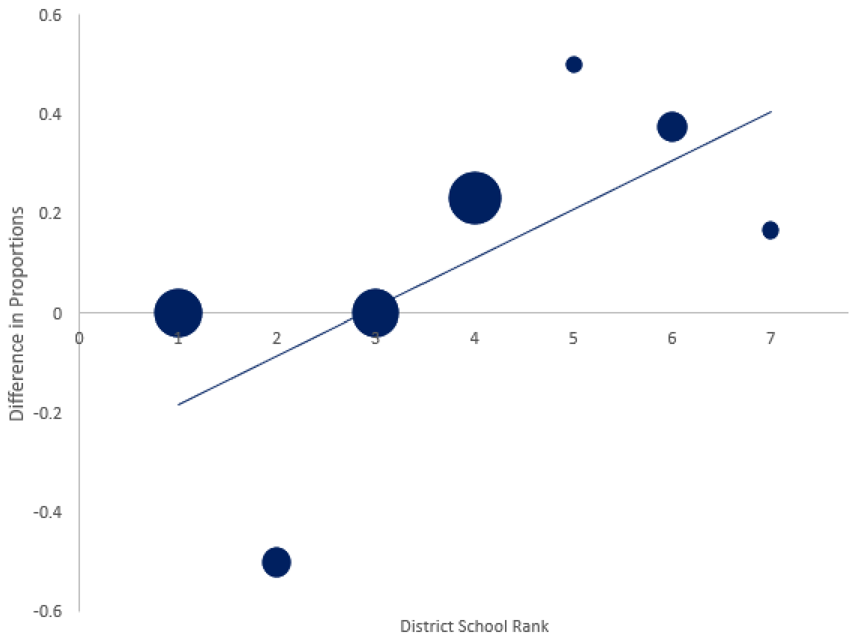
Participants who report trying to follow the advice are also much more likely to successfully do so in IA versus DA. Slightly fewer participants (83%) report trying to follow advice compared to DA, but 83% of those who try to follow the advice successfully do so by implementing one of the two recommended strategies.

Participants respond to the strategy advice heterogeneously by the ranking (in terms of payoff) of their district school. The increase in the proportion of participants who choose the Risky strategy is entirely due to participants with their district school ranked fourth or lower. Figure 3.1 graphs the difference in the proportion of participants who choose the Risky strategy in the advice treatment versus the no advice treatment by district school payoff ranking.

On the other hand, only participants with district schools ranked third or fourth contribute to the increase in the percentage of participants who choose the Safe strategy in the

[14]There are two of these strategy profiles in each session. Each of these participants chose the dominant strategy for their profile of listing their district school first.

Figure 3.1: Difference in Proportion of Participants in Advice and No Advice Treatment who Play Risky Strategy by District School Payoff Ranking



Notes: The size of the bubble represents how many participants of each district school rank played the Risky strategy.

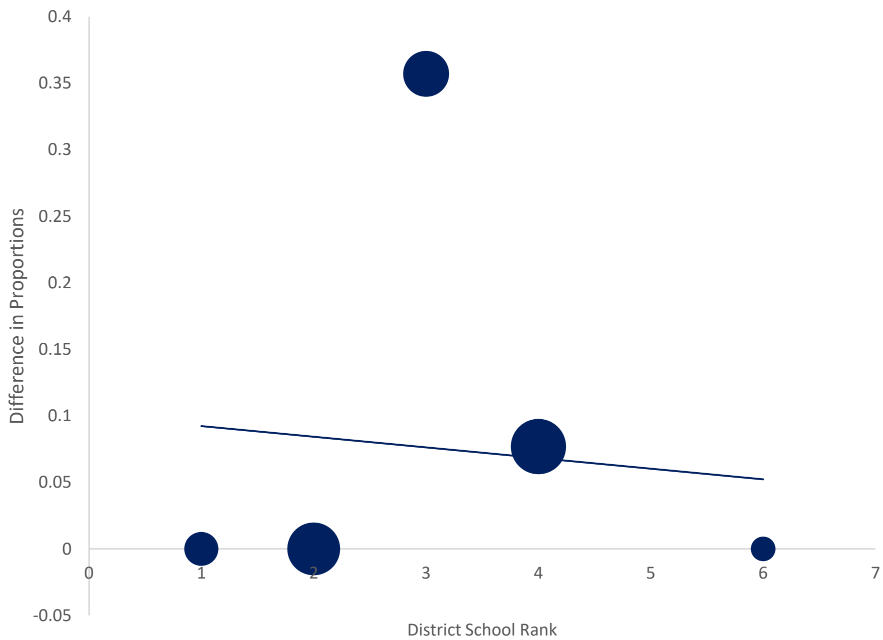
advice treatment. Figure 3.2 graphs the difference in the proportion of participants who choose the Safe strategy in the advice treatment versus the no advice treatment by district school payoff rank. Note that (fortunately) no participants with district school ranked seventh played the Safe strategy in either session.

Compared to Chen and Sönmez (2006), our replication finds a higher rate of truth-telling in the IA mechanism. In our no advice treatment, which replicates Chen and Sönmez's experiment, we find that 26% compared to their 14% of participants truthfully reveal their preferences. Again, truth-telling rates and strategic behavior in general may vary across populations.

3.4.3 Not Following the Advice

The participants who did not implement the recommended strategy may have been rational to do so or the incentives we offered may have been too low to induce participants to care about them. To address both of these concerns, we calculate the expected payoff from playing

Figure 3.2: Difference in Proportion of Participants in Advice and No Advice Treatment who Play Safe strategy by District School Payoff Ranking



Notes: The size of the bubble represents how many participants of each district school rank played the Safe strategy.

each of the recommended strategies for those who do not follow the advice.[15] To calculate this expected payoff, as well as the expected payoff to the empirical strategy the participant used, we use recombinant estimation to smooth over two sources of randomness. The first source of randomness is session effects that arise from a participant being in one session of a treatment rather than another. Second, recombinant estimation averages out a participant's good or bad luck from a particular tiebreaker.[16]

We follow the recombinant estimation technique from Reiley and Mullin (2000) to calculate expected payoffs. Define a student profile ρ_i $\forall$ $i = \{1, ..., n\}$, where n is 36 in our experiment, as a list of student i's preference profile and district school. For each of our eight sessions (two sessions each for four treatments), one participant is assigned to each

[15]Note that we cannot calculate something comparable for participants who play one of the recommended strategies because we do not know the counter-factual strategy.

[16]Recall that the participants do not know the result of the tiebreaker until after they have submitted their preferences.

of the 36 student profiles. So, for every student profile ρ_i, we observe two strategies per treatment. The recombinant technique addresses the first source of randomness, session effects, by randomly choosing one of those two strategies for all student profiles except for the participant of interest. Recombinant estimation addresses the second source of randomness, the tiebreaker, by drawing a new tiebreaker for each recombination.

The recombinant estimation proceeds as follows:

1. Fix the strategy Q_i of the participant of interest with student profile ρ_i.
2. For each other student profile $\rho_j \; \forall \; j \in \{1, ..., n\}/i$, draw an observed strategy Q_j from one of the two treatment sessions.
3. Draw a tiebreaker.
4. Implement the mechanism ϕ and record the payoff of the participant of interest $\phi_i(Q_i, Q_{-i})$.
5. Repeat for r recombinations.

For example, suppose we fix participant 1's strategy. Then, we draw one of the strategies played by a participant with student profile 2 from either session one or session two of the treatment. We do the same for student profile 3, ..., through student profile 36. Then, we calculate the payoff to participant 1 by drawing a tiebreaker and implementing the mechanism. We repeat this process for the number of desired recombinations to obtain an estimate of the participant's expected payoff, given the play of other participants in this experiment.

We list the average difference between the expected payoff from playing each recommended strategy and the observed strategy for each participant who did not choose any of the recommended strategies in Table 3.5. Note that there are no statistical tests because standard errors are driven to zero by the recombinations.

As expected, a participant would always be better off playing the dominant strategy of truth-telling in DA. The size of the difference is economically significant. For example, the \$1.80 average gains from telling the truth in the no advice treatment is approximately 13% of the total amount a participant could earn.[17] The expected payoff difference increases to \$1.92 in the advice treatment.

In stark contrast, participants who chose not to implement one of the heuristic strategies in IA would have had much worse outcomes playing a recommended strategy. For example, the cost of playing the Safe strategy in the IA no advice treatment to a participant who chose not to play either of the heuristic strategies is \$2.78, or about 20% of the total dollars

[17] Participant payoffs ranged from \$5 to \$19, for a total of \$14 that a participant could earn apart from guaranteed payments.

Table 3.5: Recombinant Estimation of Difference between Expected Payoffs from Empirical Strategy and Recommended Strategy

	Risky		Safe		Truthful	
	IA-NA	IA-A	IA-NA	IA-A	DA-NA	DA-A
Mean (in $)	-1.32	-2.53	-2.78	-3.43	1.80	1.92
Std. Dev.	4.07	2.81	4.26	3.17	2.21	2.46
Number	33	16	33	16	50	36

Notes: Expected payoff to each strategy is calculated from 10,000 recombinations per participant. We report the standard deviation of expected payoff differences between participants who did not play one of the recommended strategies and the number of participants in each treatment who did not play one of the recommended strategies. There are no statistical tests because standard errors are driven to zero by the recombinations.

a participant could earn. This result suggests that the participants who chose not to play one of the recommended strategies in IA appear to have been rational in doing so.

3.5 Mechanism Performance under Sub-Optimal Strategies

To evaluate mechanism performance under sub-optimal play, we use the recombinant technique detailed above to estimate the efficiency and stability of DA and IA in the no advice treatments versus the advice treatments. Table 3.6 lists the recombinant estimation of efficiency for each treatment.

Contrary to theoretical predictions (Abdulkadiroğlu et al., 2011; Miralles, 2009; Troyan, 2012), but consistent with other experiments (e.g. Chen and Sönmez, 2006, Calsamiglia et al., 2010), we find that DA is more efficient than IA in our experiment. The difference between the two mechanisms in the no advice treatments ($13.95 to $13.72) is about half the size Chen and Sönmez found ($11.71 to $11.15). In the advice treatment, however, the gap nearly triples ($13.87 to $13.23). While we show that DA is more efficient than IA when participants in both mechanisms choose sub-optimal strategies,[18] it is important to note that the differences in efficiency between our treatments are small relative to the difference between efficiency in our experiment and efficiency under optimal strategies. When we estimate efficiency under universal truth-telling, IA outperforms DA as theory shows it must ($15.55 to $15.07), but efficiency drops considerably for both mechanisms as sub-optimal strategies are added. The larger decrease in IA suggests that sub-optimal strategies have a more pronounced effect on the efficiency of IA compared to DA.

[18] Even though we do not know which strategies are sub-optimal in our data, we know that at least one is sub-optimal because the resulting matching is less efficient than the DA matching.

Table 3.6: Recombinant Estimation of Average Per Capita Payoff

	IA*	IA-NA	IA-A	DA*	DA-NA	DA-A
Mean	15.55	13.72	13.23	15.07	13.95	13.87
Std. Dev.	0.11	0.393	0.238	0.24	0.353	0.333

Notes: IA* and DA* are average per capita payoffs under optimal strategies. The mean is calculated from 10,000 recombinations per participant in each treatment for a total of 72,000 total assignments. We report the standard deviation of the average per capita payoff between replications. There are no statistical tests because standard errors are driven to zero by the recombinations.

A within-mechanism comparison of treatments shows that the greater efficiency of DA is driven more by IA's loss of efficiency than DA's efficiency gains from fewer sub-optimal strategies in the advice treatment. In fact, in DA, we observe a small efficiency decline in the advice treatment ($13.95 to $13.87). To the contrary, in IA, the decline in efficiency from the no advice treatment to the advice treatment is nearly $0.50. That is, when more participants are induced to choose a heuristic strategy, the average welfare of participants declined by 3.5% of the total they could earn.

When we turn to stability, DA again outperforms IA. Table 3.7 lists the average number of blocking pairs from the recombinant estimation— the higher the number of blocking pairs (i.e. the more students who are justifiably envious), the less stable the assignment is. In the no advice treatment, IA has approximately 50% more blocking pairs than DA. Since DA is theoretically stable and IA's Pareto efficiency is inconsistent with stability (Abdulkadiroğlu and Sönmez, 2003), we expect this result; however, when participants choose sub-optimal strategies, DA could be less stable than IA in practice. Although the DA matchings are not completely stable in our experiment, we find that DA is still more stable than IA in the presence of sub-optimal strategies.

Table 3.7: Recombinant Estimation of Average Number of Blocking Pairs

	IA, NA	IA, A	DA, NA	DA, A
Mean	13.53	15.10	9.42	7.55
Std. Dev.	2.65	3.52	2.64	2.01

Notes: The mean is calculated from 10,000 recombinations per participant in each treatment for a total of 72,000 total assignments. We report the standard deviation of the average per capita payoff between replications. There are no statistical tests because standard errors are driven to zero by the recombinations.

Strategy advice has an opposite effect on the stability of DA and IA. While the number of blocking pairs increases by about 1.5 in the IA advice treatment, the number of blocking pairs decreases by approximately 2 in the DA advice treatment.

In our experiment, we find that DA outperforms IA in welfare measures that average over all participants. Furthermore, our strategy advice exacerbates the difference in performance between the two mechanisms. We turn next to the welfare of individual participants. To evaluate the relative welfare of participants, we must formulate counter-factual strategies since each participant only participates in one mechanism. We do so by developing a new typology to partially order sub-optimal strategies according to their relative dominance in DA.

3.5.1 A *Dominance-Ordered* Typology of Strategies in DA

We characterize the dominant strategy in DA as the strategy resulting from understanding three strategic aspects of DA. The first strategic aspect is that a participant's district school acts as a "safety" school.[19] That is, a participant can never be assigned to a school they rank lower than their district school. A participant who understands this aspect of DA should therefore never rank a school they prefer less than their district school above their district school. Note that the student's reported preferences Q_i is the strategy, so we refer to them as such in this section.

District-Consistent: A strategy Q_i is *District-Consistent* for student i if Q_i does not rank any school above i's district school that their true preferences P_i rank below i's district school.

The second strategic aspect of DA is that a participant can never lose from truthfully ordering the schools they rank above their district school. Note that a participant also cannot lose from truthfully ordering the schools they rank below their district school, but that ordering does not matter strategically.[20]

Ordered: A strategy Q_i is *Ordered* for student i if the schools Q_i ranks above i's district school are ranked in the same order as in their true preferences P_i.

The last strategic aspect in our typology of DA is that, when a student's strategy is Ordered, the student can never lose from ranking more schools they prefer to their district school above their district school. When their strategy is *not* Ordered, a participant can also always improve their strategy by ranking one additional school they prefer to their district

[19]This is specific to the school choice environment where every district school has enough seats to accept all students in the district.

[20]As noted earlier, a participant can never be assigned to a school they rank lower than their district school.

school *directly above* their district school. Overall, it is always possible for a student to improve a strategy by ranking more schools they prefer to their district school above their district school.

x**-Strategy:** If a_i is the number of schools student i prefers to their district school, then for any $x \in \{1, \ldots, a_i\}$, Q_i is an x*-Strategy* if Q_i ranks x schools student i prefers to their district school above i's district school.

We denote District-Consistent strategies as DC and NDC otherwise. Ordered strategies are O and not-Ordered strategies as NO. Then a District-Consistent, Ordered, x-Strategy is denoted $Q_i^{\mathrm{DC|O|}x}$. The example below illustrates the typology.

Suppose that i's preferences are $P_i : s_1\ s_2\ s_3\ \underline{s_4}\ s_5\ s_6\ s_7$, where i's district school s_4 is underlined. Then $a_i = 3$ for this i and P_i. Some example strategies and their typology are:

- Strategy $Q_i^{\mathrm{DC|O|3}} : s_1\ s_2\ s_3\ \underline{s_4}\ s_7\ s_6\ s_5$ is a 3-strategy that is both District-Consistent and Ordered. It is one of i's dominant strategies.
- Strategy $Q_i^{\mathrm{DC|NO|3}} : s_2\ s_1\ s_3\ \underline{s_4}\ s_5\ s_6\ s_7$ is a 3-strategy that is District-Consistent, but not Ordered.
- Strategy $Q_i^{\mathrm{DC|O|2}} : s_1\ s_3\ \underline{s_4}\ s_5\ s_2\ s_6\ s_7$ is a 2-strategy that is both District-Consistent and Ordered.
- Strategy $Q_i^{\mathrm{NDC|NO|2}} : s_3\ s_5\ s_1\ \underline{s_4}\ s_5\ s_2\ s_6\ s_7$ is a 2-strategy that is neither District-Consistent nor Ordered.
- Strategy $Q_i^{\mathrm{NDC|NO|2}} : s_3\ s_5\ \underline{s_4}\ s_5\ s_2\ s_1\ s_6\ s_7$ is a 2-strategy that is Ordered but not District-Consistent.

Observe that all DC|O|a_i are dominant strategies. DC|O|0 are District-first strategies in which a participant ranks their district school first. Strategies DC|O|x for $x \in \{1, \ldots, a_i - 1\}$ exhibit what the literature calls *district school bias* (Chen and Sönmez, 2006). Note that a 1-strategy or a 0-strategy that is also District-Consistent is necessarily Ordered. As a consequence, there are no DC|NO|1 and DC|NO|0 strategies (the lowest x for which DC|NO|x strategies exists is $x = 2$).

The strategies in our typology are connected through dominance-based partial orderings. For any two sets of strategies X and Y, let $X \to Y$ indicate that *all* strategies in X weakly dominate all strategies in Y. Let $X \dashrightarrow Y$ indicate that, for *every* strategy $y \in Y$, *there exists* a strategy in X that weakly dominates Y. Observe that $X \to Y$ and $Y \dashrightarrow Z$ imply $X \to Z$, but $X \dashrightarrow Y$ and $Y \to Z$ only imply $X \dashrightarrow Z$ (and *not* $X \to Z$).

Table 3.8: Distribution of Strategies Played in DA

Strategy types	Advice	No-Advice	Random
Dominant strategies DC|O|a_i	50%	30.6%	1.3%
District-first strategies DC|O|0	15.3 %	27.8%	4%
Other DC|O|x strategies $\{$DC|O|1, $\ldots$, DC|O|$(a_i-1)\}$,	12.5%	15.3%	25.7%
DC|NO|x strategies $\{$DC|NO|2, $\ldots$, DC|NO|$a_i\}$	12.5%	15.3%	21%
Other strategies	9.7 %	11%	48%

Notes: The percentages in the last column are the sample averages over 10,000 draws of randomly drawn reported preferences and random student profiles (among the 36 profiles in our experimental design).

Proposition 1 (Dominance relations between sets of strategies in DA)**.** *For any student i and any preferences P_i, the following binary relations hold:*

$$\begin{tikzcd}
\mathrm{DC|O|}a_i \arrow[r] \arrow[d] & \mathrm{DC|O|}(a_i-1) \arrow[r, dashed] \arrow[d, dashed] & \ldots \arrow[r, dashed] \arrow[d, phantom, "\cdots"] & \mathrm{DC|O|}2 \arrow[r, dashed] \arrow[d, dashed] & \mathrm{DC|O|}1 \arrow[r, dashed] & \mathrm{DC|O|}0 \\
\mathrm{DC|NO|}a_i \arrow[r, dashed] & \mathrm{DC|NO|}(a_i-1) \arrow[r, dashed] & \ldots \arrow[r, dashed] & \mathrm{DC|NO|}2
\end{tikzcd}$$

Classifying strategies by this typology can reveal which strategic aspects of DA participants do, and do not, understand. To this end, Table 3.8 compares the distribution of strategies played in our advice and no advice DA treatments using this typology. The last column of Table 3.8 shows the distribution of strategies based on randomly drawn reported preferences and a randomly drawn student type from the 36 profiles in our experimental design.

Table 3.8 shows that, although participants often fail to choose the dominant strategy, about 90% of the participants understand enough about the mechanism to play a District-

Consistent strategy in both treatments. Participants choosing randomly would only choose District-Consistent strategies about half of the time. Furthermore, the majority (about 75%) of participants choose an Ordered strategy. These distributions suggest that 1) participants in our experiment are not playing randomly and 2) most understand the "safety" school and truthful ordering aspects of DA. Most participants who choose a sub-optimal strategy appear to fail to understand the x-Strategy aspect of DA. In other words, they exhibit district school bias.

Our experiment was not designed to study the impact of strategy advice on the distribution of sub-optimal strategies in DA. It is informative, however, to note that the advice does not substantially increase the number of participants who play District-Consistent strategies. Rather, the advice seems to steer participants who would have played a District-Consistent strategy anyways, particularly a District-first strategy, toward the dominant strategy of truth-telling.

The remainder of this section considers the individual welfare of participants who chose a heuristic strategy in IA. We build a series of increasingly irrational, though District-Consistent[21], counter-factual strategies in DA for the participant to play using the dominance-ordered typology in this section. Then, we use recombinant estimation to estimate expected payoffs to each strategy in their respective mechanisms. We first consider the participants who chose the Safe strategy in IA.

3.5.2 Safe Strategy

A participant who chooses the Safe strategy in IA is always assigned to their district school. In DA, the same participant is assured a seat at a school they like *at least as well* as their district school by reporting a District-Consistent strategy, regardless of whether the strategy is Ordered $o \in \{O, NO\}$. If, in addition, the participant ranks at least one school they prefer to their district school above their district school, then there is a positive probability that they are assigned to a school they like *strictly better* than their district school in DA. We denote the weak preference relationship for student i as R_i.

Proposition 2. *(i) For any preferences P_i, any District-Consistent strategy $Q_i^{DC|o|x}$, any Safe strategy Q_i^{Safe}, any strategies played by students $j \in \{1, ..., n\}/i \quad Q_{-i}$, and any priority profile that is drawn with positive probability,*

$$DA(Q_i^{DC|o|x}, Q_{-i}) \ R_i \ IA(Q_i^{Safe}, Q_{-i}).$$

[21] This assumption seems reasonable since approximately 90% of participants play District-Consistent strategy.

(ii) If, in addition, $x \neq 0$, then there exists $\overline{Q}_i^{DC|o|x}$ such that for any priority profile that is drawn with positive probability,

$$DA(\overline{Q}_i^{DC|o|x}, Q_{-i})\ P_i\ IA(Q_i^{Safe}, Q_{-i}),$$

Proposition 2 says that any student playing the Safe strategy in IA can be no worse-off in DA as long as they play a District-Consistent strategy, which around 90% of our participants do (see Table 3.8). In addition, if the student plays a District-Consistent strategy with at least one school ranked above their district school, they could be strictly better-off in DA. How much and how often participants are *strictly* better-off under DA compared to IA depends on (a) the preferences reported by other students (Q_{-i}), and (b) the "level of rationality" of the strategy they implement in DA.[22] We use recombinant estimation to empirically estimate this welfare gain with our experimental data.

Quantifying Proposition 2 provides a measure of the welfare gain to a participant who chooses the Safe strategy in IA of participating in the DA mechanism instead. We use the same recombinant estimation technique described above to determine a participant's expected payoff for each possible District-Consistent strategy. Specifically, for each participant who chose Q_i^{Safe} in the IA Advice treatment, we implement the following procedure:

1. Set $Q_i = Q_i^{\mathrm{DC}|o|x}$ for $o \in \{\mathrm{O,NO}\}$ and $x \in \{1, ..., a_i\}$.

2. For each other participant profile Q_{-i}, draw one of the observed strategies from the DA Advice treatment.[23]

3. Draw a tie breaker.

4. Implement the DA mechanism and record the payoff of the participant of interest $DA_i(Q_i, Q_{-i})$.

We repeat this procedure 50,000 times for each x,o combination. Note that we do not need to estimate the participant's payoff from playing the Safe strategy in IA since it will always be the payoff from their district school.

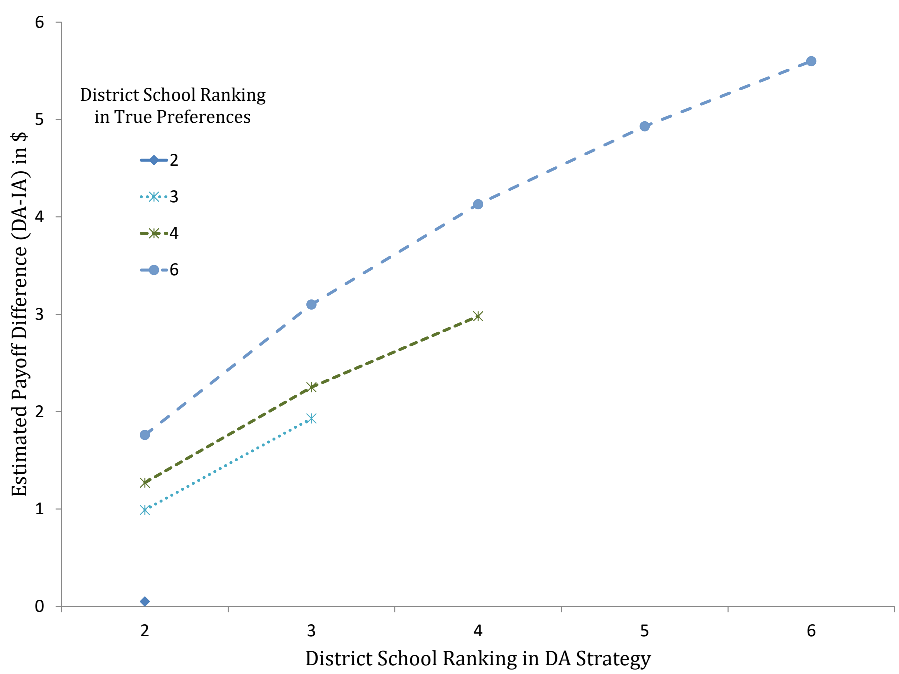
Figures 3.3 and 3.4 show the average welfare gain for each Ordered and Not-Ordered strategy by the participant's district school rank. The x-axis tracks the rank of the district

[22] Suppose that A and B are two sets of District-Consistent strategies for student i, and $A \dashrightarrow B$. By definition, for any strategy in $Q_i^B \in B$, there exists a strategy $Q_i^A \in A$ that weakly dominates Q_i^B. That is, student i is better-off more often in DA compared to IA when playing Q_i^A than when playing Q_i^B (and i is never worse-off under Q_i^A than under Q_i^B).

[23] Out of two, since there are two sessions for each treatment. See Appendix S for results using the No Advice treatments.

school in the counter-factual DA strategy. For example, for a participant whose district school is ranked fourth (the green starred line), the first point on the x-axis corresponds to the welfare difference between playing a counter-factual strategy with their district school ranked second in DA and playing the Safe strategy in IA. The second point corresponds to a counter-factual strategy with their district school ranked third and the third point to the dominant strategy. Participants with lower ranked district schools have more points on the graph because there are more possible District-Consistent strategies. There are fewer points on the Not-Ordered graph because District-Consistent 1-strategies are necessarily ordered.

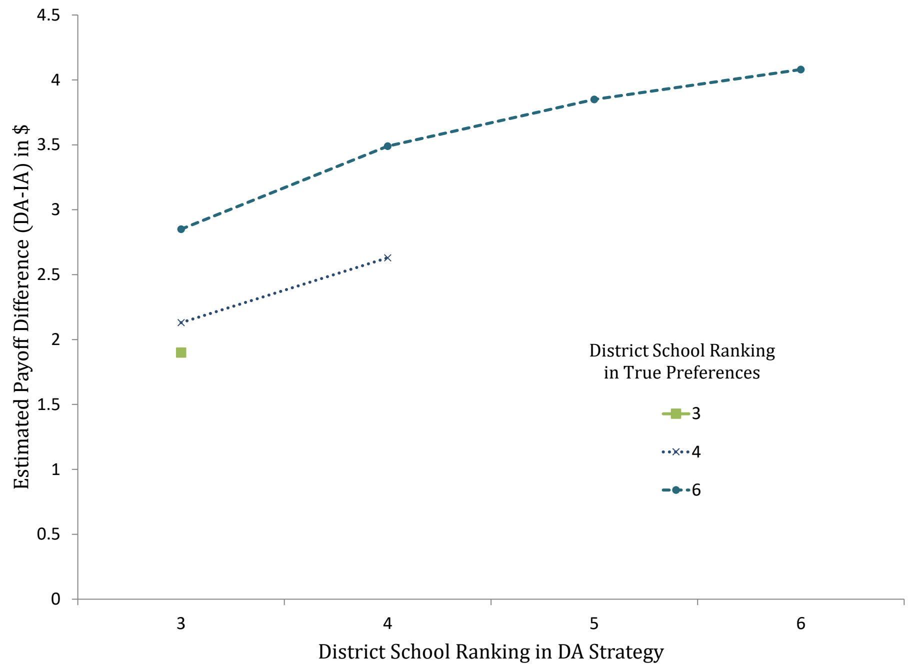
Figure 3.3: Expected Payoff Difference between District-Consistent Ordered Strategy in DA and Safe Strategy in IA



Notes: Each line traces the estimated welfare change to participants with the same district school rank who chose the Safe strategy in an IA session of playing a District-Consistent Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

The graphs show that participants who chose to play the Safe strategy in IA would have been unequivocally better off in DA. Note that all points are above the x-axis, meaning that all welfare gains are positive. We also find empirical evidence for Proposition 1 by observing the increasing welfare gains from DC|O|1 or DC|NO|2 to DC|O|a_i. Note also that

Figure 3.4: Expected Payoff Difference between District-Consistent Not-Ordered Strategy in DA and Safe Strategy in IA



Notes: Each line traces the estimated welfare change to participants with the same district school rank who chose the Safe strategy in an IA session of playing a District-Consistent Not-Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

the largest welfare gains are to those participants with district schools ranked lower in their true preferences. In other words, the most disadvantaged participants in the game have the most to gain from switching to any District-Consistent strategy in the DA mechanism from the Safe strategy in the IA mechanism.

3.5.3 Risky Strategy

Theoretical results are weaker for the Risky strategy. In principle, a participant could be better-off playing the Risky strategy in IA than playing *any* District-Consistent strategy in DA, including a dominant strategy. It is also possible for a participant to be better-off under a District-Consistent strategy in DA than under the Risky strategy in IA. Comparisons between the District-Consistent strategy in DA and the Risky strategy in IA are entirely dependent on other participants' reported preferences.

The only systematic advantage of District-Consistent strategies in DA over the Risky strategy in IA is in terms of the worst-case scenario: a District-Consistent strategy in DA guarantees a lower-bound on the participant's assignment that is at least as good as the lower-bound under the Risky strategy in IA. When the participant's district school is not their least or most preferred school, this statement can be strengthened to say that the lower-bound is strictly higher under any District-Consistent strategy in DA compared to the lower-bound under the Risky strategy in IA. In other words, the lower-bound assignment for the Risky strategy in IA is the participant's least preferred school. The lower-bound assignment for a District-Consistent strategy in DA is the district school. So long as these two do not intersect and the assignment is not trivial, the lower-bound in a District-Consistent strategy in DA is strictly higher than the lower-bound in the Risky strategy in IA.

Proposition 3. *(i) For any preferences P_i, any District-Consistent strategy $Q_i^{DC|o|x}$, any Risky strategy Q_i^{Risky}, and any priority profile that is drawn with positive probability,*

$$\min_{P_i}\{DA_i(Q_i^{DC|o|x}, Q_{-i}) \textit{ for all } Q_{-i}\} \; R_i \; \min_{P_i}\{IA_i(Q_i^{Risky}, Q_{-i}) \textit{ for all } Q_{-i}\}.$$

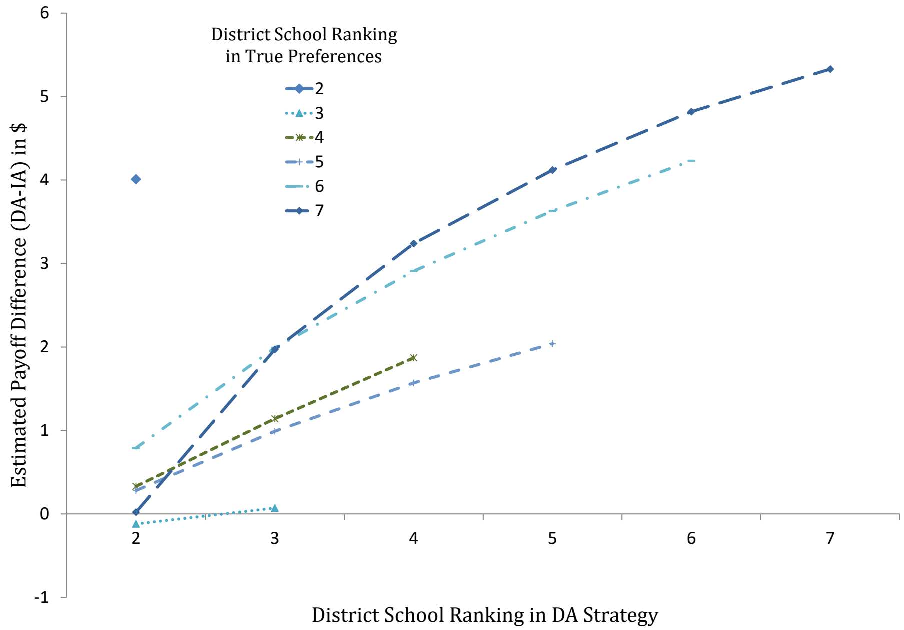
(ii) If in addition, i's district school is not their first- or least-preferred school, then

$$\min_{P_i}\{DA_i(Q_i^{DC|o|x}, Q_{-i}) \textit{ for all } Q_{-i}\} \; P_i \; \min_{P_i}\{IA_i(Q_i^{Risky}, Q_{-i}) \textit{ for all } Q_{-i}\}.$$

As the name suggests, participants take a chance in choosing the Risky strategy. By ranking a school above their district school, a participant risks losing their priority at their district school if their district school fills its quota in the first round. Then, participants who choose the Risky strategy may be assigned to a school they like less than their district school, which is impossible in DA for students who play a District-Consistent strategy in DA. It is even plausible — and we observe it in our simulations — for participants implementing the Risky strategy to be matched to their least preferred school when their district school is ranked considerably higher in their true preferences. At the same time, participants who have well-calibrated beliefs about other participants' strategies could obtain much higher payoffs than would otherwise be plausible in DA, particularly when other participants implement the Safe strategy in IA. We again turn to recombinant estimation to empirically evaluate whether participants who choose the Risky strategy in IA would be better off on average than by playing a District-Consistent strategy in the DA mechanism.

We calculate the welfare change of participants who chose the Risky strategy in an IA treatment of instead implementing a District-Consistent strategy in DA. We replicate the recombinant estimation of expected payoffs to playing a series of increasingly irrational strategies in DA from the previous subsection. Then, since the participant's payoff to playing the Risky strategy is not known ex-ante, we use recombinant estimation again to estimate the participant's expected payoff in IA. We repeat the procedure 50,000 times for each strategy.

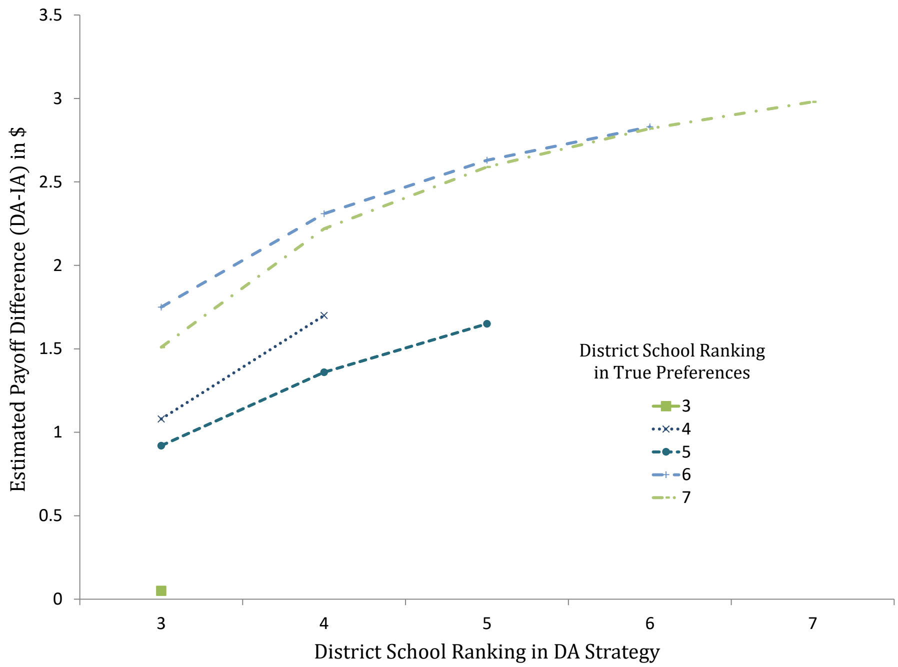
Figure 3.5: Expected Payoff Difference between District-Consistent Ordered Strategy in DA and Risky Strategy in IA



Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose the Risky strategy in an IA session of playing a District-Consistent Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

Figures 3.5 and 3.6 show that most participants would be better off playing any District-Consistent strategy in DA than playing the Risky strategy in IA. Participants whose district school is ranked third in their true preferences and choose an Ordered strategy stand out as the exception, but the size of the difference is small compared to the payoff increases that other participants expect to receive. We, again, observe that the greatest welfare gains of switching to the DA mechanism are realized by participants with low-ranked district schools.

Figure 3.6: Expected Payoff Difference between District-Consistent Not-Ordered Strategy in DA and Risky Strategy in IA



Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose the Risky strategy in an IA session of playing a District-Consistent Not-Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

3.5.4 Other Strategies

Almost any strategy can be justified in the IA mechanism by a participant's beliefs about the strategies of other players. Participants who choose strategies that correspond neither to our Risky or Safe strategy are not guaranteed their district school because they do not rank it first. The same as for the Risky strategy, the minimum payoff is lower in IA compared to DA for all participants whose district school is not ranked last. On the other hand, these other strategies may yield higher payoffs than playing even the dominant strategy in DA depending on the strategy choices of the other players and the tie breaker. We use recombinant estimation to evaluate how the "other" strategies we observe in our IA sessions compare to District-Consistent strategies in the DA mechanism.

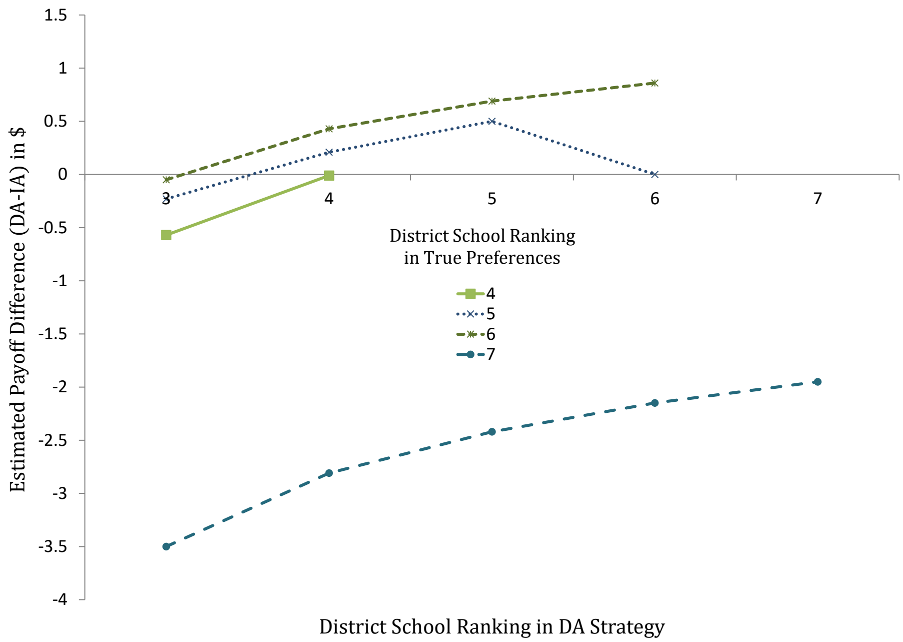
We calculate the welfare change of participants who choose neither the Risky nor Safe strategy in an IA session of instead implementing increasingly irrational District-Consistent strategies in a DA session. Our recombinant estimation, again, estimates the participant's expected payoff from their choosen strategy in IA and their expected payoff from playing a series of counter-factual strategies in DA. We repeat the procedure 50,000 times to estimate the welfare difference for each participant.

Figure 3.7: Expected Payoff Difference between District-Consistent Ordered Strategy in DA and Other Strategies in IA

Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose neither the Risky strategy nor the Safe strategy in an IA session of playing a District-Consistent Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

Figures 3.19 and 3.20 show mixed results for participants who chose non-heuristic strategies in IA. For participants with district schools ranked 5^{th} or 6^{th}, ranking more than two schools above their district school generates higher welfare in the DA mechanism; however, for participants with district schools ranked 4^{th} and 7^{th}, their welfare is near uniformly better playing their chosen strategy in an IA session. These contrary results suggest that, as would be expected, the expected welfare gains from participating in the DA mechanism compared to the IA mechanism is largely dependent on the participant's chosen strategy. For example, participants who ranked one of their middle-preferred schools first, when that school is

Figure 3.8: Expected Payoff Difference between District-Consistent Not-Ordered Strategy in DA and Other Strategies in IA



Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose neither the Risky strategy nor the Safe strategy in an IA session of playing a District-Consistent Not-Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

also always under-subscribed in the first round of IA, did particularly well relative to other participants with the same district school rank. In summary, these results show that while participants likely to choose a heuristic strategy fare worse in the IA mechanism, disadvantaged participants who play sophisticated strategies in the IA mechanism allows them to obtain higher payoffs compared to even the dominant strategy in DA.

3.6 Discussion

Our strategy advice decreases the proportion of sub-optimal strategies in DA. The increase in the proportion of participants who choose the dominant strategy in DA is large and statistically significant, but our strategy advice fails to achieve truth-telling rates close to the theoretically optimal 100%. Participants may be sensitive to the phrasing of the strategy advice, how it is communicated, and how trustworthy they find the person giving the advice. Future research integrating insights from behavioral economics and marketing research may shed light on how to produce more compelling strategy advice. On the other hand, the existence of sub-optimal play in the medical residency match (Rees-Jones, 2018) should temper any expectations of designing an information environment that induces optimal play for all participants.

The increase in the proportion of participants who choose one of the recommended heuristic strategies in the IA advice treatment is disconcerting when taken together with the overall welfare loss in that treatment. We cannot say for certain that the participants who were induced to choose one of the heuristic strategies in the advice treatment would have done better to not follow the advice. One, we do not know which participants would have played the heuristic strategy regardless of strategy advice and which were induced by the advice. Two, we do not know the counter-factual strategy the participant would have chosen. We do know that the strategy advice has a negative and significant effect on overall welfare. This result is concerning due to the prevalence of strategy advice like ours in the field.

A more important issue may be that we do not have a replacement for the "bad" advice that is currently pervasive. There is no generally "good" advice since the optimal strategy varies greatly based on participants' beliefs, preferences, and priorities. Moreover, journalists and bloggers are unlikely to stop giving strategy advice even though there is not universally good advice. This problem with strategy advice in IA adds to the evidence we find that favors the performance of DA relative to IA. In our experiment, DA is more efficient and more stable, even when only 31% of participants choose the dominant strategy. Overall, our results suggest that DA outperforms IA under sub-optimal strategies. In other words, while it may be theoretically possible for IA to outperform DA, we are unlikely to observe such an outcome in real-world applications that resemble the game played in this experiment.

There are important caveats to this conclusion. First, our efficiency analysis is inherently ex post; therefore, we do not address the theoretical result that IA ex-ante Pareto dominates any strategy-proof mechanism (Troyan, 2012). Second, our preferences do not vary in cardinality. Miralles (2009) and Abdulkadiroğlu et al. (2011) argue that IA may be more efficient due to varying levels of preference intensities. Those with higher preference intensities may take more risks when manipulating their preferences and obtain higher payoffs. We cannot address this issue with our current experiment and leave that to future research.

Our analysis of individual welfare shows that almost any participant who chose a heuristic strategy in IA would have preferred to walk into one of our DA sessions as long as they play a District-Consistent strategy (which almost all of our participants do). Participants who would benefit the most from switching to the DA mechanism from the IA mechanism are those who start the game at a disadvantage. These results have important equity implications for policymakers since they often emphasize the expected equity gains from implementing a school choice program. Disadvantaged students who are likely to play heuristic strategies in IA, particularly when these strategies are widely recommended, may actually be worse off under school choice.

Suppose a student in a disadvantaged area chose the Risky strategy. That student may end up at a worse school, one that is further away and not better, or even unassigned. If they do not have an outside option such as private school, the risk of being unassigned may make the gamble even higher stakes. Disadvantaged students, then, may be even more likely to play the Safe strategy. The Safe strategy, of course, leaves these disadvantaged students at the same school they would attend without school choice. The equity quality of the DA mechanism takes on greater meaning in light of these results.

The welfare gains from switching to DA only extend to some participants who chose neither the Risky nor the Safe strategy in the IA mechanism. Participants who choose sophisticated strategies, especially those at a particular disadvantage in our experiment, fare better playing their chosen strategies in IA. This result qualifies our conclusions for individual welfare in an intuitive way: participants who take advantage of the manipulability of the IA mechanism using a sophisticated strategy and have well-calibrated beliefs about other participants' strategies fare better in the less equitable mechanism.

References

Abdulkadiroğlu, A., Che, Y.-K., and Yasuda, Y. (2011). Resolving conflicting preferences in school choice: The "boston mechanism" reconsidered. *American Economic Review*, 101(1):399–410.

Abdulkadiroğlu, A., Pathak, P., Roth, A. E., and Sönmez, T. (2006). Changing the boston school choice mechanism. Technical report, National Bureau of Economic Research.

Abdulkadiroğlu, A. and Sönmez, T. (2003). School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747.

Aguiar, F., Brañas-Garza, P., Cobo-Reyes, R., Jimenez, N., and Miller, L. M. (2009). Are women expected to be more generous? *Experimental Economics*, 12(1):93–98.

Albrecht, K., Von Essen, E., Parys, J., and Szech, N. (2013). Updating, self-confidence, and discrimination. *European Economic Review*, 60:144–169.

Alston, M. (2019). The (perceived) cost of being female: An experimental investigation of strategic responses to discrimination. *Working paper*.

Amuedo-Dorantes, C. and De la Rica, S. (2006). The role of segregation and pay structure on the gender wage gap: evidence from matched employer-employee data for spain. *The BE Journal of Economic Analysis & Policy*, 5(1).

Andersen, S., Fountain, J., Harrison, G. W., and Rutström, E. E. (2014). Estimating subjective probabilities. *Journal of Risk and Uncertainty*, 48(3):207–229.

Babcock, L., Recalde, M. P., Vesterlund, L., and Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *American Economic Review*, 107(3):714–47.

Babin, J. J. (2019). Detecting group gender stereotypes: Opinion-mining vs. incentivized coordination games. *Journal of Economic Perspectives*, 45(1):21–42.

Barbulescu, R. and Bidwell, M. (2013). Do women choose different jobs from men? mechanisms of application segregation in the market for managerial workers. *Organization Science*, 24(3):737–756.

Basteck, C. and Mantovani, M. (2018). Cognitive ability and games of school choice. *Games and Economic Behavior*, 109:156–183.

Bayard, K., Hellerstein, J., Neumark, D., and Troske, K. (2003). New evidence on sex segregation and sex differences in wages from matched employee-employer data. *Journal of labor Economics*, 21(4):887–922.

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., and Doms, M. E. (2011). Women in stem: A gender gap to innovation. *Economics and Statistics Administration Issue Brief*, (04-11).

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

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2019). Beliefs about gender. *American Economic Review*, 109(3):739–73.

Braun, S., Dwenger, N., Kübler, D., and Westkamp, A. (2014). Implementing quotas in university admissions: An experimental analysis. *Games and Economic Behavior*, 85:232–251.

Brier, G. W. (1950). The statistical theory of turbulence and the problem of diffusion in the atmosphere. *Journal of Atmospheric Sciences*, 7(4):283–290.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American economic review*, 110(10):2997–3029.

Calsamiglia, C., Haeringer, G., and Klijn, F. (2010). Constrained school choice: An experimental study. *American Economic Review*, 100(4):1860–74.

Card, D., Cardoso, A. R., and Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, 131(2):633–686.

Cardoso, A. R., Guimarães, P., and Portugal, P. (2016). What drives the gender wage gap? a look at the role of firm and job-title heterogeneity. *Oxford Economic Papers*, 68(2):506–524.

Carrington, W. J. and Troske, K. R. (1998). Sex segregation in us manufacturing. *ILR Review*, 51(3):445–464.

Castillo, M. and Petrie, R. (2010). Discrimination in the lab: Does information trump appearance? *Games and Economic Behavior*, 68(1):50–59.

Charness, G., Cobo-Reyes, R., Meraglia, S., and Sánchez, Á. (2020). Anticipated discrimination, choices, and performance: Experimental evidence. *European Economic Review*, page 103473.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Chen, Y. and Sönmez, T. (2006). School choice: an experimental study. *Journal of Economic theory*, 127(1):202–231.

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

Crawford, V. P., Costa-Gomes, M. A., and Iriberri, N. (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic*

Literature, 51(1):5–62.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, 47(2):448–74.

Danz, D., Vesterlund, L., and Wilson, A. J. (2020). Belief elicitation: Limiting truth telling with information on incentives. Technical report, National Bureau of Economic Research.

Dianat, A., Echenique, F., and Yariv, L. (2018). Statistical discrimination and affirmative action in the lab.

Ding, T. and Schotter, A. (2017). Matching and chatting: An experimental study of the impact of network communication on school-matching mechanisms. *Games and Economic Behavior*, 103:94–115.

Ding, T. and Schotter, A. (2019). Learning and mechanism design: An experimental test of school matching mechanisms with intergenerational advice. *The Economic Journal*, 129(623):2779–2804.

Dubins, L. E. and Freedman, D. A. (1981). Machiavelli and the gale-shapley algorithm. *The American Mathematical Monthly*, 88(7):485–494.

Dustan, A., Koutout, K., and Leo, G. (2020). Second-order beliefs and gender. *Available at SSRN.*

Eckel, C., De Oliveira, A. C., and Grossman, P. J. (2008). Gender and negotiation in the small: are women (perceived to be) more cooperative than men? *Negotiation Journal*, 24(4):429–445.

Eckel, C. C. and Grossman, P. J. (2001). Chivalry and solidarity in ultimatum games. *Economic inquiry*, 39(2):171–188.

Eckel, C. C. and 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.

Fafchamps, M., Söderbom, M., and Benhassine, N. (2009). Wage gaps and job sorting in african manufacturing. *Journal of African Economies*, 18(5):824–868.

Fang, H. and Moro, A. (2010). *Theories of Statistical Discrimination and Affirmative Action: A Survey.* Handbook of Social Economics.

Featherstone, C. R. and Niederle, M. (2016). Boston versus deferred acceptance in an interim setting: An experimental investigation. *Games and Economic Behavior*, 100:353–375.

Fernandez, R. M. and Campero, S. (2017). Gender sorting and the glass ceiling in high-tech firms. *Ilr Review*, 70(1):73–104.

Fernandez, R. M. and Friedrich, C. (2011). Gender sorting at the application interface. *Industrial Relations: A Journal of Economy and Society*, 50(4):591–609.

Fernandez, R. M. and Mors, M. L. (2008). Competing for jobs: Labor queues and gender sorting in the hiring process. *Social Science Research*, 37(4):1061–1080.

Fershtman, C. and Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics*, 116(1):351–377.

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

Fortin, N. M. and Huberman, M. (2002). Occupational gender segregation and women's wages in canada: An historical perspective. *Canadian Public Policy/Analyse de Politiques*, pages S11–S39.

Friedenberg, A. (2019). Bargaining under strategic uncertainty: The role of second-order optimism. *Econometrica*, 87(6):1835–1865.

Gallen, Y., Lesner, R. V., and Vejlin, R. (2019). The labor market gender gap in denmark: Sorting out the past 30 years. *Labour Economics*, 56:58–67.

Glover, D., Pallais, A., and Pariente, W. (2017). Discrimination as a self-fulfilling prophecy: Evidence from french grocery stores. *The Quarterly Journal of Economics*, 132(3):1219–1260.

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

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125.

Groshen, E. L. (1991). The structure of the female/male wage differential: Is it who you are, what you do, or where you work? *Journal of Human Resources*, pages 457–472.

Guillen, P. and Hakimov, R. (2018). The effectiveness of top-down advice in strategy-proof mechanisms: A field experiment. *European Economic Review*, 101:505–511.

Guillen, P. and Hing, A. (2014). Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism. *European Economic Review*, 70:178–185.

Hakimov, R. and Kübler, D. (2019). Experiments on matching markets: A survey. WZB Discussion Paper SP II 2019-205, Berlin.

Harsanyi, J. C. (1967). Games with incomplete information played by "bayesian" players, i–iii part i. the basic model. *Management science*, 14(3):159–182.

Hassidim, A., Romm, A., and Shorrer, R. I. (2016). " strategic" behavior in a strategy-proof environment. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 763–764.

Hossain, T. and Okui, R. (2013). The binarized scoring rule. *Review of Economic Studies*, 80(3):984–1001.

Hyde, J. S., Fennema, E., and Lamon, S. J. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychological bulletin*, 107(2):139.

Kahneman, D. and Tversky, A. (2013a). Choices, values, and frames. In *Handbook of the fundamentals of financial decision making: Part I*, pages 269–278. World Scientific.

Kahneman, D. and Tversky, A. (2013b). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World Scientific.

Karni, E. (2009). A mechanism for eliciting probabilities. *Econometrica*, 77(2):603–606.

Klijn, F., Pais, J., and Vorsatz, M. (2013). Preference intensities and risk aversion in school choice: A laboratory experiment. *Experimental Economics*, 16(1):1–22.

Kneeland, T. (2015). Identifying higher-order rationality. *Econometrica*, 83(5):2065–2079.

Leo, G. (2020). Belief elicitation through rank-ordering.

Lundberg, S. J. (2017). Father absence and the educational gender gap.

Machina, M. J. and Schmeidler, D. (1992). A more robust definition of subjective probability. *Econometrica: Journal of the Econometric Society*, pages 745–780.

Manski, C. F. and Neri, C. (2013). First-and second-order subjective expectations in strategic decision-making: Experimental evidence. *Games and Economic Behavior*, 81:232–254.

Mantonakis, A., Rodero, P., Lesschaeve, I., and Hastie, R. (2009). Order in choice: Effects of serial position on preferences. *Psychological Science*, 20(11):1309–1312.

Marianne, B. (2011). New perspectives on gender. In *Handbook of labor economics*, volume 4, pages 1543–1590. Elsevier.

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102(5):2183–2205.

Miralles, A. (2009). School choice: The case for the boston mechanism. In *International conference on auctions, market mechanisms and their applications*, pages 58–60. Springer.

Mobius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2011). Managing self-confidence: Theory and experimental evidence. Technical report, National Bureau of Economic Research.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the national academy of sciences*, 109(41):16474–16479.

Murphy, J., Hofacker, C., and Mizerski, R. (2006). Primacy and recency effects on clicking behavior. *Journal of Computer-Mediated Communication*, 11(2):522–535.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, 122(3):1067–1101.

Perea, A. (2012). *Epistemic game theory: reasoning and choice*. Cambridge University Press.

Qu, X. (2012). A mechanism for eliciting a probability distribution. *Economics Letters*, 115(3):399–400.

Rees-Jones, A. (2018). Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match. *Games and Economic Behavior*, 108:317–330.

Reiley, D. and Mullin, C. H. (2000). Recombinant estimation for normal-form games, with applications to auctions and bargaining. Technical report, Econometric Society.

Reuben, E., Sapienza, P., and Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, 111(12):4403–4408.

Roth, A. E. and Malouf, M. W. (1979). Game-theoretic models and the role of information in bargaining. *Psychological review*, 86(6):574.

Salant, Y. and Rubinstein, A. (2008). (a, f): choice with frames. *The Review of Economic Studies*, 75(4):1287–1296.

Schlag, K. H. et al. (2013). Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality.

Schniter, E. and Shields, T. W. (2014). Ageism, honesty, and trust. *Journal of Behavioral and Experimental Economics*, 51:19–29.

Smith, C. A. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society: Series B (Methodological)*, 23(1):1–25.

Solnick, S. J. (2001). Gender differences in the ultimatum game. *Economic Inquiry*, 39(2):189–200.

Trautmann, S. T. and van de Kuilen, G. (2015). Belief elicitation: A horse race among truth serums. *The Economic Journal*, 125(589):2116–2135.

Troyan, P. (2012). Comparing school choice mechanisms by interim and ex-ante welfare. *Games and Economic Behavior*, 75(2):936–947.

Van Erkel, P. F. and Thijssen, P. (2016). The first one wins: Distilling the primacy effect. *Electoral Studies*, 44:245–254.

Wilson, A. and Vespa, E. (2018). Paired-uniform scoring: Implementing a binarized scoring rule with non-mathematical language. Technical report, Working paper.

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

Zhu, M. (2015). Experience transmission: Truth-telling adoption in matching. *Available at SSRN 2631442*.

A Task 1 Instructions

Participation ID ________

Task 1

In this task, you will be paired with a random partner. Your earnings will depend on the choice you make and the choice your partner makes. One of you will be assigned to be "Person 1" and the other to be "Person 2". The partner assigned to be Person 1 will propose how to split a total of $10 between the two partners. In other words, Person 1 proposes how much of the $10 to give to Person 2 and how much to keep for him or herself.

Person 2 then decides whether to accept or reject the split proposed by Person 1. If Person 2 accepts the proposal, the money is divided between Person 1 and Person 2 as proposed. If Person 2 rejects the proposal, both partners earn $0.

You must decide on the actions you will take in this game before knowing whether you will be Person 1 or Person 2. At the end of the experiment, we will pair you randomly with a partner and make choices on your behalf based on what you submit below. You will not know who your partner is and your partner will not know who you are. While your choices in this task will be used to determine your earnings, your choices will not be revealed during or after the experiment.

If you are **Person 1,** how much of the $10 would you like to propose to give to Person 2 (circle one)?

I propose to give Person 2:

$0 $1 $2 $3 $4 $5 $6 $7 $8 $9 $10

If you are **Person 2,** what is the smallest amount that Person 1 could propose to give you that you would accept (circle one)? If you are in the role of Person 2 and Person 1 offers you any amount equal to or larger than the number you circle below, you will automatically accept the split. If Person 1 offers you any amount less than the number you circle below, you will automatically reject the split and you will both earn $0.

The smallest amount that I would accept from Person 1 is:

$0 $1 $2 $3 $4 $5 $6 $7 $8 $9 $10

B Task 2 Instructions

Participation ID ________

Task 2

During this task you earn money by correctly summing 2-digit numbers. You will be shown several sets of five two-digit numbers. Each set will be arranged in a row. For example, you could see:

60	71	41	75	81	

For each set, you will write your answer in the empty box on the right. In the above example, the correct answer is 60 + 71 + 41 + 75 + 81 = 328. You would write 328 in the empty box.

For each correct answer, you will earn $0.50. You will not be penalized for incorrect answers. You have 5 minutes to solve as many of the summations as you can. You will be told when time is up, but no time warnings will be issued.

When the experimenter instructs you to do so, please turn to the next page and begin.

C Experiment Screenshots

Instructions

Today, you will be asked to make educated guesses about how people performed on two tasks in an experiment conducted recently here at the Vanderbilt University Experimental Economics Lab.

In the previous experiment, participants completed two tasks. All participants completed Task 1 first and could take as much time as they wanted to make two decisions. You will be asked about **one** of these decisions. Task 2 was a timed math exercise. Participants were paid for both tasks plus a $5 show-up fee.

We will now hand out a copy of the instructions used in this previous experiment.

Your Payment and Anonymity

In this experiment, your payment will be based on a lottery in which you receive either $15 or $0. The likelihood that you receive the larger amount of $15 is determined by how accurate your educated guess is compared to the actual outcome. (If you are interested, the lottery system is carefully designed so that it is mathematically optimal to submit your best guess about the median outcome.) So, **it is in your best interest to submit your true best guess.**

The decisions you make in this experiment are completely anonymous. Your identity will not be linked in any way to your decisions in this experiment. Your decisions are linked to your participant ID for payment purposes, but there is no record that matches your participant ID to your name.

An Example

In this example, you are asked to make an educated guess about which geographic location is closer to Vanderbilt University. You will not be paid for this example; it is only to ensure that you understand how to make your guess.

You must guess **which geographic location is closer** to our location at Vanderbilt University and **how much closer** it is.

Which is closer to our location at Vanderbilt University: the Titans Stadium or the Mall at Green Hills?

Titans Stadium	**I believe that they are the same distance**	Mall at Green Hills

Suppose you believe the Titans Stadium is 1.5 miles closer to our location at Vanderbilt University than the Mall at Green Hills. You would move the slider in to the section that says "Titans Stadium" until it says "1.5."

Titans Stadium	**I believe that the Titans Stadium is closer by 1.5 miles**	Mall at Green Hills

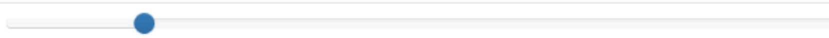
A chart shows your probability of winning $15 based on what the actual distance is.

If the actual distance is:	The Titans Stadium is closer by							The Mall at Green Hills is closer by					
	3+	**2.5**	**2**	**1.5**	**1**	**0.5**	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3+**
You win $15 with probability:	75%	83%	92%	100%	92%	83%	75%	67%	58%	50%	42%	33%	25%

For example, if your guess is accurate and the Titans Stadium is 1.5 miles closer than the Mall at Green Hills, you win $15 for sure (100%). On the other hand, if the Titans Stadium is actually 0.5 miles closer, you have a 83% chance of winning $15. If the Mall at Green Hills is closer than the Titans Stadium by 1.5 miles, your chance of winning $15 falls to 50%.

As you move the slider, the chart will update to show the probabilities of winning $15 at each possible value of the actual distance. So, if you decided the Titans Stadium was actually 2 miles closer than the Mall at Green Hills, the chart would change when you moved the slider.

Titans Stadium	**I believe that the Titans Stadium is closer by 2 miles**	Mall at Green Hills



If the actual distance is:	The Titans Stadium is closer by							The Mall at Green Hills is closer by					
	3+	**2.5**	**2**	**1.5**	**1**	**0.5**	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3+**
You win $15 with probability:	83%	92%	100%	92%	83%	75%	67%	58%	50%	42%	33%	25%	17%

You will now have an opportunity to test the slider and make your guess. Remember, this example is just for practice and you will not be paid for the results.

An Example

In this example, you are asked to make an educated guess about which geographic location is closer to Vanderbilt University. You will not be paid for this example; it is only to ensure that you understand how to make your guess.

Which is closer to our location at Vanderbilt University: the Titans Stadium or the Mall at Green Hills?

Titans Stadium	**I believe that they are the same distance**	Mall at Green Hills

If the actual distance is:	The Titans Stadium is closer by							The Mall at Green Hills is closer by					
	3+	**2.5**	**2**	**1.5**	**1**	**0.5**	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3+**
You win $15 with probability:	50%	58%	67%	75%	83%	92%	100%	92%	83%	75%	67%	58%	50%

Next

Example Results

The Titans Stadium is actually 0.5 miles closer to our location at Vanderbilt University than the Mall at Green Hills. You would have won $15 with 92% probability.

Titans Stadium	**You guessed that they are the same distance**	Mall at Green Hills

If the actual distance were:	The Titans Stadium is closer by							The Mall at Green Hills is closer by					
	3+	**2.5**	**2**	**1.5**	**1**	**0.5**	**0**	**0.5**	**1**	**1.5**	**2**	**2.5**	**3+**
You win $15 with probability:	50%	58%	67%	75%	83%	92%	100%	92%	83%	75%	67%	58%	50%

Now you will make educated guesses that determine your payment in this experiment. Consider your choices carefully. One of the guesses you make will be randomly chosen by a computer to determine your payment. Each guess is equally likely to be selected but you will not know which guess is chosen for payment until the end of the experiment. It is in your best interest to treat each guess as if it is the one that determines your payment.

Task 1

A computer will randomly draw one man and one woman from the previous experiment. Consider the decision each of these individuals made in the role of Person 2 in Task 1. You must guess **which individual chose the larger amount in the role of Person 2** and **how much larger** that amount was. In other words, who chose a larger amount in response to "The smallest amount that I would accept from Person 1 is:" and how much larger was that amount?

Woman	**I believe that they chose the same amount**	Man

If the actual outcome is:	The woman chose a larger amount by										The man chose a larger amount by									
	$10	**$9**	**$8**	**$7**	**$6**	**$5**	**$4**	**$3**	**$2**	**$1**	**$0**	**$1**	**$2**	**$3**	**$4**	**$5**	**$6**	**$7**	**$8**	**$9**	**$10**
You win $15 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Task 2

A computer will randomly draw one man and one woman from the previous experiment. You must guess **which individual answered more of the math sums correctly** and **how many more.**

Woman	**I believe that they answered the same amount correctly**	Man

If the actual outcome is:	The woman answered more sums correctly by										The man answered more sums correctly by									
	10+	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10+**
You win $15 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

More Instructions

In an earlier session of this experiment, participants made the same two choices you just made. They were given the same instructions and asked to make their best guess. You must now make educated guesses about what those participants chose as their guesses. Consider your choices carefully. Again, any one of your guesses could be randomly chosen to determine your payment in this experiment and each is equally likely.

Task 2, Man

A computer will randomly draw one man from a previous session of this experiment. You must guess what he reported as his guess when asked if the man or the woman **answered more of the math sums correctly** and **how many more.**

Woman	**I believe that the man guessed they answered the same amount correctly**	Man

If the man reported:	The woman answered more sums correctly by											The man answered more sums correctly by									
	10+	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10+**
You win $15 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Task 2, Woman

A computer will randomly draw one woman from a previous session of this experiment. You must guess what she reported as her guess when asked if the man or the woman **answered more of the math sums correctly** and **how many more.**

Woman	**I believe that the woman guessed they answered the same amount correctly**	Man

If the woman reported:	The woman answered more sums correctly by											The man answered more sums correctly by									
	10+	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10+**
You win $15 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Task 1, Man

A computer will randomly draw one man from a previous session of this experiment. You must guess what he reported as his guess when asked if the man or the woman **chose the larger number in the role of Person 2 in Task 1** and **how much larger.**

Woman	**I believe that the man guessed they chose the same amount**	Man

If the man reported:	The woman chose a larger amount by											The man chose a larger amount by									
	$10	**$9**	**$8**	**$7**	**$6**	**$5**	**$4**	**$3**	**$2**	**$1**	**$0**	**$1**	**$2**	**$3**	**$4**	**$5**	**$6**	**$7**	**$8**	**$9**	**$10**
You win $15 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

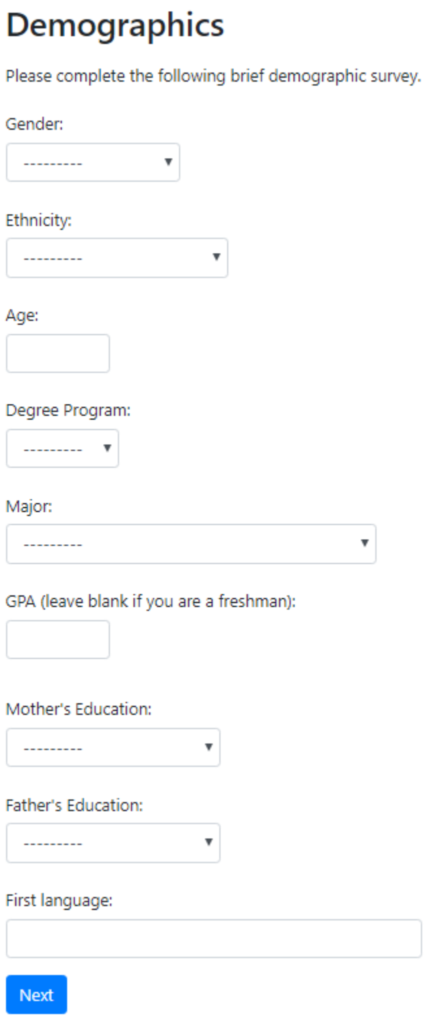
Task 1, Woman

A computer will randomly draw one woman from a previous session of this experiment. You must guess what she reported as her guess when asked if the man or the woman **chose the larger number in the role of Person 2 in Task 1** and **how much larger.**

Woman	**I believe that the woman guessed they chose the same amount**	Man

If the woman reported:	The woman chose a larger amount by											The man chose a larger amount by									
	$10	**$9**	**$8**	**$7**	**$6**	**$5**	**$4**	**$3**	**$2**	**$1**	**$0**	**$1**	**$2**	**$3**	**$4**	**$5**	**$6**	**$7**	**$8**	**$9**	**$10**
You win $15 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next



Demographics

Please complete the following brief demographic survey.

Gender:

Ethnicity:

Age:

Degree Program:

Major:

GPA (leave blank if you are a freshman):

Mother's Education:

Father's Education:

First language:

Next

Results

Your payment will be based on the following choice:

A computer will randomly draw one man and one woman from the previous experiment. You must guess **which individual answered more of the math sums correctly** and **how many more.**

Woman	**You guessed that the man answered more sums correctly by 5**	Man

If the actual outcome were:	The woman answered more sums correctly by											The man answered more sums correctly by									
	10+	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10+**
You win $15 with probability:	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%

You guessed that the man answered more sums correctly by 5 and the actual outcome was the woman answered more sums correctly by 5, so you have a 50% chance of winning $15.

When you press this button, a random number between 0 and 100 will be chosen. If that number is less than 50 (your percent chance of winning), you win $15. Otherwise, you win $0.

Generate random number **20**

Next

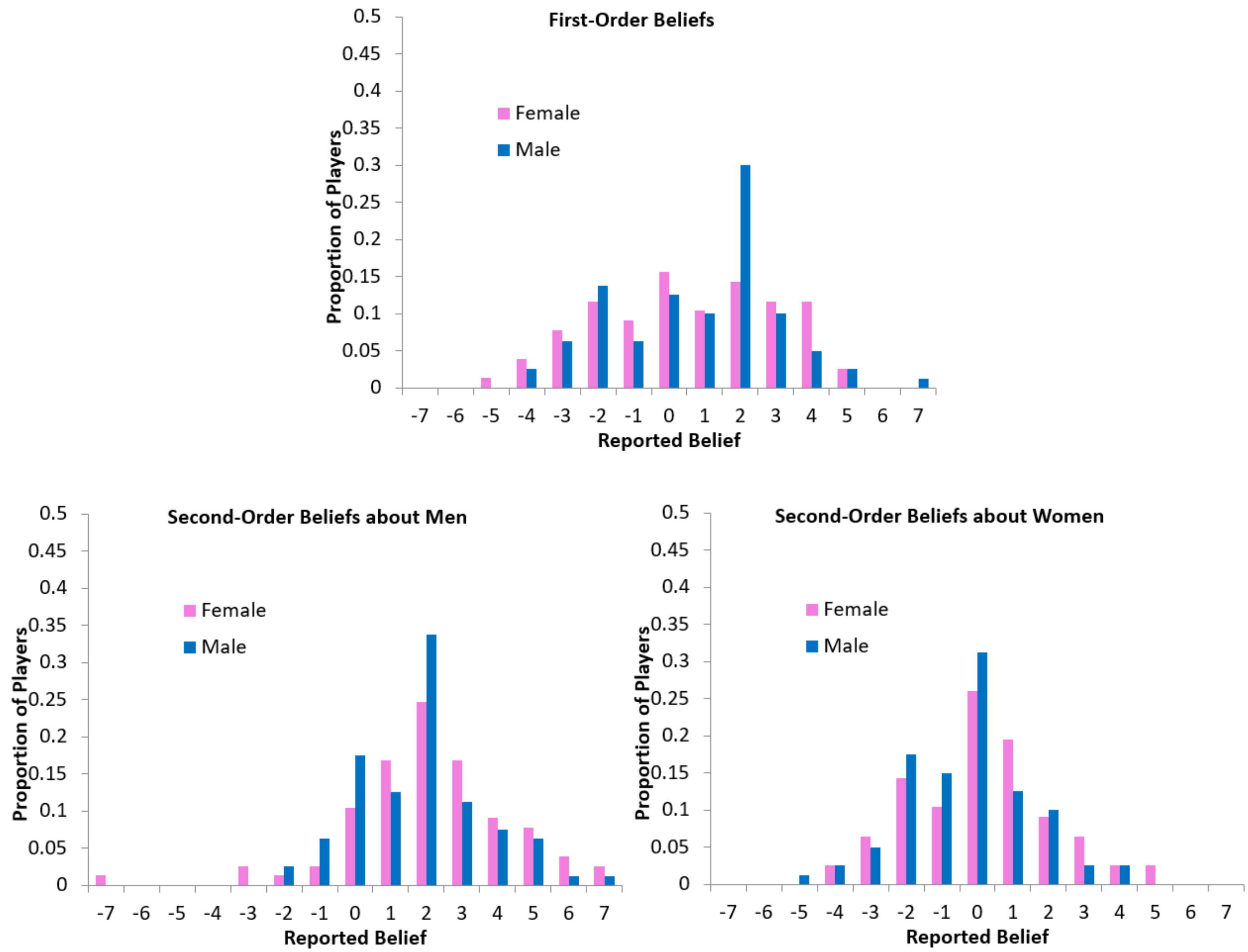
D Additional Figures and Tables

Table 3.9: Proportion of participants reporting same-gender second-order beliefs matching their own first-order beliefs, by first-order belief

	All	Men	Women	Difference
Math Task				
W>M	0.408	0.261	0.538	-0.278
	(0.071)	(0.094)	(0.100)	(0.134)
	[49]	[23]	[26]	
W=M	0.409	0.400	0.417	-0.017
	(0.107)	(0.163)	(0.149)	(0.210)
	[22]	[10]	[12]	
W<M	0.698	0.830	0.538	0.291
	(0.050)	(0.055)	(0.081)	(0.097)
	[86]	[47]	[39]	
Ultimatum Task				
W>M	0.357	0.556	0.000	0.556
	(0.133)	(0.176)	(0.000)	(0.166)
	[14]	[9]	[5]	
W=M	0.500	0.500	0.500	0.000
	(0.090)	(0.121)	(0.139)	(0.178)
	[32]	[18]	[14]	
W<M	0.775	0.755	0.793	-0.038
	(0.040)	(0.060)	(0.054)	(0.080)
	[111]	[53]	[58]	

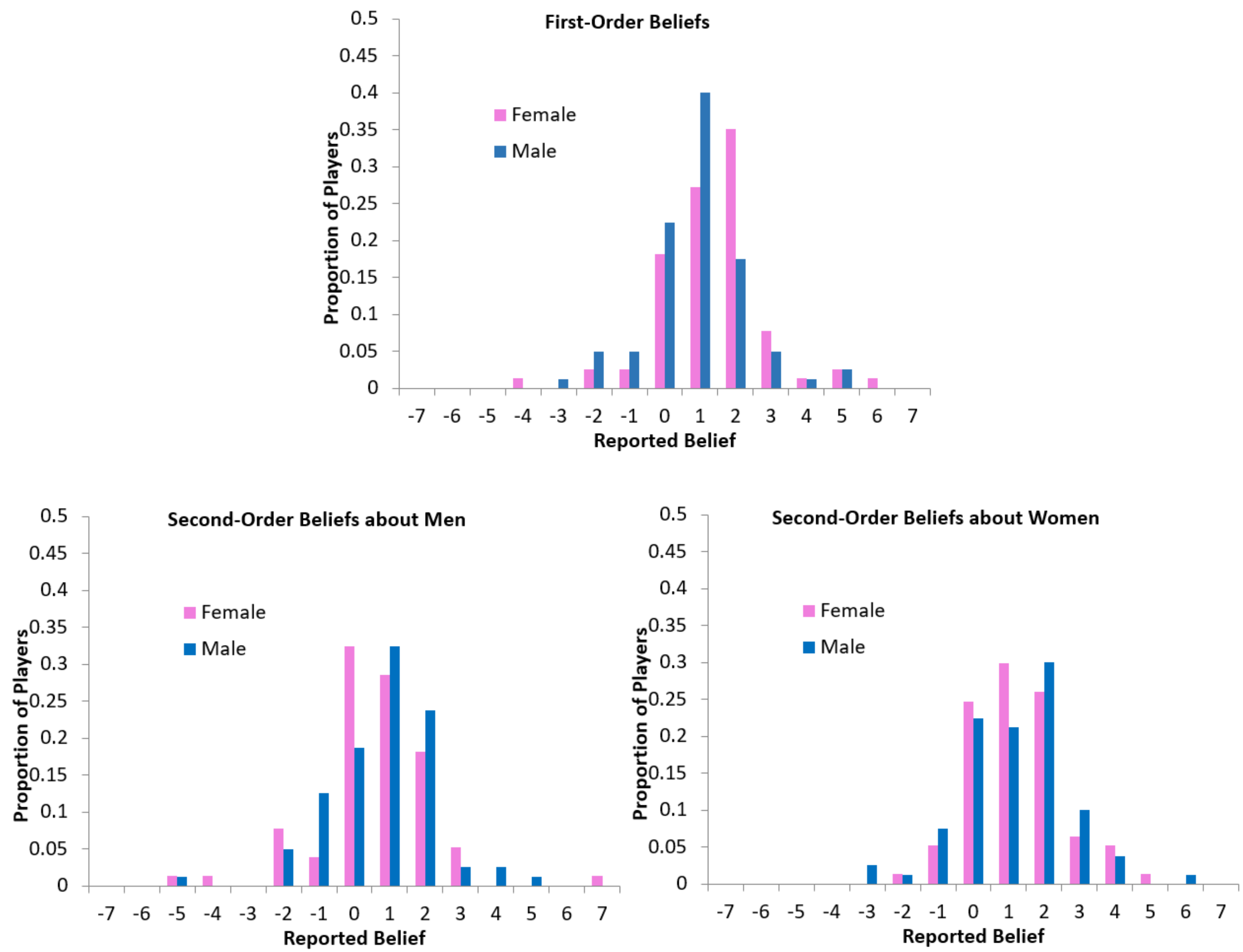
Notes: Columns (1) to (3) reference subsamples. Column (4) reports the differences between the men and women subsamples. Standard errors are reported in parentheses underneath the proportions. The number of participants in each cell are reported in brackets underneath the standard errors.

Figure 3.9: Belief Elicitation about the Math Task



Notes: The bars represent the proportion of players who report each median difference between a man and a woman on the math task. A negative difference means that the woman answered more math summations correctly, while a positive difference means that the man answered more correctly.

Figure 3.10: Belief Elicitations about the Bargaining Task



Notes: The bars represent the proportion of players who report each median difference between a man and a woman on the bargaining task. A negative difference means that the woman chose a higher MAO, while a positive difference means that the man chose a higher MAO.

E Job Advertisement

We are recruiting top talent for multiple outside business to business sales positions in and around (City). The positions involve:

- interfacing with new and existing customers
- developing leads
- cold calling prospects
- working on quotations
- assisting with order fulfillment
- entertaining clients
- partnering with vendor representatives

Starting package includes salary commensurate with experience and comprehensive benefits, with the potential to move to an incentive-based system with unlimited six-figure earning potential. Positions require a minimum of a bachelor's degree, with some sales experience preferred.

Press 'Apply' to express your interest in receiving company information and detailed job descriptions. Thank you for your interest in these competitive positions!

F Treatment E-mail

Dear «First Name» «Last Name»,

Thank you for your interest in hearing more about our open positions!

We are currently hiring new outside business to business sales representatives in «City» for «Firm Name». «Firm Name» is a wholesale distributor of disposable goods including janitorial and office products with a strong regional presence.[24] The position we are hiring for involves interfacing with new and existing customers, developing leads, cold calling prospects, working on quotations, assisting with order fulfillment, entertaining clients, and partnering with vendor representatives.

«Firm Name» has a solid infrastructure in both their «City & Corporate Offices» to support your sales efforts, so that maximum time can be devoted to interfacing with customers. Most meetings take place in-person, but both hours and place of work are flexible. Compensation for the position consists of a competitive base salary plus expenses, potentially moving to an incentive-based system with unlimited six-figure earning potential. The comprehensive benefits package includes:

- Health insurance
- Generously funded Health Savings Account
- 401(k) plan with matching contributions offered
- Employer paid life insurance provided
- Dental, vision, long term disability and AFLAC plans offered
- Employee Assistance Program offered and paid 100%
- Paid Time Off

[24] The order of the product sectors here and later in the treatment e-mail was randomized.

Minimum Requirements:

- 4 year degree from an accredited university or college.
- Evidence of achievement and progression of results in school, career, or outside interests.
- Basic understanding of business transactions- bids and proposals, requisitions, purchase orders, logistics, payments, etc.
- Firm understanding of gross profit margins.
- Excellent written and verbal communication skills- able to recognize and recall nuance from conversations as well as present written data/information.
- Able to prospect, persuade, and close business accounts- ideal candidate will be able to get commitments from potential customers even when the initial answer is no.
- Excellent time management and prioritization.
- Previous experience in outside B2B sales is preferred.

Janitorial Products
You should be comfortable working with janitorial product buyers such as facility managers, operations, and maintenance workers. The most successful team members assess problems and offer solutions that improve the productivity and cost efficiency of customers' facilities, while being profitable to the business.

To apply for this position, please send your resume and cover letter (optional) to the hiring manager **«Manager Name & E-mail»**, then complete the application at the link below.

http://www.«Recruiting Firm Name».com/job-application-«Firm Abbreviation»

Office Products

You should be comfortable working with office product buyers such as office managers, receptionists, and administrative workers. The most successful team members assess problems and offer solutions that improve the productivity and cost efficiency of customers' offices, while being profitable to the business.

To apply for this position, please send your resume and cover letter (optional) to the hiring manager **«Manager Name & E-mail»**, then complete the application at the link below.

http://www.«Recruiting Firm Name».com/job-application-«Firm Abbreviation»

Please **submit your application to only one manager.** Application materials sent to more than one manager will not be processed.[25]

We look forward to reviewing your application. Thank you again for your interest!

«Recruiting Firm Name»

[25] In the first wave, these two sentences were instead "To avoid duplication, **please submit your application to only one manager."** . The change was made in the second round to reduce administrative burden from contacting workers who submitted application materials to more than one manager and increase worker compliance with protocol.

G Recruitment E-mail

Dear (First Name) (Last Name),

Vanderbilt University is conducting a paid economics study, in which you will be asked to make educated guesses using your career experience in sales. You are guaranteed a $5 payment just for completing the study, but you can earn $15 more based on how accurate your guesses are. In addition, you will have the opportunity to earn tickets into a drawing for a $500 Amazon gift card!

The survey takes just 10-15 minutes to complete, including listening to instructions about making your educated guesses. Your participation is crucial to obtaining a representative sample, as only a limited number of individuals have been invited.

You will be paid via PayPal within 24 hours and can deposit the money directly into your bank account or spend it online. You do not need a Paypal account to receive payment. Thank you in advance for assisting Vanderbilt University in this important research!

Note that your participation in the survey is completely voluntary and your responses are kept anonymous. Your responses to the survey are used for research purposes only- there will be no commercial use of the survey and your data will never be shared. This study has been approved by the Vanderbilt University Institutional Review Board. If you have any questions, contact Kristine Koutout at kristine.f.koutout@vanderbilt.edu.

To participate, click on your personal link below. Please be sure to take the survey on a large screen like a laptop or tablet rather than a phone. If you need to leave the survey at any time, you can return to your place using the link below.

(Link)

Thank you again for your participation! If you want to opt out of receiving reminders about this survey, please reply to this e-mail.

Kristine Koutout
Vanderbilt University

H Experiment Screenshots

Thank you for participating!

Please click on the audio file to listen to the instructions. You can follow along in the text below.

Press Play.

0:00 / 2:44

Note that the "Next" button on pages with instructions (like this one) will only become available after the audio file has played through. If you prefer to read the instructions, please press the button to mute the audio while it plays or silence your device.

Please also note that the link you used to reach this survey is unique to you. If you need to leave the survey, you can always return to exactly where you left off by clicking on the link again.

The "Progress Bar" at the bottom of each page tracks how far along you are in the survey. Once you reach the end, you will learn how much money you earned total. You are guaranteed to earn at least $5.00. See **Payment** information and **Instructions** below for more details on how your total earnings will be determined.

Anonymity

Your responses on this survey are completely anonymous and will be used for research purposes only. Information like your name and e-mail address is not stored with your responses to the survey.

This survey has been reviewed by the Institutional Review Board (IRB) at Vanderbilt University to ensure your privacy is protected. You may direct any questions or concerns about this survey to the Vanderbilt IRB at (866) 224-8273. For more information, click here.

Payment

For simply completing this 10 to 15 minute survey, you will be paid $5.00. You may earn an additional $15.00 prize based on your educated guesses in the next section of this survey. Your payment will be sent via PayPal to the e-mail address of your choice within 24 hours.

In addition, you will have the opportunity to be entered into a drawing for a $500.00 Amazon gift card. The drawing for the $500.00 Amazon gift card will occur at the conclusion of this online survey. You will receive more information about winning the $500.00 Amazon gift card after you make your guesses.

Instructions

In this survey, you will be asked to make five educated guesses about how people performed on an assessment of sales ability. One guess out of those five will be randomly selected by a computer at the end of the survey to determine your total payment.

Your selected guess will be compared to the *actual performance* of randomly selected people who took the assessment of sales ability as a part of their job application for a sales position. The closer your educated guess is compared to the actual outcome of those randomly selected people, the more likely you are to earn the $15.00 prize. A lottery system determines how likely you are to earn the $15.00 prize based on the accuracy of your guess. (If you are interested, the lottery system is carefully designed so that it is mathematically optimal to submit your best guess about the median outcome.) **You are most likely to earn the $15.00 prize if you submit your true best guess.**

To demonstrate how the lottery system works, I will go through an example with you.

Next

Progress Bar

Example

Press Play.

1:30 / 1:30

In this example, you are asked to make an educated guess about which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona. You will not be paid for this example; it is only to ensure that you understand how to make your guess.

You will see a screen like the following:

Image 1

Which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona?

	Washington, D.C.					**I believe that they are the same distance**					Phoenix, Arizona
If the actual distance is:		Washington, D.C. is closer by						Phoenix, Arizona is closer by			
	1000	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000**
You win $15.00 with probability:	50%	60%	70%	80%	90%	100%	90%	80%	70%	60%	50%

Suppose you believe Washington, D.C. is 400 miles closer to St. Louis, Missouri than Phoenix, Arizona. You would move the slider in to the section that says "Washington, D.C." until it says "400." Note that the slider now says "I believe that Washington, D.C. is closer by 400 miles."

Image 2

Which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona?

	Washington, D.C.					**I believe that Washington, D.C. is closer by 400 miles**					Phoenix, Arizona
If the actual distance is:		Washington, D.C. is closer by						Phoenix, Arizona is closer by			
	1000	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000**
You win $15.00 with probability:	70%	80%	90%	100%	90%	80%	70%	60%	50%	40%	30%

The chart below the slider shows your probability of winning the $15.00 prize based on what the actual distance is and your guess. For example, if your guess is accurate and Washington, D.C. is 400 miles closer than Phoenix, Arizona, you win the $15.00 prize for sure (100%).

Image 3

Which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona?

	Washington, D.C.					**I believe that Washington, D.C. is closer by 400 miles**					Phoenix, Arizona
If the actual distance is:		Washington, D.C. is closer by						Phoenix, Arizona is closer by			
	1000	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000**
You win $15.00 with probability:	70%	80%	90%	100%	90%	80%	70%	60%	50%	40%	30%

If Washington, D.C. is actually 200 miles closer, you have a 90% chance of winning the $15.00 prize.

Image 4

Which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona?

	Washington, D.C.					**I believe that Washington, D.C. is closer by 400 miles**					Phoenix, Arizona
If the actual distance is:		Washington, D.C. is closer by							Phoenix, Arizona is closer by		
	1000	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000**
You win $15.00 with probability:	70%	80%	90%	100%	90%	80%	70%	60%	50%	40%	30%

On the other hand, if Phoenix, Arizona is actually closer than Washington, D.C. by 200 miles, your chance of winning the $15.00 prize falls to 70%.

Image 5

Which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona?

	Washington, D.C.					**I believe that Washington, D.C. is closer by 400 miles**					Phoenix, Arizona
If the actual distance is:		Washington, D.C. is closer by							Phoenix, Arizona is closer by		
	1000	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000**
You win $15.00 with probability:	70%	80%	90%	100%	90%	80%	70%	60%	50%	40%	30%

As you move the slider, the chart will update to show the probabilities of winning the $15.00 prize at each possible value of the actual distance. So, if you decided Washington, D.C. was actually 800 miles closer than Phoenix, Arizona, the chart would change when you moved the slider.

Image 6

Which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona?

	Washington, D.C.					**I believe that Washington, D.C. is closer by 800 miles**					Phoenix, Arizona
If the actual distance is:		Washington, D.C. is closer by							Phoenix, Arizona is closer by		
	1000	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000**
You win $15.00 with probability:	90%	100%	90%	80%	70%	60%	50%	40%	30%	20%	10%

You will now have an opportunity to test the slider and make your guess. Remember, this example is just for practice and you will not be paid for the results.

Next

Progress Bar

Example Task

In this example, you are asked to make an educated guess about which city is closer to St. Louis, Missouri. You will not be paid for this example; it is only to ensure that you understand how to make your guess.

Which city is closer to St. Louis, Missouri: Washington, D.C. or Phoenix, Arizona?

	Washington, D.C.					**I believe that they are the same distance**					Phoenix, Arizona
If the actual distance is:	Washington, D.C. is closer by						Phoenix, Arizona is closer by				
	1000	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000**
You win $15.00 with probability:	50%	60%	70%	80%	90%	100%	90%	80%	70%	60%	50%

Next

Progress Bar

Example Results

The results show you the actual outcome and your probability of winning the $15.00 prize based on your guess.

Washington, D.C. is actually 600 miles closer to St. Louis, Missouri than Phoenix, Arizona is. You would have won $15.00 with 70% probability.

	Washington, D.C.					**You guessed that they are the same distance**					Phoenix, Arizona
If the actual distance were:	Washington, D.C. is closer by						Phoenix, Arizona is closer by				
	1000+	**800**	**600**	**400**	**200**	**0**	**200**	**400**	**600**	**800**	**1000+**
You win $15.00 with probability:	50%	60%	70%	80%	90%	100%	90%	80%	70%	60%	50%

Now, press the random number generator (RNG). The RNG selects a random number between 1 and 100. If that number is equal to or lower than your probability of winning, you would earn the $15.00 prize. In other words, if you have an 80% chance of winning, then you would win the $15.00 prize if the RNG selects a number that is 80 or lower.

RNG **93**

I'm sorry, you would not have won the $15.00 prize.

Next

Progress Bar

Task Instructions

Press Play.

0:00 / 1:15

You will now make five educated guesses that determine whether you earn the $15.00 prize. Remember that each guess is related to the performance of individuals on an assessment of sales ability. This sales assessment is used by employers to evaluate job candidates for positions in outside sales.

The assessment evaluates job candidates in three broad categories: "will to sell" (e.g. desire and commitment to success), "sales DNA" (e.g. controls emotions and handles rejection), and "selling competencies" (e.g. hunting and closing sales).

Individual performance on the assessment is measured as a percentile in the total population of job candidates who have taken this assessment (more than 2 million people). So, an individual who scored 90% on the sales assessment performed better than 90% of the people who completed this assessment. An individual who scored 25% performed better than 25% of people who completed the assessment.

Consider your choices carefully. Remember that one of the guesses you make will be randomly selected by a computer to determine your payment. Each guess is equally likely to be selected but you will not know which guess is selected for payment until the end of the survey. It is in your best interest to treat each guess as if it is the one that determines your payment.

Next

Progress Bar

Task 1

A computer at the end of this survey will randomly select a man and a woman who completed the sales ability assessment when applying for a sales position. **Who do you believe did better on the sales ability assessment, the randomly selected man or the randomly selected woman, and by how much?**

	Woman										**I believe that they did the same**										Man
If the actual outcome is:				The woman did better by												The man did better by					
	50+	**45**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50+**
You win $15.00 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Progress Bar

Additional Task Instructions

Press Play.

0:00 / 0:41

In an earlier version of this survey, managers responsible for hiring and supervising outside sales representatives answered the same question you just did. They were given the same instructions you received at the beginning of this survey, including the information about the anonymity of their responses, and asked to make their true best guess. These managers were also paid based on how close their educated guess was to the actual outcome when a computer selected a random man and a random woman. **You must now make educated guesses about what those managers chose as their guesses.** You will make your last four guesses, each corresponding to a different type of manager. Consider your choices carefully. Again, any one of your guesses could be randomly selected to determine your payment at the end of this survey and each is equally likely.

Next

Progress Bar

Task 2, Female Manager in a Traditionally Female Product Sector

A computer at the end of this survey will randomly select a female manager responsible for hiring and managing sales representatives in a traditionally female product sector. **What did she choose when asked "Who do you believe did better on the sales ability assessment, the randomly selected man or the randomly selected woman, and by how much?"**

	Woman										**I believe that the manager guessed they did the same**										Man
If the manager guessed:				The woman did better by												The man did better by					
	50+	**45**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50+**
You win $15.00 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Progress Bar

Task 3, Male Manager in a Traditionally Female Product Sector

A computer at the end of this survey will randomly select a male manager responsible for hiring and managing sales representatives in a traditionally female product sector. **What did he choose when asked "Who do you believe did better on the sales ability assessment, the randomly selected man or the randomly selected woman, and by how much?"**

	Woman										**I believe that the manager guessed they did the same**										Man
If the manager guessed:				The woman did better by												The man did better by					
	50+	**45**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50+**
You win $15.00 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Progress Bar

Task 4, Female Manager in a Traditionally Male Product Sector

A computer at the end of this survey will randomly select a female manager responsible for hiring and managing sales representatives in a traditionally male product sector. **What did she choose when asked "Who do you believe did better on the sales ability assessment, the randomly selected man or the randomly selected woman, and by how much?"**

	Woman										I believe that the manager guessed they did the same										Man
If the manager guessed:	The woman did better by																				The man did better by
	50+	**45**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50+**
You win $15.00 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Progress Bar

Task 5, Male Manager in a Traditionally Male Product Sector

A computer at the end of this survey will randomly select a male manager responsible for hiring and managing sales representatives in a traditionally male product sector. **What did he choose when asked "Who do you believe did better on the sales ability assessment, the randomly selected man or the randomly selected woman, and by how much?"**

	Woman										I believe that the manager guessed they did the same										Man
If the manager guessed:	The woman did better by																				The man did better by
	50+	**45**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50+**
You win $15.00 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Next

Progress Bar

Task 6

Recall the last educated guess you made and answer the questions below.

In the last question, did we ask you about a male or female manager?

In the last question, did we ask you about a traditionally male or traditionally female product sector?

Next

Progress Bar

Task 7

Press Play.

0:00 / 0:42

You will now have the opportunity to earn tickets that will be entered into a drawing for a $500.00 Amazon gift card. At the conclusion of this study, one ticket will be drawn to determine the winner, who will receive the gift card via e-mail. **The more tickets you earn, the greater chance you have of winning the $500.00 Amazon gift card.**

In an earlier version of this survey, managers responsible for hiring and supervising outside sales representatives were asked the following question:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

You must now guess what each type of manager chose. For each correct guess you make below, you earn 5 tickets for the drawing to win the $500.00 Amazon gift card.

Consider a **male manager in a traditionally male sector.** What do you believe he chose when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

Consider a **female manager in a traditionally male sector.** What do you believe she chose when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

Consider a **male manager in a traditionally female sector.** What do you believe he chose when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

Consider a **female manager in a traditionally female sector.** What do you believe she chose when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

Next

Progress Bar

Task 8

Press Play.

0:00 / 0:34

You will now have another opportunity to earn tickets that will be entered into the drawing for a $500.00 Amazon gift card.

You must guess whether the product sectors listed below are traditionally male or traditionally female. If your guess matches the most commonly guessed answer at the end of this study, you earn 5 tickets into the drawing for the $500.00 Amazon gift card. Each guess is worth 5 tickets, so you can earn up to a total of 45 tickets into the drawing.

For example, if most people who take this survey guess that "pharmaceuticals" is a traditionally male product sector, then you earn 5 tickets if you also guessed that the product sector is traditionally male.

Banking Services:

Beauty Products:

Financial Services:

Janitorial Products:

Manufacturing Machinery:

Office Products:

Packaging Products:

Pharmaceuticals:

Travel Services:

Next

Progress Bar

Task 9

Press Play.

0:00 / 0:40

You will now have **one more opportunity to win cash**, paid via PayPal within 24 hours.

Choose one of the seven gambles listed below. Each gamble has two possible outcomes: Low and High. The two outcomes are equally likely. At the end of this survey, a random number generator (RNG) will select a number between 1 and 100. If the random number is 50 or lower, you win the "Low" dollar amount. If the number is 51 or higher, you win the "High" dollar amount.

For example, suppose you choose gamble 4. Then, if the random number generated is 50 or lower, you win $0.45. If the random number generated is 51 or higher, you win $2.15.

Gamble	Low	High
○ 1	$1.00	$1.00
○ 2	$0.75	$1.45
○ 3	$0.60	$1.80
○ 4	$0.45	$2.15
○ 5	$0.30	$2.50
○ 6	$0.15	$2.65
○ 7	$0	$2.80

Next

Progress Bar

Task 10

In general, do you prefer to work with mostly men, mostly women, an equal number of each, or have no preferences either way?

In general, do you prefer a male manager, a female manager, or have no preferences either way?

In general, do you prefer to work in a traditionally male product sector, a traditionally female product sector, or have no preferences either way?

We welcome any comments you have about your preferences. Remember that all responses in this survey are anonymous.

Next

Progress Bar

Task 11

Consider a hypothetical situation in which you and 9 other sales representatives are hired at the same time to promote a new product line. At the end of one year, how do you think you will rank in total sales relative to the other sales representatives hired at the same time as you?

Rank 1 for highest through 10 for lowest.

Hypothetical rank:

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Next

Progress Bar

Survey Questions

How accurate do you think the educated guesses you made in this survey are?

How often do you take risks in general?

How often do you take risks in your career?

Are you currently employed full-time, employed part-time, or unemployed?

If currently employed, are you paid by salary, hourly, or commission?

What was your total earned income last year?

Are you actively seeking employment or looking to change jobs?

If you are not currently employed, answer the questions based on your last job.

Is your direct supervisor male, female, or other?

Do you work with mostly women, mostly men, or an equal number of each?

What is your age?

What is your gender?

What is your race/ethnicity?

What is your marital status?

How many children do you have?

How many children under 5 do you have?

See Results

Progress Bar

Results

Your payment will be based on the following choice:

You guessed that they did the same and the actual outcome was the man did better by 15% , so you have a 85% chance of winning $15.00.

	Woman										You guessed that they did the same										Man
If the actual outcome were:				The woman did better by												The man did better by					
	50+	**45**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**0**	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50+**
You win $15.00 with probability:	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%

Now, press the random number generator (RNG). The RNG selects a random number between 1 and 100. If that number is equal to or lower than your probability of winning, you would earn the additional $15.00. In other words, you have a 85% chance of winning, so you win the $15.00 if the RNG selects a number that is 85 or lower.

RNG **89**

Task 9

You chose gamble 5 from the table below. Now, press the random number generator (RNG). The RNG selects a random number between 1 and 100. If that number is equal to or lower than 50, you win $0.30. If the RNG selects a number greater than 50, you win $2.50.

Gamble	Low	High
1	$1.00	$1.00
2	$0.75	$1.45
3	$0.60	$1.80
4	$0.45	$2.15
5	$0.30	$2.50
6	$0.15	$2.65
7	$0	$2.80

RNG **5**

Based on the gamble you chose, **you earn $0.30**.

Results

Task 7

You said you thought a male manager in a traditionally male sector chose "Male" when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

In fact, the manager chose "Male", so based on your answer **you earn 5 tickets** for the $500.00 Amazon gift card drawing.

You said you thought a female manager in a traditionally male sector chose "Male" when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

In fact, the manager chose "No Preference", so based on your answer **you earn 0 tickets** for the $500.00 Amazon gift card drawing.

You said you thought a male manager in a traditionally female sector chose "Female" when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

In fact, the manager chose "Male", so based on your answer **you earn 0 tickets** for the $500.00 Amazon gift card drawing.

You said you thought a female manager in a traditionally female sector chose "Male" when asked:

"In general, do you think customers in your product sector prefer working with a male sales representative, a female sales representative, or have no preferences either way?"

In fact, the manager chose "No Preference", so based on your answer **you earn 0 tickets** for the $500.00 Amazon gift card drawing.

You have earned a total of 5 tickets for the $500.00 Amazon gift card drawing, in addition to those you may earn based on your answers about whether certain product sectors are traditionally male or female. If your ticket is drawn at the conclusion of this study, you will be contacted via e-mail to receive the gift card.

Thank you for completing the survey! You have earned a total of $5.30. **Please fill out the below receipt to receive payment.** This information is collected separately to ensure that your responses are anonymous. Your payment will be sent within 24 hours of receiving this receipt.

Receipt

I Theoretical Framework

I outline a simple theoretical framework to show how workers' second-order beliefs could affect the job application decision. The model incorporates design elements from the experiment to better represent the context of the worker's decision. Workers may be male or female, but the model is symmetric by gender, so I analyze the case of one gender.

I study the behavior of risk-neutral workers in a market in which there are four jobs. The jobs are characterized by manager type, as listed in Table 3.10. Manager type is determined by the manager's own gender, indexed by $g \in \{m, f\}$, and the product sector, indexed by $s \in \{j, o\}$ where j and o correspond to the male (janitorial) and female (office) product sectors from the field experiment respectively.

Table 3.10: Manager Types

Type	Manager Gender	Product Sector
$\{m, j\}$	Male	Janitorial (M)
$\{f, j\}$	Female	Janitorial (M)
$\{m, o\}$	Male	Office (F)
$\{f, o\}$	Female	Office (F)

There are two time periods $t \in \{1, 2\}$. In the first time period, there are two job advertisement slots $a \in \{1, 2\}$. For every worker $i \in I$, Nature randomly draws one manager type from the office sector $\{g, o\} = \{\{m, o\}, \{f, o\}\}$ and one manager type from the janitorial sector $\{g, j\} = \{\{m, j\}, \{f, j\}\}$. Nature then randomly assigns each manager type drawn to a job slot. In total, there are eight possible combinations of draws in time period 1 listed in Table 3.11. In the second time period, Nature takes another draw of one of the four manager types in Table 2.1. The manager type randomly drawn in time period t advertised in slot a (if in time period 1) is $\{g, s\}_i^{t,a}$. There are an equal proportion of each manager type in the population.

Table 3.11: Eight Possible Job Advertisement Draws in Time Period 1

Draw	First Job Ad Janitorial Sector (M) Manager Gender	Second Job Ad Office Sector (F) Manager Gender	Draw	First Job Ad Office Sector (M) Manager Gender	Second Job Ad Janitorial Sector (F) Manager Gender
1	Male $\{m, j\}$	Male $\{m, o\}$	5	Male $\{m, j\}$	Male $\{m, o\}$
2	Female $\{f, j\}$	Female $\{m, o\}$	6	Female $\{f, j\}$	Female $\{m, o\}$
3	Male $\{m, j\}$	Female $\{f, o\}$	7	Male $\{m, j\}$	Female $\{f, o\}$
4	Female $\{f, j\}$	Male $\{f, o\}$	8	Female $\{f, j\}$	Male $\{f, o\}$

In time period 1, workers face the manager type in job slot 1 $\{g, s\}_i^{1,1}$. They may choose to either 1) apply to the first manager type $y\left(\{g, s\}_i^{1,1}\right) = 1$, 2) pay a cost c to view the manager type in slot 2 $\{g, s\}_i^{1,2}$, or 3) proceed to time period 2 without applying. If the workers chooses to pay c to view the manager type in slot 2, they may then choose to either 1) apply to the first manager type $y\left(\{g, s\}_i^{1,1}\right) = 1$, 2) apply to the second manager type $y\left(\{g, s\}_i^{1,2}\right) = 1$, or 3) proceed to time period 2 without applying.

In time period 2, workers who did not apply in time period 1 face the manager type $\{g, s\}_i^2$. They may choose to either 1) apply to the manager $y\left(\{g, s\}_i^2\right) = 1$, or 2) accept their reservation wage r.

This set-up mimics the experiment design, in which treatment is the assignment of managers to sectors and the order in which those sectors appear in the job advertisement. The first period corresponds to the period in which workers participate in the experiment. The cost of viewing the job advertised second is due to workers' limited attention.

Workers have second-order beliefs $b_i\left(\{g, s\}^{t,a}\right) \in \mathbb{R}$, which are randomly drawn from normally distributed manager type-specific distributions $b_i\left(\{g, s\}^{t,a}\right) \sim N(\mu_{gs}, \sigma_{gs}^2)$. The random variable $b_i\left(\{g, s\}^{t,a}\right)$ is defined to be a function of the worker's subjective distribution of managers' $j \in J$ first-order beliefs b_j about the gender-specific distribution of worker productivity that serves as a sufficient statistic for expected probability of being hired for worker i.[26] Specifically, $b_i\left(\{g, s\}^{t,a}\right) = f[\mathbb{E}_i^{gs}(b_j)]$, where $f(\cdot)$ is a one-to-one mapping with $f(\cdot) > 0$ and the expected value is taken over worker i's manager type-specific subjective distribution of b_j.

The worker's utility function is defined to be a function of second-order beliefs $U\left(b_i\left(\{g, s\}^{t,a}\right)\right)$, where $U'(\cdot) > 0$. So, worker i's utility from applying for a job with manager type $\{g, s\}_i^{t,a}$ is completely determined by the worker's expectation about the probability of being hired based on their second-order beliefs about that manager type $b_i\left(\{g, s\}^{t,a}\right)$.

The worker's application decision can be determined through backwards induction. To simplify the expressions of the workers' decision rules, I use $b^{t,a}$ to represent $U\left(b_i(\{g, s\}^{t,a})\right)$. In time period 2, worker i applies for a job with manager type $\{g, s\}_i^2$ when the utility from applying is greater than the utility from taking the reservation wage $b^2 > U(r)$.

In time period 1, workers who chose to pay c must choose whether to apply for a job with one of the manager types drawn in period 1, $\{g, s\}_i^{1,1}$ or $\{g, s\}_i^{1,2}$, or to proceed to time period 2, where utility is discounted by a factor of δ. Worker i chooses to apply when

[26] I abstract away from worker gender since the cases are symmetric. The distribution of worker productivity is the only point that differs for the genders. Women workers consider managers' first-order beliefs about the distribution of women workers, while men workers consider managers' first-order beliefs about the distribution of men workers.

$$\max(b^{1,1}, b^{1,2}) > \delta \max\left(\mathbb{E}(b_i^2), U(r)\right) \tag{3.1}$$

Lastly, I consider workers at the beginning of time period 1. The worker has three choices. One, the worker can apply to the first manager type $y\left(\{g, s\}_i^{1,1}\right) = 1$ and receive utility $b^{1,1}$. Two, the worker can proceed directly to time period 2 and receive, in expectation, $\delta \max\left(\mathbb{E}(b_i^2), U(r)\right)$. Three, the worker can pay the cost c of observing the second manager type. The expected utility from paying the cost c to observe the second manager type is

$$\frac{1}{2}b^{1,2}(m)\mathbb{1}[b^{1,2}(m) > b^{1,1}] + \frac{1}{2}b^{1,2}(f)\mathbb{1}[b^{1,2}(f) > b^{1,1}] + b^{1,2}[1 - Pr(b^{1,2} > b^{1,1})] - c \tag{3.2}$$

There are three cases. One, the expected utility from either of the two manager types in job slot 2 is higher than the utility from the known manager type in slot 1 $b^{1,2}(m) > b^{1,1}$ and $b^{1,2}(f) > b^{1,1}$. In this case, the worker would choose the manager type in slot 2 over the manager type in slot 1 regardless of the gender of the manager in slot 2; therefore, $1 - Pr(b^{1,2} > b^{1,1})$ is zero and the worker's decision is based on whether the difference between the expected utility of the manager type in slot 2 and the known utility of the manager type in slot 1 is greater than the cost of observing the second manager type $\frac{1}{2}b^{2,2}(m) + \frac{1}{2}b^{2,2}(f) - b^{1,2} > c$.

In the second case, utility from one manager type in job slot 2 is greater than the utility from the manager type in job slot 1 $b^{1,2}(g) > b^{1,1}$, but the utility from the other possible manager type in slot 2 is lower $b^{1,2}(g') < b^{1,1}$.[27] Then, there is an equal probability that the worker chooses the manager type in slot 1 and the manager type in slot 2, based on their draw of the manager type in slot 2. For the worker to pay the cost c to observe that manager type, half the difference between the utility from the manager type in slot 2 that the worker would choose if drawn and the known manager type in slot 1 must be greater than cost of observing the manager type in slot 2. The half captures the probability that the worker draws this more favorable manager type, since the worker would pay the cost of observing the manager type in slot 2, but still choose the manager type in slot 1, if the worker drew the less favorable manager type.

Lastly, it could be the case that $b^{1,2}(m) < b^{1,1}$ and $b^{1,2}(f) < b^{1,1}$. Then, the worker would never pay the cost of observing the manager type in slot 2 based on that draw of manager type in slot 1. Note that this last case must necessarily be true for at least one of the four manager types that the worker can draw since there must exist a maximum, but

[27] Where g is one gender and g' the other.

there is not necessarily the converse case. Even though there must be a minimum, it is not necessarily true that there is a sufficiently large difference between the minimum and the expected utility from the manager types in the opposite sector to induce the worker to pay the cost to observe the manager type in slot 2.

In summary, the worker applies in period 1 when either $b^{1,1}$ or $\max(b^{1,1}, b^{1,2})$, conditional on paying the cost c to observe $b^{1,2}$, are larger than the expected utility from taking another draw of manager types in the next time period $\delta \max\left(\mathbb{E}(b_i^2), U(r)\right)$. There are, therefore, three functions of beliefs that I include in the empirical specification: $b^{1,1}$, $\max(b^{1,1}, b^{1,2})$, and $\delta \max\left(\mathbb{E}(b_i^2), U(r)\right)$. To make the last function tractable, I summarize $\delta \max\left(\mathbb{E}(b_i^2), U(r)\right)$ with the average of all beliefs except the manager in job slot 1 in time period 1, since a worker who proceeds to time period 2 must be searching for a manager type not observed in time period 1. I do not similarly exclude the manager type in job slot 2 in the average since the worker decides endogenously whether to observe that manager type.

J Regression Tables

Table 3.12: Main Specification, Men

	(1)	(2)	(3)	(4)
Belief about First Manager	0.002**	0.002**	0.002**	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
T1: Male Managers	0.142***	0.175***	0.181***	0.205
	(0.038)	(0.052)	(0.055)	(0.176)
T2: Female Managers	0.235***	0.235***	0.250***	0.263
	(0.050)	(0.050)	(0.061)	(0.183)
T3: Sector-Matching Managers	0.162***	0.184***	0.196***	0.200
	(0.037)	(0.048)	(0.052)	(0.178)
T4: Non-Sector-Matching Managers	0.197***	0.211***	0.223***	0.251
	(0.041)	(0.042)	(0.049)	(0.175)
Most Favorable Belief of Facing Managers		-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)
Average Belief about Non-First Manager		0.003	0.003	0.003
		(0.002)	(0.002)	(0.002)
Prefer Gender of First Manager			-0.010	0.003
			(0.052)	(0.051)
Prefer Opposite-Gender of First Manager			-0.034	-0.012
			(0.053)	(0.055)
Prefer Sector of First Manager			-0.044	-0.058
			(0.055)	(0.058)
Prefer Opposite Sector of First Manager			0.031	0.009
			(0.053)	(0.051)
Controls	N	N	N	Y
Observations	351	351	351	341
R-Squared	0.193	0.198	0.202	0.257

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Robust standard errors are reported in parentheses. Controls include the wave, city, and job board in which the worker was recruited, as well as whether the worker reports actively searching for a job.

Table 3.13: Main Specification, Women

	(1)	(2)	(3)	(4)
Belief about First Manager	-0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
T1: Male Managers	0.055*	0.063*	0.079*	0.026
	(0.032)	(0.036)	(0.046)	(0.054)
T2: Female Managers	0.097**	0.125***	0.130**	0.075
	(0.039)	(0.047)	(0.051)	(0.066)
T3: Sector-Matching Managers	0.050*	0.070*	0.082	0.041
	(0.028)	(0.041)	(0.053)	(0.064)
T4: Non-Sector-Matching Managers	0.108**	0.123**	0.135**	0.086
	(0.045)	(0.051)	(0.061)	(0.073)
Most Favorable Belief of Facing Managers		-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)
Average Belief about Non-First Manager		0.002	0.002	0.003*
		(0.001)	(0.001)	(0.001)
Prefer Gender of First Manager			-0.084**	-0.082**
			(0.034)	(0.036)
Prefer Opposite-Gender of First Manager			-0.010	-0.040
			(0.051)	(0.051)
Prefer Sector of First Manager			0.039	0.054
			(0.051)	(0.050)
Prefer Opposite Sector of First Manager			0.015	0.016
			(0.050)	(0.046)
Controls	N	N	N	Y
Observations	223	223	223	220
R-Squared	0.084	0.095	0.111	0.309

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Robust standard errors are reported in parentheses. Controls include the wave, city, and job board in which the worker was recruited, as well as whether the worker reports actively searching for a job.

Table 3.14: Other Specification of Treatment

	(1)	(2)
Belief about First Manager	0.002*	-0.000
	(0.001)	(0.001)
Most Favorable Belief of Facing Managers	-0.001	-0.001
	(0.002)	(0.001)
Average Belief about Non-First Manager	0.003	0.003*
	(0.002)	(0.002)
T1: Male Managers, Jan First	0.176	0.035
	(0.185)	(0.081)
T2: Female Managers, Jan First	0.200	0.077
	(0.192)	(0.082)
T3: Sector-Matching Managers, Jan First	0.207	-0.045
	(0.182)	(0.085)
T4: Non-Sector-Matching Managers, Jan First	0.311*	0.042
	(0.184)	(0.086)
T1: Male Managers, Off First	0.178	-0.016
	(0.184)	(0.058)
T2: Female Managers, Off First	0.271	0.054
	(0.196)	(0.076)
T3: Sector-Matching Managers, Off First	0.135	0.082
	(0.193)	(0.070)
T4: Non-Sector-Matching Managers, Off First	0.134	0.119
	(0.183)	(0.097)
Prefer Gender of First Manager	0.013	-0.078**
	(0.051)	(0.037)
Prefer Opposite-Gender of First Manager	-0.017	-0.041
	(0.056)	(0.053)
Prefer Sector of First Manager	-0.071	0.044
	(0.061)	(0.050)
Prefer Opposite Sector of First Manager	0.031	0.017
	(0.054)	(0.045)
Controls	Y	Y
Observations	341	220
R-Squared	0.270	0.326

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Robust standard errors are reported in parentheses. Controls include the wave, city, and job board in which the worker was recruited, as well as whether the worker reports actively searching for a job.

Table 3.15: Other Specifications of Alternative Beliefs, Men

	(1)	(2)	(3)	(4)	(5)
Belief about First Manager	0.003**	0.002**	0.002**	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
First-Order Belief	-0.001				
	(0.001)				
Most Favorable Belief of Facing Managers	-0.001			-0.001	
	(0.001)			(0.001)	
Average Belief about Non-First Manager	0.003	0.004*			
	(0.002)	(0.002)			
Most Favorable Belief of All Other Managers			0.000		
			(0.001)		
Belief about Second Manager		-0.002			-0.000
		(0.001)			(0.001)
Most Favorable Belief of Non-Facing Managers				0.002*	
				(0.001)	
Average Belief of Non-Facing Managers					0.003*
					(0.002)
T1: Male Managers	0.191	0.210	0.168	0.174	0.210
	(0.175)	(0.171)	(0.173)	(0.178)	(0.171)
T2: Female Managers	0.255	0.247	0.252	0.218	0.247
	(0.182)	(0.178)	(0.185)	(0.189)	(0.178)
T3: Sector-Matching Managers	0.187	0.196	0.171	0.165	0.196
	(0.177)	(0.172)	(0.175)	(0.180)	(0.172)
T4: Non-Sector-Matching Managers	0.237	0.237	0.228	0.191	0.237
	(0.174)	(0.171)	(0.178)	(0.181)	(0.171)
Prefer Gender of First Manager	0.004	0.005	0.006	0.011	0.005
	(0.051)	(0.051)	(0.052)	(0.051)	(0.051)
Prefer Opposite-Gender of First Manager	-0.007	-0.008	-0.013	-0.005	-0.008
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
Prefer Sector of First Manager	-0.056	-0.064	-0.064	-0.068	-0.064
	(0.058)	(0.058)	(0.056)	(0.058)	(0.058)
Prefer Opposite Sector of First Manager	0.011	0.010	0.008	0.015	0.010
	(0.051)	(0.051)	(0.052)	(0.052)	(0.051)
Controls	Y	Y	Y	Y	Y
Observations	341	341	341	341	341
R-Squared	0.259	0.260	0.252	0.259	0.260

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Robust standard errors are reported in parentheses. Controls include the wave, city, and job board in which the worker was recruited, as well as whether the worker reports actively searching for a job. Column (1) adds first-order beliefs. Column (2) replaces the most favorable belief of the two facing managers with the belief about the second manager type. Column (3) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the most favorable belief of all managers, excepting the first. Column (4) replaces the average belief of all other managers, except the first, with the most favorable belief of the non-facing managers. Column (5) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the belief about the second manager type and the average belief of the non-facing managers.

Table 3.16: Other Specifications of Alternative Beliefs, Women

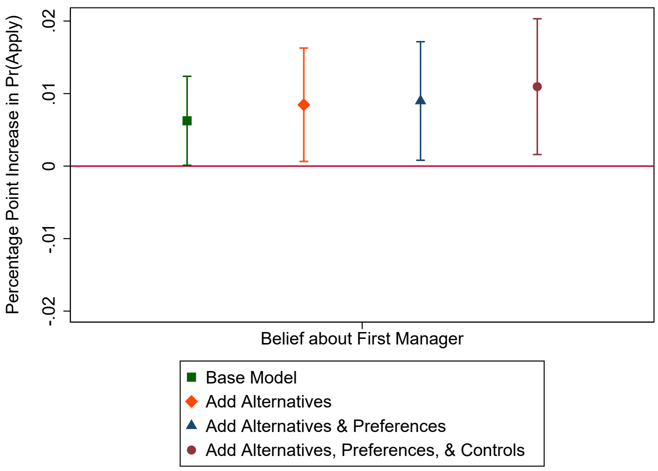
	(1)	(2)	(3)	(4)	(5)
Belief about First Manager	0.000	-0.001	-0.000	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
First-Order Belief	0.000				
	(0.001)				
Most Favorable Belief of Facing Managers	-0.001			-0.001	
	(0.001)			(0.001)	
Average Belief about Non-First Manager	0.002	0.004**			
	(0.002)	(0.002)			
Most Favorable Belief of All Other Managers			0.002**		
			(0.001)		
Belief about Second Manager		-0.002**			-0.001
		(0.001)			(0.001)
Most Favorable Belief of Non-Facing Managers				0.002**	
				(0.001)	
Average Belief of Non-Facing Managers					0.003**
					(0.001)
T1: Male Managers	0.027	-0.036	-0.049	-0.050	-0.036
	(0.054)	(0.058)	(0.052)	(0.059)	(0.058)
T2: Female Managers	0.075	0.052	-0.019	0.019	0.052
	(0.067)	(0.063)	(0.054)	(0.060)	(0.063)
T3: Sector-Matching Managers	0.040	-0.018	-0.045	-0.019	-0.018
	(0.065)	(0.059)	(0.051)	(0.059)	(0.059)
T4: Non-Sector-Matching Managers	0.086	0.051	-0.004	0.009	0.051
	(0.074)	(0.072)	(0.068)	(0.073)	(0.072)
Prefer Gender of First Manager	-0.079**	-0.083**	-0.091**	-0.083**	-0.083**
	(0.036)	(0.037)	(0.038)	(0.036)	(0.037)
Prefer Opposite-Gender of First Manager	-0.040	-0.036	-0.035	-0.034	-0.036
	(0.051)	(0.050)	(0.052)	(0.050)	(0.050)
Prefer Sector of First Manager	0.053	0.051	0.050	0.049	0.051
	(0.050)	(0.051)	(0.050)	(0.051)	(0.051)
Prefer Opposite Sector of First Manager	0.018	0.015	0.011	0.012	0.015
	(0.047)	(0.045)	(0.046)	(0.044)	(0.045)
Controls	Y	Y	Y	Y	Y
Observations	220	220	220	220	220
R-Squared	0.310	0.321	0.305	0.314	0.321

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$
Robust standard errors are reported in parentheses. Controls include the wave, city, and job board in which the worker was recruited, as well as whether the worker reports actively searching for a job. Column (1) adds first-order beliefs. Column (2) replaces the most favorable belief of the two facing managers with the belief about the second manager type. Column (3) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the most favorable belief of all managers, excepting the first. Column (4) replaces the average belief of all other managers, except the first, with the most favorable belief of the non-facing managers. Column (5) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the belief about the second manager type and the average belief of the non-facing managers.

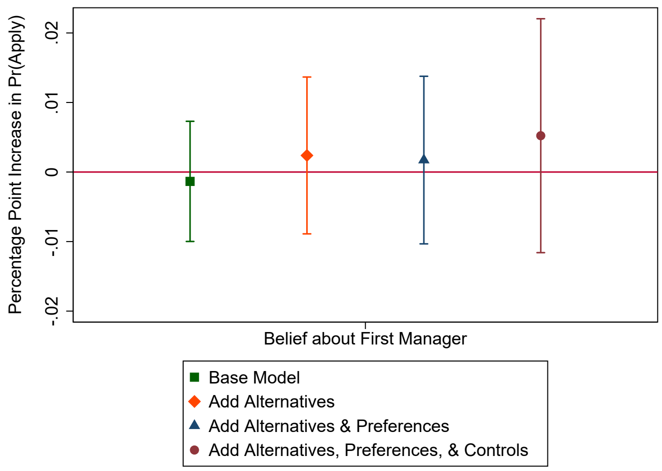
K Probit Regressions

Figure 3.11: Regression Coefficient on Belief about First Manager by Gender, Main Specification

(a) Men



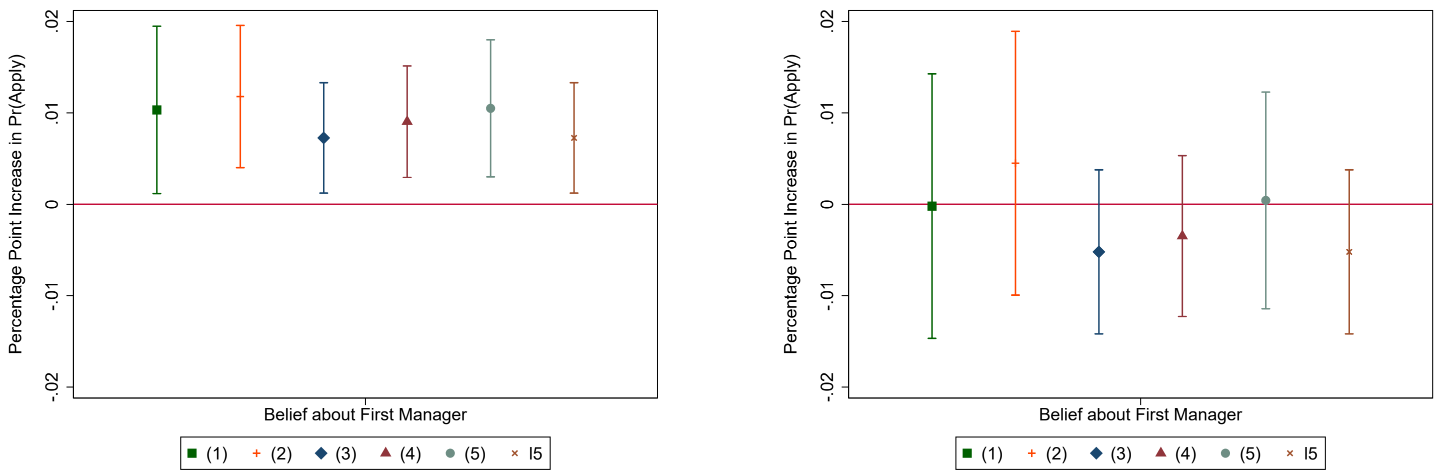
(b) Women



Notes: 95% confidence intervals are shown. Standard errors are robust. Coefficient is scaled to a one unit increase in the favorability of beliefs. Panel (a) reports the coefficient on the regression with men only and panel (b) reports the same for women only. In each graph, the first estimate includes only the vector of treatment variables and the belief of interest. The second estimate adds measures of alternative beliefs, the third estimate adds preferences, and the fourth estimate includes controls for the wave, city, and job board in which the worker was recruited and whether the worker reports actively searching for a job.

Figure 3.12: Regression Coefficient on Belief about First Manager by Gender, Other Specifications

(a) Men (b) Women

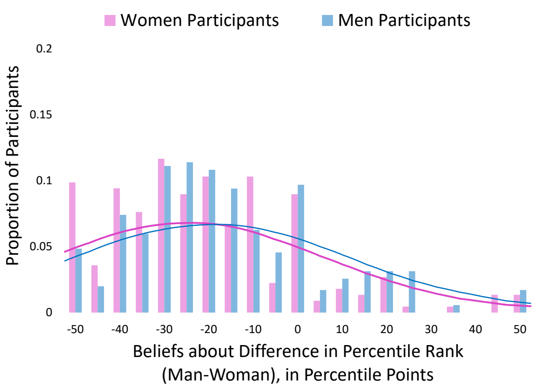


Notes: 95% confidence intervals are shown. Standard errors are robust. Coefficient is scaled to a one unit increase in the favorability of beliefs. Panel (a) reports the coefficient on the regression with men only and panel (b) reports the same for women only. In each graph, specification (1) interacts treatment with first manager type. Specification (2) adds first-order beliefs. Specification (3) replaces the most favorable belief of the two facing managers with the belief about the second manager type. Specification (4) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the most favorable belief of all managers, excepting the first. Specification (5) replaces the average belief of all other managers, except the first, with the most favorable belief of the non-facing managers. Specification (6) replaces the most favorable belief of the two facing managers and the average belief of all other managers, except the first, with the belief about the second manager type and the average belief of the non-facing managers.

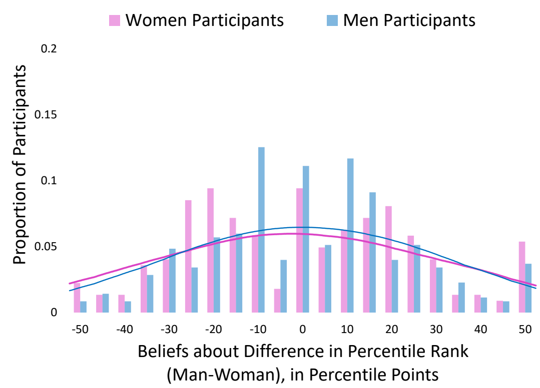
L Cardinal Belief Distributions

Figure 3.13: Cardinal Distributions of Second-Order Beliefs

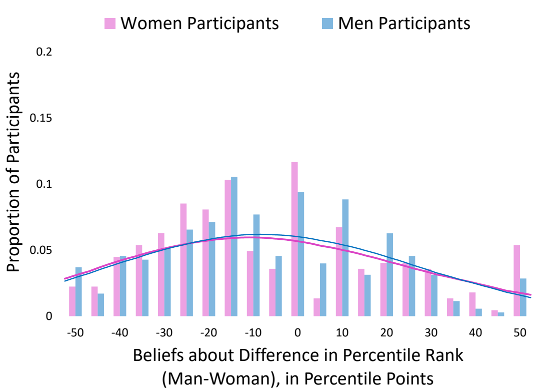
(a) Female Manager in Female Sector



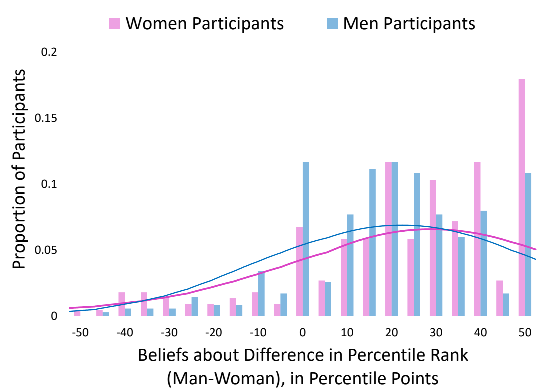
(b) Female Manager in Male Sector



(c) Male Manager in Female Sector

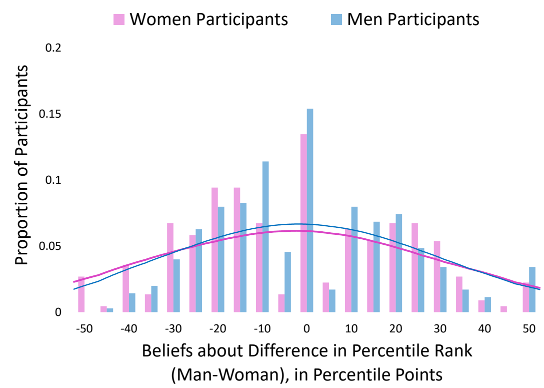


(d) Male Manager in Male Sector



Notes: Distributions are fitted with lines estimated using kernel densities.

Figure 3.14: Cardinal Distribution of First-Order Beliefs



Notes: Distribution is fitted with a line estimated using a kernel density.

M Proofs

Proof of Proposition 1.

Step 1: $DC|O|a_i \to DC|O|(a_i-1)$ and $DC|O|a_i \to DC|NO|a_i$.

These binary relations follow from the fact that all strategies in $DC|O|a_i$ are dominant strategies.

Step 2: $DC|O|k \dashrightarrow DC|O|(k-1)$ for every $k \in \{2, \ldots, a_i - 1\}$.

Take any strategy $Q_i \in \mathrm{DC|O|}(k-1)$, and let d be i's district school. Let $\hat{S}$ be the set of schools student i ranks above their district school in Q_i. Because $(k-1) < a_i$, there exists a school s^* that student i prefers to their district school such that $s^* \notin \hat{S}$. Let Q_i^* be the strategy in $\mathrm{DC|O|}k$ for which $\hat{S}^* = \hat{S} \cup \{s^*\}$ is the set of schools i ranks above their district school (i.e., Q_i^* is constructed from Q_i by "moving" s^* from below to above the district school and ranking s^* truthfully).

Preferences in (3.3) to (3.6) hold for all strict priority profiles that can result from the tiebreaker. Because DA is strategy-proof (Dubins and Friedman, 1981),

$$DA_i(Q_i^*, Q_{-i})\ Q_i^*\ DA_i(Q_i, Q_{-i}), \qquad \text{for all } Q_{-i}. \tag{3.3}$$

Then, since Q_i^* ranks schools $\hat{S} \cup \{s^*, d\}$ identically to student i's true preferences P_i,

$$\begin{aligned} &DA_i(Q_i^*, Q_{-i})\ R_i\ DA_i(Q_i, Q_{-i}), \\ &\qquad \text{for all } Q_{-i} \text{ such that } DA_i(Q_i^*, Q_{-i}), DA_i(Q_i, Q_{-i}) \in \hat{S}^* \cup \{d\}. \end{aligned} \tag{3.4}$$

But because $\hat{S}$ is the set of schools student i ranks above their district school in Q_i, and $\hat{S}^*$ the set of schools student i ranks above their district school in Q_i^*, we have $DA_i(Q_i, Q_{-i}), DA_i(Q_i^*, Q_{-i}) \in \hat{S}^* \cup \{d\}$ for all Q_{-i}. Hence, (3.4) implies

$$DA_i(Q_i^*, Q_{-i})\ R_i\ DA_i(Q_i, Q_{-i}), \qquad \text{for all } Q_{-i}. \tag{3.5}$$

Lastly, consider any Q_{-i}^* s.t. all students except i and j rank their district schools first. Let student $j \neq i$ with district school s^* rank d above s^* in Q_j^*. Then, all students except i and j are assigned to their district schools and there is one slot open in s^* and one slot open in d. Under Q_i, student i is assigned to d since they apply to d before s^* and are admitted. Under Q_i^*, student i is assigned to s^* since they apply to s^* before d and hold a seat there until student j applies to d and is admitted. Then, we have

$$DA_i(Q_i^*, Q_{-i}^*) = s^*\ P_i\ d = DA_i(Q_i, Q_{-i}^*). \tag{3.6}$$

Together, (3.5) and (3.6) show that Q_i^* dominates Q_i. Because Q_i is an arbitrary strategy in DC|O|$(k-1)$, this completes the proof of this step.

Step 3: DC/NO/k $\dashrightarrow$ DC/NO/$(k-1)$ for every $k \in \{2, \dots, a_i\}$. Take any strategy $Q_i \in$ DC|NO|$(k-1)$, and let d be i's district school. Let $\hat{S}$ be the set of schools i ranks above their district school in Q_i. Because $(k-1) < a_i$, there exists a school s^* that i prefers to their district school such that $s^* \notin \hat{S}$. Let Q_i^* be the strategy in DC|NO|k for which

(a) $\hat{S}^* = \hat{S} \cup \{s^*\}$ is the set of schools i ranks above their district school,

(b) schools in $\hat{S}$ are ranked exactly as in Q_i, and

(c) s^* is ranked directly above the district school, i.e., $s^* \; Q_i^* \; d$, and $s \; Q_i^* \; s^*$ for all $s \in \hat{S}$ (i.e., Q_i^* is constructed from Q_i by "moving" s^* from below to right above the district school, and not changing any other rankings compared to Q_i).

Observe that, by construction, Q_i^* is indeed in DC|NO|k (in particular, by (b), Q_i^* is not-Ordered because Q_i is not-Ordered). Preferences in (3.7) to (3.11) hold for all strict priority profiles that can result from the tiebreaker. By construction,

$$\begin{aligned} &DA_i(Q_i^*, Q_{-i}) = DA_i(Q_i, Q_{-i}), \\ &\qquad \text{for all } Q_{-i} \text{ such that } DA_i(Q_i^*, Q_{-i}), DA_i(Q_i, Q_{-i}) \in \hat{S}. \end{aligned} \tag{3.7}$$

Also, because i prefers s^* to d,

$$\begin{aligned} &DA_i(Q_i^*, Q_{-i}) \; P_i \; DA_i(Q_i, Q_{-i}), \\ &\qquad \text{for all } Q_{-i} \text{ such that } DA_i(Q_i^*, Q_{-i}) = s^* \text{ and } DA_i(Q_i, Q_{-i}) = d. \end{aligned} \tag{3.8}$$

and

$$DA_i(Q_i^*, Q_{-i}) = d \text{ implies } DA_i(Q_i, Q_{-i}) = d. \tag{3.9}$$

Together, (3.7) to (3.9) imply

$$DA_i(Q_i^*, Q_{-i}) \; R_i \; DA_i(Q_i, Q_{-i}), \qquad \text{for all } Q_{-i}. \tag{3.10}$$

Lastly, consider any Q_{-i}^* s.t. all students except i and j rank their district schools first. Let student $j \neq i$ with district school s^* rank d above s^* in Q_j^*. Then, all students except i and j are assigned to their district schools in round 1 and there is one slot open in s^* and one slot open in d. Under Q_i, student i is assigned to d since they apply to d before s^* and are admitted. Under Q_i^*, student i is assigned to s^* since they apply to s^* before d and are admitted, while student j is assigned to d. Then, we have

$$DA_i(Q_i^*, Q_{-i}^*) = s^* \ P_i \ d = DA_i(Q_i, Q_{-i}^*). \tag{3.11}$$

Together, (3.5) and (3.6) show that Q_i^* dominates Q_i. Because Q_i is an arbitrary strategy in DC|O|$(k-1)$, this completes the proof of this step.

Step 4: DC|O|k $\dashrightarrow$ DC|NO|k for every $k \in \{2, \ldots, a_i - 1\}$. Take any strategy $Q_i \in$ DC|NO|k, and let d be i's district school. Let $\hat{S}$ be the set of schools i ranks above their district school in Q_i. Let Q_i^* be the strategy in DC|O|k for which $\hat{S}$ is the set of schools i ranks above their district school (i.e., Q_i^* is constructed from Q_i by truthfully re-ordering the schools in $\hat{S}$).

By the same argument that lead to (3.5) in Step 2, we have

$$DA_i(Q_i^*, Q_{-i}) \ R_i \ DA_i(Q_i, Q_{-i}), \qquad \text{for all } Q_{-i}. \tag{3.12}$$

Again, (3.12) holds for all strict priority profiles that can result from the tiebreaker.

Because Q_i is District-Consistent but not-Ordered, there exists two schools $\tilde{s}, \hat{s} \in \hat{S}$ such that $\tilde{s} \ Q_i \ \hat{s}$ but $\hat{s} \ P_i \ \tilde{s}$.

Consider any Q_{-i}^* in which

(a) every student $j \neq i$ whose district school is ranked above $\tilde{s}$ in Q_i ranks their district school first,

(b) every student $j \neq i$ whose district school is $\tilde{s}$ ranks $\tilde{s}$ first, *except* for one of these students who ranks $\hat{s}$ first and

(c) every student $j \neq i$ whose district school is $\hat{s}$ ranks $\hat{s}$ first , *except* for one of these students who ranks d first,

(d) every student $j \neq i$ whose district school is d ranks d first.

Let $h \neq i$ be the student whose district school is $\tilde{s}$ and who ranks $\hat{s}$ first. Let $l \neq i$ be the student whose district school is $\hat{s}$ and who ranks d first. In both $DA(Q_i, Q_{-i}^*)$ and $DA(Q_i^*, Q_{-i}^*)$, every student but i, h, and l is assigned to their district school. This implies that, in both $DA(Q_i, Q_{-i}^*)$ and $DA(Q_i^*, Q_{-i}^*)$, student i is necessarily rejected from any school they applies to other than $\tilde{s}$, $\hat{s}$, and d.

Under both $DA(Q_i, Q_{-i}^*)$ and $DA(Q_i^*, Q_{-i}^*)$, l initially applies to d and is held at d as no more than q_d students initially apply to d (only students with d as their district school apply to d and i who has d as their district school does not). Similarly, h initially applies to $\hat{s}$ and is held at $\hat{s}$ as no more than $q_{\hat{s}}$ students initially apply to $\hat{s}$ (only students with $\hat{s}$ as their district school apply to $\hat{s}$ and l who has $\hat{s}$ as their district school does not). This remains true at least until h applies to either d or $\hat{s}$.

In $DA(Q_i, Q^*_{-i})$, after having been rejected from all the schools student i ranks above $\tilde{s}$, i applies to $\tilde{s}$. Because there is a seat available at $\tilde{s}$, i is held at $\tilde{s}$, all students are assigned, and the assignment is therefore final. That is,

$$DA_i(Q_i, Q^*_{-i}) = \tilde{s}. \tag{3.13}$$

In contrast, in $DA(Q^*_i, Q^*_{-i})$, after having been rejected from all the schools student i ranks above $\hat{s}$, i applies to $\hat{s}$. When i applies to $\hat{s}$, all seats at $\hat{s}$ are already occupied, one of which by h who does not have $\hat{s}$ as their district school. Thus, i is held at $\hat{s}$ exactly half the time, i.e., when i is given higher priority than h by the tie-breaking rule. That is,

$$DA_i(Q^*_i, Q^*_{-i}) = \hat{s}, \qquad \text{with probability one-half.} \tag{3.14}$$

If student i is rejected from $\hat{s}$ (which occurs with probability one half), i ends up applying to school $\tilde{s}$, and by the above argument, i is assigned to $\tilde{s}$. That is,

$$DA_i(Q^*_i, Q^*_{-i}) = \tilde{s}, \qquad \text{with probability one-half.} \tag{3.15}$$

Because, by construction, student i prefers $\hat{s}$ to $\tilde{s}$, (3.13), (3.14), and (3.15) together imply i's expected utility under $DA(Q^*_i, Q^*_{-i})$ is larger than under $DA(Q_i, Q^*_{-i})$. Together with the fact that (3.12) holds for every realization of the tiebreaker, this shows that Q^*_i dominates Q_i. Because Q_i is an arbitrary strategy in DC-k, this completes the proof of this step.

Proof of Proposition 2.

Let $\bar{Q}_i$ be any strategy $Q_i \in \mathrm{DC|O|}x$. (i) Because i's district school d has as many seats are there are students in d's district,

$$IA_i(Q^{safe}_i, Q_{-i}) = d \qquad \text{for all } Q_{-i}. \tag{3.16}$$

Also, because DA is strategy-proof Dubins and Freedman (1981),

$$DA_i(\bar{Q}^{DA}_i, Q_{-i}) \; \bar{Q}^{DA}_i \; DA_i(Q_i, Q_{-i}), \qquad \text{for all } Q_{-i} \text{ and all } Q_i. \tag{3.17}$$

In particular,

$$DA_i(\bar{Q}_i^{DA}, Q_{-i}) \ \bar{Q}_i^{DA} \ DA_i(Q_i^{safe}, Q_{-i}) = d, \qquad \text{for all } Q_{-i}. \tag{3.18}$$

But because $\bar{Q}_i^{DA}$ is District-Consistent, this implies

$$DA_i(\bar{Q}_i^{DA}, Q_{-i}^{DA}) \ R_i \ d = IA_i(Q_i^{safe}, Q_{-i}^{IA}), \qquad \text{for all } Q_{-i}^{DA}, Q_{-i}^{IA}, \tag{3.19}$$

the desired result.

(ii) Let s^* be the school that is ranked first in $\bar{Q}_i^{DA}$. Because $\bar{Q}_i^{DA}$ is District-Consistent but is not District-First, $s^* \ P_i \ d$. Let $\tilde{Q}_{-i}^{DA}$ be such that

(a) every student $j \neq i$ whose district school is different from d and s^* ranks their district school first,

(b) all but one student $j \neq i$ whose district is s^* ranks s^* first, with the student who does not rank s^* first ranking d first, and

(c) every student $j \neq i$ whose district school is d ranks d first.

By (c) and because i ranks s^* first, the student whose district school is s^* and who ranks d first is assigned to d. Every other student $j \neq i$ is assigned to their district school. This implies that there is an available seat school s^* and that i is assigned to s^* (as i ranks s^* first). But then we have

$$DA_i(\bar{Q}_i^{DA}, \tilde{Q}_{-i}^{DA}) = s^* \ P_i \ d = IA_i(Q_i^{safe}, Q_{-i}^{IA}), \qquad \text{for all } Q_{-i}^{IA}, \tag{3.20}$$

the desired result.

Proof of Proposition 3. Let f be student i's most-preferred school, and l student i's least-preferred school.

(i) Let $\tilde{Q}_i^{IA}$ be such that every student $j \neq i$ ranks their district school first. Thus, i is rejected from f when applying to f in the first round of IA and others report $\overline{Q}_i^{IA}$. Because only $q_d - 1$ students apply to d in the first round, i is then assigned to d, that is,

$$IA_i(Q_i^{Risky}, \overline{Q}_i^{IA}) = d, \tag{3.21}$$

which implies

$$d \ R_i \ \min_{R_i}\{IA_i(Q_i^{Risky}, Q_{-i}) \text{ for all } Q_{-i}\}. \tag{3.22}$$

But by (3.19), we have

$$\min_{R_i}\{DA_i(\bar{Q}_i^{DA}, Q_{-i}) \text{ for all } Q_{-i}\}\ R_i\ d, \tag{3.23}$$

which combined with (3.22) implies the desired result.

(ii) Let $\hat{Q}_i^{IA}$ be such that

(a) every student $j \neq i$ ranks their district school first, except for one student j^* whose district school is l,

(b) student j^* ranks school d first.

By (a), i is rejected from school f when applying to f in the first round. By (a) and (b), all seats are occupied at the end of the first round, except for one seat at school l. Thus,

$$IA_i(Q_i^{Risky}, \hat{Q}_i^{IA}) = l, \tag{3.24}$$

which implies

$$\min_{R_i}\{IA_i(Q_i^{Risky}, Q_{-i}) \text{ for all } Q_{-i}\} = l. \tag{3.25}$$

But by (3.19), we have

$$\min_{R_i}\{DA_i(\bar{Q}_i^{DA}, Q_{-i}) \text{ for all } Q_{-i}\}\ R_i\ d. \tag{3.26}$$

By assumption, d is not i least preferred school, i.e., $d\ P_i\ l$. Therefore (3.25) and (3.26) imply the desired result.

N DA Instructions

Instructions

This is an experiment in the economics of decision making. The instructions are simple, and if you follow them carefully and make good decisions, you might earn a considerable amount of money. In this experiment, we simulate a procedure to allocate students to schools. The procedure, payment rules, and student allocation method are described below. Do not communicate with each other during the experiment. If you have questions at any point during the experiment, raise your hand and the experimenter will help you.

Procedure

- There are 36 participants in this experiment. You are participant 1.
- In this simulation, 36 school slots are available across seven schools. These schools differ in size, geographic location, specialty, and quality of instruction in each specialty. Each school slot is allocated to one participant. There are three slots each at schools A and B, and six slots each at schools C, D, E, F and G.
- **Your payoff** amount depends on the school slot you hold at the end of the experiment. These amounts reflect the desirability of the school in terms of location, specialty and quality of instruction.

Slot received at School:	A	B	C	D	E	F	G
Payoff to Participant 1 (in dollars)	16	19	12	5	8	14	10

The table is explained as follows:

- You will be paid $16 if you hold a slot at school A at the end of the experiment.
- You will be paid $19 if you hold a slot at school B at the end of the experiment.
- You will be paid $12 if you hold a slot at school C at the end of the experiment.
- You will be paid $5 if you hold a slot at school D at the end of the experiment.
- You will be paid $8 if you hold a slot at school E at the end of the experiment.
- You will be paid $14 if you hold a slot at school F at the end of the experiment.

- You will be paid $10 if you hold a slot at school G at the end of the experiment.

***NOTE* different participants might have different payoff tables.** That is, payoff by school might be different for different participants.

- During the experiment, each participant first completes the Decision Sheet by indicating school preferences. The Decision Sheet is the last page of this packet. Note that you need to rank all seven schools in order to indicate your preferences.
- After all participants have indicated their preferences, the experimenter will collect the preferences and start the allocation process.
- Once the allocations are determined, the experimenter will inform each participant of his/her allocation slot and respective payoff.
- In this experiment, participants are defined as belonging to the following school districts.
    - Participants #1 - #3 live within the school district of school A,
    - Participants #4 - #6 live within the school district of school B,
    - Participants #7 - #12 live within the school district of school C,
    - Participants #13 - #18 live within the school district of school D,
    - Participants #19 - #24 live within the school district of school E,
    - Participants #25 - #30 live within the school district of school F,
    - Participants #31 - #36 live within the school district of school G,
- A priority order is determined for each school. Each participant is assigned a slot at the **best possible** school he/she reported that is consistent with the priority order below.
- The priority order for each school is separately determined as follows:
    - **High Priority Level:** Participants who live within the school district. Since the number of High priority participants at each school is equal to the school capacity, each High priority participant is guaranteed an assignment which is at least as good as his/her district school based on the ranking indicated in his/her Decision Sheet.

- **Low Priority Level:** Participants who do not live within the school district. The priority among the Low priority students is based on their respective order in a fair lottery. This means each participant has an equal chance of being the first in the line, the second in the line, ..., as well as the last in the line. To determine this lottery, the experimenter will draw participant's ID number randomly, one at a time. The sequence of the draw determines the order in the lottery.

- Once the priorities are determined, the allocation of school slots is obtained as follows:
  - An application to the first ranked school in the Decision Sheet is sent for each participant.
  - Throughout the allocation process, a school can hold no more applications than its number of slots.
    If a school receives more applications than its capacity, then it rejects the students with lowest priority orders. The remaining applications are retained.
  - Whenever an applicant is rejected at a school, his/her application is sent to the next highest school he/she reported.
  - Whenever a school receives new applications, these applications are considered together with the retained applications for that school. Among the retained and new applications, the lowest priority ones in excess of the number of the slots are rejected, while remaining applications are retained.
  - The allocation is finalized when no more applications can be rejected. Each participant is assigned a slot at the school that holds his/her application at the end of the process.

An Example:
We will go through a simple example to illustrate how the allocation method works.

Students and Schools: In this example, there are six students, 1-6, and four schools, Clair Erie, Huron and Ontario.

Student ID Numbers: 1,2,3,4,5,6	Schools: Clair, Erie, Huron, Ontario

Slots and Residents: there are two slots at each Clair and Erie, and one slot each at Huron and Ontario. Residents of districts are indicated in the table below.
Lotteries: In this example, the lottery produced the following order
Submitted School Rankings: The students have submitted the following school rankings

Schools	Slot 1	Slot 2	District Residents
Clair	☐	☐	1 2
Erie	☐	☐	3 4
Huron	☐		5
Ontario	☐		6

1-2-3-4-5-6

	1st Choice	2nd Choice	3rd Choice	Last Choice
Student 1	Huron	Clair	Ontario	Erie
Student 2	Huron	Ontario	Clair	Erie
Student 3	Ontario	Clair	Erie	Huron
Student 4	Huron	Clair	Ontario	Erie
Student 5	Ontario	Huron	Clair	Erie
Student 6	Clair	Erie	Ontario	Huron

Priority: School priorities first depend on whether the school is a district school, and next on the lottery order:

	Resident	Non-Resident
Priority order at Clair	**1-2**	3-4-5-6
Priority order at Erie	**3-4**	1-2-5-6
Priority order at Huron	**5**	1-2-3-4-6
Priority order at Ontario	**6**	1-2-3-4-5

The allocation method consists of the following steps:
Step 1: Each student applies to his/her first choice: students 1, 2 and 4 apply to Huron, students 3 and 5 apply to Ontario, and student 6 applies to Clair.

- Clair holds the application of student 6.
- Huron holds the application of student 1 and rejects students 2 and 4
- Ontario holds the application of student 3 and rejects student 5

Step 2: Each student rejected in Step 1 applies to his/her next choice: student 2 applies to Ontario, student 4 applies to Clair, and student 5 applies to Huron.

Applicants		School		Hold	Reject
6	$\longrightarrow$	Clair	$\longrightarrow$	$\boxed{6}$ $\square$	
	$\longrightarrow$	Erie	$\longrightarrow$	$\square$ $\square$	
1,2,4	$\longrightarrow$	Huron	$\longrightarrow$	$\boxed{1}$	2,4
3,5	$\longrightarrow$	Ontario	$\longrightarrow$	$\boxed{3}$	5

- Clair considers the application of student 4 together with the application of student 6, which was on hold. It holds both applications.
- Huron considers the application of student 5 together with the application of student 1, which was on hold. It holds the application of student 5 and rejects student 1.
- Ontario considers the application of student 2 together with the application of student 3, which was on hold. It holds the application of student 2 and rejects student 3.

Hold	New Applicants		School		Hold	Reject
$\boxed{6}$ $\square$	4	$\longrightarrow$	Clair	$\longrightarrow$	$\boxed{6}$ $\boxed{4}$	
$\square$ $\square$		$\longrightarrow$	Erie	$\longrightarrow$	$\square$ $\square$	
$\boxed{1}$	5	$\longrightarrow$	Huron	$\longrightarrow$	$\boxed{5}$	1
$\boxed{3}$	2	$\longrightarrow$	Ontario	$\longrightarrow$	$\boxed{2}$	3

Step 3: Each student rejected in Step 2 applies to his/her next choice: Students 1 and 3 apply to Clair.

- Clair considers the applications of students 1 and 3 together with the applications of students 4 and 6, which were on hold. It holds the applications of students 1 and 3 and rejects students 4 and 6.

Hold	New Applicants		School		Hold	Reject
$\boxed{6}$ $\boxed{4}$	1,3	$\longrightarrow$	Clair	$\longrightarrow$	$\boxed{1}$ $\boxed{3}$	4,6
$\square$ $\square$		$\longrightarrow$	Erie	$\longrightarrow$	$\square$ $\square$	
$\boxed{5}$		$\longrightarrow$	Huron	$\longrightarrow$	$\boxed{5}$	
$\boxed{2}$		$\longrightarrow$	Ontario	$\longrightarrow$	$\boxed{2}$	

Step 4: Each student rejected in Step 3 applies to his/her next choice: Student 4 applies to Ontario and student 6 applies to Erie.

- Ontario considers the application of student 4 together with the application of student 2, which was on hold. It holds the application of student 2 and rejects student 4.

- Erie holds the application of student 6.

Hold	New Applicants		School		Hold	Reject
$\boxed{1}$ $\boxed{3}$		$\longrightarrow$	Clair	$\longrightarrow$	$\boxed{1}$ $\boxed{3}$	
□ □	6	$\longrightarrow$	Erie	$\longrightarrow$	$\boxed{6}$ □	
$\boxed{5}$		$\longrightarrow$	Huron	$\longrightarrow$	$\boxed{5}$	
$\boxed{2}$	4	$\longrightarrow$	Ontario	$\longrightarrow$	$\boxed{2}$	4

Step 5: Each student rejected in Step 4 applies to his/her next choice: student 4 applies to Erie.

- Erie considers the application of student 4 together with the application of student 6, which was on hold. It holds both applications.

Hold	New Applicants		School		Hold	Reject
$\boxed{1}$ $\boxed{3}$		$\longrightarrow$	Clair	$\longrightarrow$	$\boxed{1}$ $\boxed{3}$	
$\boxed{6}$ □	4	$\longrightarrow$	Erie	$\longrightarrow$	$\boxed{6}$ $\boxed{4}$	
$\boxed{5}$		$\longrightarrow$	Huron	$\longrightarrow$	$\boxed{5}$	
$\boxed{2}$		$\longrightarrow$	Ontario	$\longrightarrow$	$\boxed{2}$	

No application is rejected at Step 5. Based on this method the final allocations are:

Student	1	2	3	4	5	6
School	Clair	Ontario	Clair	Erie	Huron	Erie

You will have 15 minutes to go over the instructions at your own pace, and make your decisions. Feel free to earn as much cash as you can. Are there any questions?

Decision Sheet

- Recall: You are participant 1 and you live within the school district of School A.
- Recall: **Your payoff** amount depends on the school slot you hold at the end of the experiment. Payoff amounts are outlined in the following table.

Slot received at School:	A	B	C	D	E	F	G
Payoff to Participant 1 (in dollars)	16	19	12	5	8	14	10

The table is explained as follows:

- You will be paid \$16 if you hold a slot at school A at the end of the experiment.
- You will be paid \$19 if you hold a slot at school B at the end of the experiment.
- You will be paid \$12 if you hold a slot at school C at the end of the experiment.
- You will be paid \$5 if you hold a slot at school D at the end of the experiment.
- You will be paid \$8 if you hold a slot at school E at the end of the experiment.
- You will be paid \$14 if you hold a slot at school F at the end of the experiment.
- You will be paid \$10 if you hold a slot at school G at the end of the experiment.

Please write down your ranking of the schools (A through G) from your first choice to your last choice. Please rank ALL seven schools.

1st Choice	2nd Choice	3rd Choice	4th Choice	5th Choice	6th Choice	7th Choice

Please remain seated until the experimenter collects your Decision Sheet.

After the experimenter collects all Decision Sheets, the experimenter will draw ping pong balls from an urn to generate a fair lottery. The lottery, as well as all participants' rankings will be entered into a computer after the experiment. The experimenter will inform each participant of his/her allocation slot and respective payoff once it is computed.

O IA Instructions

Instructions

This is an experiment in the economics of decision making. The instructions are simple, and if you follow them carefully and make good decisions, you might earn a considerable amount of money. In this experiment, we simulate a procedure to allocate students to schools. The procedure, payment rules, and student allocation method are described below. Do not communicate with each other during the experiment. If you have questions at any point during the experiment, raise your hand and the experimenter will help you.

Procedure

- There are 36 participants in this experiment. You are participant 1.
- In this simulation, 36 school slots are available across seven schools. These schools differ in size, geographic location, specialty, and quality of instruction in each specialty. Each school slot is allocated to one participant. There are three slots each at schools A and B, and six slots each at schools C, D, E, F and G.
- **Your payoff** amount depends on the school slot you hold at the end of the experiment. These amounts reflect the desirability of the school in terms of location, specialty and quality of instruction.

Slot received at School:	A	B	C	D	E	F	G
Payoff to Participant 1 (in dollars)	16	19	12	5	8	14	10

The table is explained as follows:

- You will be paid $16 if you hold a slot at school A at the end of the experiment.
- You will be paid $19 if you hold a slot at school B at the end of the experiment.
- You will be paid $12 if you hold a slot at school C at the end of the experiment.
- You will be paid $5 if you hold a slot at school D at the end of the experiment.
- You will be paid $8 if you hold a slot at school E at the end of the experiment.
- You will be paid $14 if you hold a slot at school F at the end of the experiment.

- You will be paid $10 if you hold a slot at school G at the end of the experiment.

***NOTE* different participants might have different payoff tables.** That is, payoff by school might be different for different participants.

- During the experiment, each participant first completes the Decision Sheet by indicating school preferences. The Decision Sheet is the last page of this packet. Note that you need to rank all seven schools in order to indicate your preferences.
- After all participants have indicated their preferences, the experimenter will collect the preferences and start the allocation process.
- Once the allocations are determined, the experimenter will inform each participant of his/her allocation slot and respective payoff.
- In this experiment, participants are defined as belonging to the following school districts.
  - Participants #1 - #3 live within the school district of school A,
  - Participants #4 - #6 live within the school district of school B,
  - Participants #7 - #12 live within the school district of school C,
  - Participants #13 - #18 live within the school district of school D,
  - Participants #19 - #24 live within the school district of school E,
  - Participants #25 - #30 live within the school district of school F,
  - Participants #31 - #36 live within the school district of school G,
- In addition, for each school, a separate priority order of the students is determined as follows:
  - **Highest Priority Level:** Participants who rank the school as their first choice AND who also live within the school district.
  - **2nd Priority Level:** Participants who rank the school as their first choice BUT who do not live within the school district.
  - **3rd Priority Level:** Participants who rank the school as their second choice AND who also live within the school district.

- **4th Priority Level:** Participants who rank the school as their second choice BUT who do not live within the school district.

  .

  .

  .

- **13th Priority Level:** Participants who rank the school as their seventh choice AND who also live within the school district.
- **Lowest Priority Level:** Participants who rank the school as their seventh choice BUT who do not live within the school district.

• The ties between participants at the same priority level are broken using a fair lottery. This means each participant has an equal chance of being the first in the line, the second in the line, ..., as well as the last in the line. To determine this lottery, the experimenter will draw participant's ID number randomly, one at a time. The sequence of the draw determines the order in the lottery.

• Therefore, to determine the priority order of a student for a school:

  - The first consideration is how highly the participant ranks the school when indicating his or her preferences,
  - The second consideration is whether the participant lives within the school district or not, and
  - The last consideration is the order in the lottery.

• Once the priorities are determined, slots are allocated in seven rounds.

Round 1. a. An application to the first ranked school in the Decision Sheet is sent for each participant.

b. Each school accepts the students with higher priority order until all slots are filled. These students and their assignments are removed from the system. The remaining applications for each respective school are rejected.

Round 2. a. The rejected applications are sent to the school he/she ranked second when indicating his/her preferences.

b. Each school accepts the students with higher priority order until all slots are filled. These students and their assignments are removed from the system. The remaining applications for each respective school are rejected.

⋮

Round 6. a. The application of each participant who is rejected by his/her top five choices is sent to his/her sixth choice.

b. If a school still has slots available, then it accepts the students with higher priority order until all slots are filled. The remaining applications are rejected.

Round 7. Each remaining participant is assigned a slot at his/her last choice.

An Example:
We will go through a simple example to illustrate how the allocation method works.

Students and Schools: In this example, there are six students, 1-6, and four schools, Clair Erie, Huron and Ontario.

Student ID Numbers: 1,2,3,4,5,6	Schools: Clair, Erie, Huron, Ontario

Slots and Residents: there are two slots at each Clair and Erie, and one slot each at Huron and Ontario. Residents of districts are indicated in the table below.

Schools	Slot 1	Slot 2	District Residents
Clair	□	□	1 2
Erie	□	□	3 4
Huron	□		5
Ontario	□		6

Lotteries: In this example, the lottery produced the following order

$\boxed{\text{1-2-3-4-5-6}}$

Submitted School Rankings: The students have submitted the following school rankings

	1st Choice	2nd Choice	3rd Choice	Last Choice
Student 1	Huron	Clair	Ontario	Erie
Student 2	Huron	Ontario	Clair	Erie
Student 3	Ontario	Clair	Erie	Huron
Student 4	Huron	Clair	Ontario	Erie
Student 5	Ontario	Huron	Clair	Erie
Student 6	Clair	Erie	Ontario	Huron

Priority: School priorities depend on: (1) how highly the student ranks the school, (2) whether the school is a district school, and (3) the lottery order.
Clair: Student 6 ranks Clair first. Students 1,3 and 4 rank Clair second; among them, student 1 lives within the Clair school district. Students 2 and 5 rank Clair third.
Using the lottery order to break ties, the priority order for Clair is 6,1,3,4,2,5.

1st Choice	2nd Choice		3rd Choice	4th Choice
6	1	3,4	2,5	None
	Resident	Non-Resident	Non-Resident	

Erie: Student 6 ranks Erie second. Student 3 ranks Erie third. Students 1, 2, 4 and 5 rank Erie fourth; among them student 4 lives within the Erie school district.
Using the lottery order to break ties, the priority for Erie is 6-3-4-1-2-5.

1st Choice	2nd Choice	3rd Choice	4th Choice	
None	6	3	4	1,2,5
			Resident	Non-Resident

Huron: Students 1, 2 and 4 rank Huron first. Student 5 ranks Huron second. Students 3 and 6 rank Huron fourth.

Using the lottery at Huron in order to break ties, the priority for Huron is 1-2-4-5-3-6.
Ontario: Students 3 and 5 rank Ontario first. Student 2 ranks Ontario second. Students 1, 4 and 6 rank Ontario third; among them student 6 lives within the Ontario school district.

1st Choice	2nd Choice	3rd Choice	4th Choice
1,2,4	5	None	3,6
Non-Resident			Non-Resident

1st Choice	2nd Choice	3rd Choice	
3,5	2	6	1,4
Non-Resident		Resident	Non-Resident

Using the lottery at Ontario order to break ties, the priority for Ontario is 3-5-2-6-1-4.

The allocation method consists of the following rounds.

Round 1: Each student applies to his/her first choice: students 1, 2 and 4 apply to Huron, students 3 and 5 apply to Ontario, and student 6 applies to Clair.

- Clair accepts student 6.
- Huron accepts student 1 and rejects students 2 and 4.
- Ontario accepts student 3 and rejects student 5.

Applicants		School		Accept	Reject	Slot 1	Slot 2
6	$\longrightarrow$	Clair	$\longrightarrow$	6		$\boxed{6}$	$\square$
	$\longrightarrow$	Erie	$\longrightarrow$			$\square$	$\square$
1,2,4	$\longrightarrow$	Huron	$\longrightarrow$	1	2,4	$\boxed{1}$	
3,5	$\longrightarrow$	Ontario	$\longrightarrow$	3	5	$\boxed{3}$	

Accepted students are removed from the subsequent process.

Round 2: Each student who is rejected in Round 1 then applies to his/her **second choice:** Student 2 applies to Ontario, student 4 applies to Clair, and student 5 applies to Huron.

- No slot is left at Ontario, so it rejects student 2.
- Clair accepts student 4 for its last slot.
- No slot is left at Huron, so it rejects student 5.

Round 3: Each student who is rejected in Round 1-2 then applies to his/her **third choice:** Students 2 and 5 apply to Clair.

Applicants		School		Accept	Reject	Slot 1	Slot 2
4	⟶	Clair	⟶	4		$\boxed{6}$	$\boxed{4}$
	⟶	Erie	⟶			□	□
5	⟶	Huron	⟶		5	$\boxed{1}$	
2	⟶	Ontario	⟶		2	$\boxed{3}$	

Applicants		School		Accept	Reject	Slot 1	Slot 2
2,5	⟶	Clair	⟶		2,5	$\boxed{6}$	$\boxed{4}$
	⟶	Erie	⟶			□	□
	⟶	Huron	⟶			$\boxed{1}$	
	⟶	Ontario	⟶			$\boxed{3}$	

- No slot is left at Clair, so it rejects student 2 and 5.

Round 4: Each remaining student is assigned a slot at his/her **last choice:**

- Student 2 and 5 receive a slot at Erie.

Applicants		School		Accept	Reject	Slot 1	Slot 2
	⟶	Clair	⟶			$\boxed{6}$	$\boxed{4}$
2,5	⟶	Erie	⟶	2,5		$\boxed{2}$	$\boxed{5}$
	⟶	Huron	⟶			$\boxed{1}$	
	⟶	Ontario	⟶			$\boxed{3}$	

on this method the final allocations are:

Student	1	2	3	4	5	6
School	Huron	Erie	Ontario	Clair	Erie	Clair

You will have 15 minutes to go over the instructions at your own pace, and make your decisions. Feel free to earn as much cash as you can. Are there any questions?

Decision Sheet

- Recall: You are participant 1 and you live within the school district of School A.
- Recall: **Your payoff** amount depends on the school slot you hold at the end of the experiment. Payoff amounts are outlined in the following table.

Slot received at School:	A	B	C	D	E	F	G
Payoff to Participant 1 (in dollars)	16	19	12	5	8	14	10

The table is explained as follows:

- You will be paid \$16 if you hold a slot at school A at the end of the experiment.
- You will be paid \$19 if you hold a slot at school B at the end of the experiment.
- You will be paid \$12 if you hold a slot at school C at the end of the experiment.
- You will be paid \$5 if you hold a slot at school D at the end of the experiment.
- You will be paid \$8 if you hold a slot at school E at the end of the experiment.
- You will be paid \$14 if you hold a slot at school F at the end of the experiment.
- You will be paid \$10 if you hold a slot at school G at the end of the experiment.

Please write down your ranking of the schools (A through G) from your first choice to your last choice. Please rank ALL seven schools.

1st Choice	2nd Choice	3rd Choice	4th Choice	5th Choice	6th Choice	7th Choice

Please remain seated until the experimenter collects your Decision Sheet.

After the experimenter collects all Decision Sheets, the experimenter will draw ping pong balls from an urn to generate a fair lottery. The lottery, as well as all participants' rankings will be entered into a computer after the experiment. The experimenter will inform each participant of his/her allocation slot and respective payoff once it is computed.

P Post-experiment Survey

1. Did you try to follow the advice on how to rank the schools that you were given?
    - Yes
    - No
    - Other (Please Specify)
2. How clear was the advice on how to rank the schools that you were given?
    - Very unclear
    - Unclear
    - Clear
    - Very clear
3. How good was the advice on how to rank the schools that you were given?
    - Very bad
    - Bad
    - Good
    - Very good
4. How did you decide how to rank the schools on the decision sheet?
5. How clear were the rules that would determine which school you would hold a slot at?
    - Very unclear
    - Unclear
    - Clear
    - Very clear
6. Was there anything about the rules that you did not understand?

7. Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?

    - Most people can be trusted
    - You can't be too careful

8. How many experiments in this lab or a similar lab have you participated in?

9. Did you have any prior experience with procedures in which students rank schools and are assigned a slot to a school, either in an experiment, in your own life, or in classes?

    - No
    - Yes, I participated in an experiment involving a similar problem
    - Yes, I have been exposed to a similar procedure in my own life
    - Yes, I have heard of these procedures in one of my classes
    - Other (Please Specify)

10. In the past, have you ever been lied to or deceived in some way in an experiment?

    - Yes
    - No
    - Other (Please Specify)

11. Do you have any additional remarks or comments about this experiment?

The next 4 questions refer to the below scenario. For each question you answer correctly, you will receive $0.50 for up to $2.00 total.

Four people H,J,K, and L have apartments in the same four-story building. Each person lives on a different floor, from the first floor up to the fourth floor. The following restrictions apply:

- Either H or K lives on the first floor.
- J lives on the floor directly below L.

1. Which of the following could be a list of the four people in order from the first floor up to the fourth floor?

   - H,K,L,J
   - H,J,K,L
   - J,L,K,H
   - K,H,J,L
   - K,J,H,L

2. Which of the following statements CANNOT be true?

   - H lives on the first floor.
   - H lives on the second floor.
   - J lives on the second floor.
   - K lives on the third floor.
   - L lives on the fourth floor.

3. If K lives on the second floor, all of the following statements are true EXCEPT:

   - H does not live on the third floor
   - H does not live on the fourth floor
   - J does not live on the third floor
   - J does not live on the first floor
   - L does not live on the third floor

4. If L lives on the third floor, which of the following statements must be true?

   - H does not live on the first floor.
   - J does not live on the first floor.
   - J does not live on the second floor.
   - K does not live on the first floor.
   - K does not live on the fourth floor.

Demographics

1. Gender
   - Male
   - Female
   - Other (Please Specify)
2. Ethnicity
   - American Indian
   - Hispanic or Latino
   - Asian
   - Black or African-American
   - Non-hispanic White
   - Pacific Islander
   - Multiple or Mixed
   - Other (Please Specify)
3. Age
4. Marital Status
   - Never Married (Single)
   - Married
   - Living Together, Partners
   - Separated
   - Divorced
   - Widowed

5. Number of children

6. Age of youngest child (if any)

7. Age of oldest child (if any)

8. Student Status

   - Full-time
   - Part-time
   - Other (Please Specify)

Q DA Advice

Advice on how to rank the schools

Whatever ranking other participants report, you will obtain the highest possible payoff by reporting first the school for which you have the highest payoff, second the school for which you have the second highest payoff, and so on. In other words, you should rank the schools in the order of your payoffs, from high to low.

This follows from the fact that your priority at the different schools does not depend on the ranking of schools that you report. Whatever ranking you report, your priority at the different schools remains the same.

If you are rejected from the first school you apply to, the allocation method always allows you to keep your priority at your later schools.

Consider for instance student 2 in the example we gave you. In Step 1, student 2 applies to school Huron and is rejected because student 1 also applies to Huron, and student 1 has a higher priority than student 2 at Huron.

In Step 2, student 2 applies to school Ontario.

Because student 2 has a higher priority at school Ontario than student 3, student 3 is rejected and student 2 is retained at Ontario, even though student 3 had applied to Huron in a prior step and been retained. There was no loss to student 2 from ranking Huron higher than Ontario.

R IA Advice

Advice on how to rank the schools

Be careful about the ranking you report: you will not necessarily obtain the highest possible payoff by reporting first the school for which you have the highest payoff, second the school for which you have the second highest payoff, and so on. In other words, it is not necessarily best for you to rank the schools in the order of your payoffs, from high to low.

One possible strategy is to rank first the school within the district of which you live.

A second strategy is to rank first the school for which you have the highest payoff, and rank the school within the district of which you live second.

The second strategy is riskier, but offers you a higher chance of holding a seat at the school for which you have the highest payoff.

The reason it is not necessarily best to rank schools in the order of your payoff is that your priority at different schools depends on the ranking of schools that you report.

For example, at any school you do not rank first, your priority at that school is lower than the priority of any other participant who would rank that school first. This is true even if you live within the school district of that school and the other participant does not.

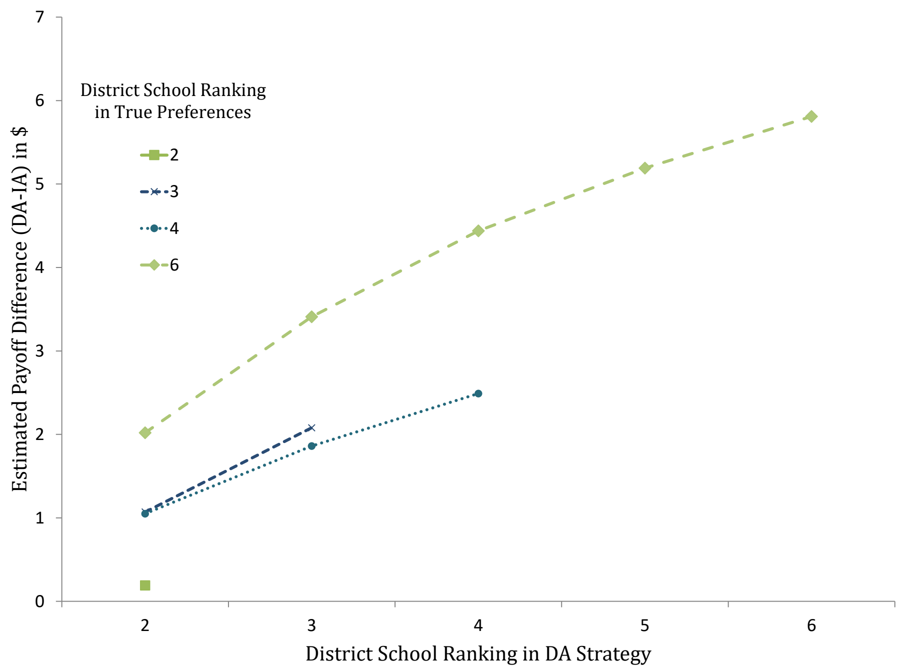
Consider for example student 2 in the example we gave you. In Step 1, student 2 applies to school Huron and is rejected because student 1 also applies to Huron and has a higher priority than student 2 at Huron.

In Step 2, student 2 applies to school Ontario.

Even though student 2 has a higher lottery at school Ontario than student 3, student 2 is rejected from Ontario and student 3 remains at Ontario because student 3 applied to Ontario in an earlier round than student 2.

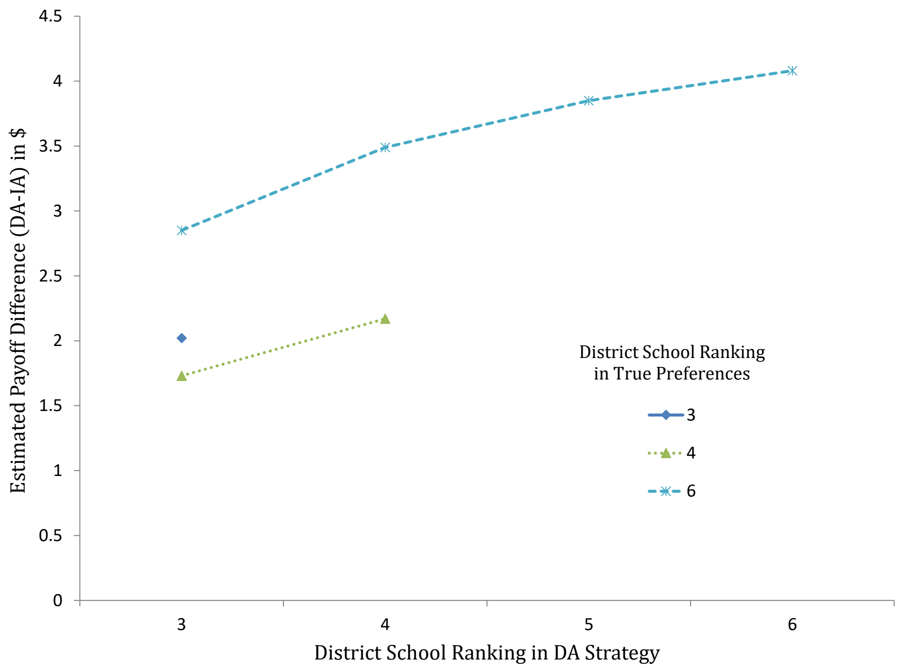
S Individual Welfare Results for Participants in No Advice Sessions

Figure 3.15: Expected Payoff Difference between District-Consistent Ordered Strategy in DA and the Safe Strategy in IA



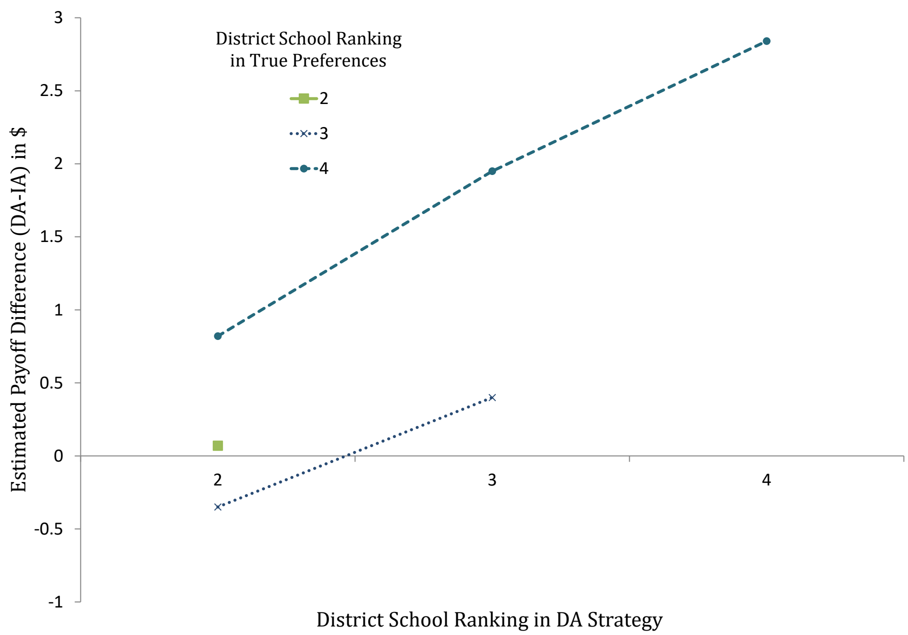
Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose the Safe strategy in an IA session of playing a District-Consistent Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

Figure 3.16: Expected Payoff Difference between District-Consistent Not-Ordered Strategy in DA and the Safe Strategy in IA



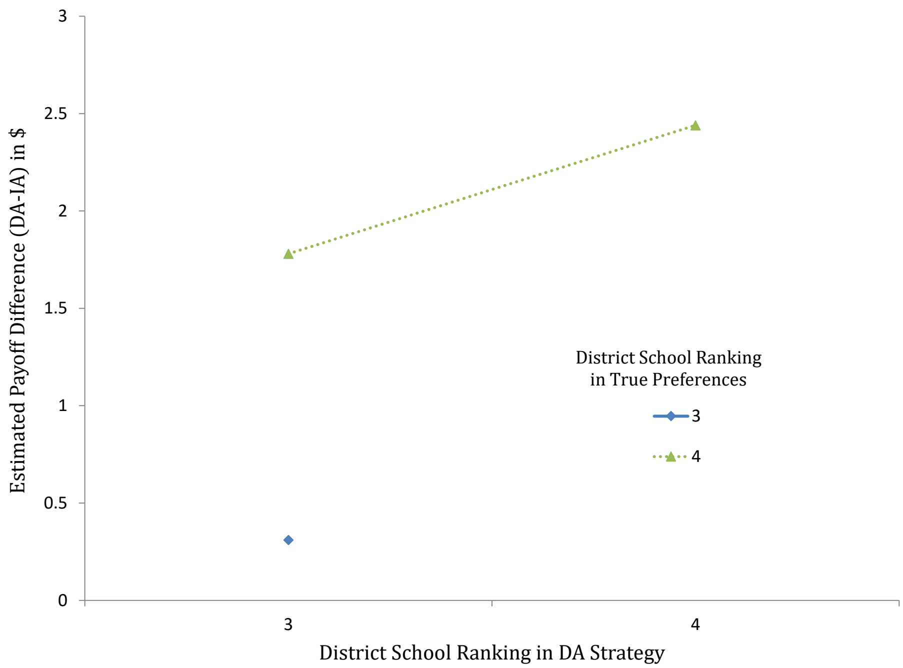
Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose the Safe strategy in an IA session of playing a District-Consistent Not-Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

Figure 3.17: Expected Payoff Difference between District-Consistent Ordered Strategy in DA and the Risky Strategy in IA



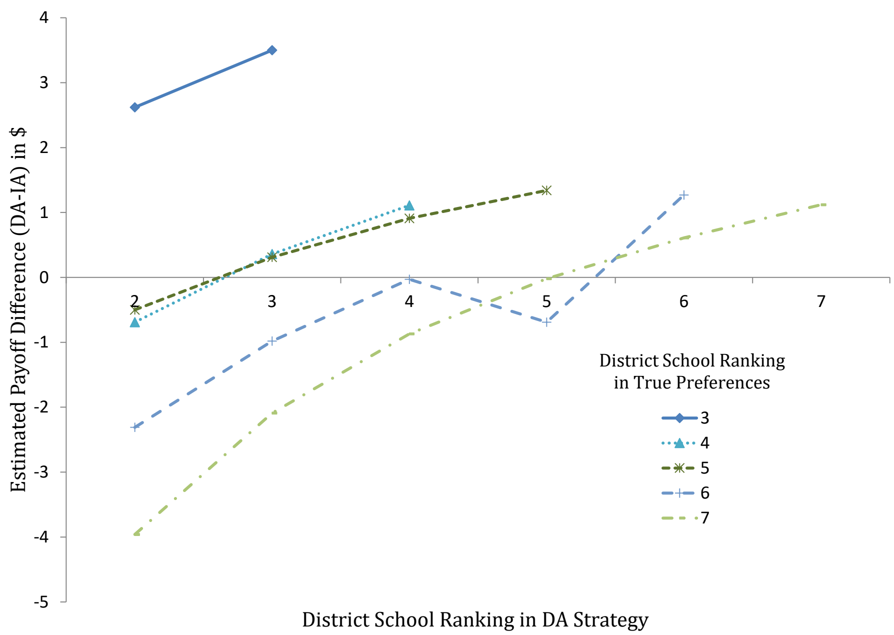
Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose the Risky strategy in an IA session of playing a District-Consistent Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

Figure 3.18: Expected Payoff Difference between District-Consistent Not-Ordered Strategy in DA and the Risky Strategy in IA



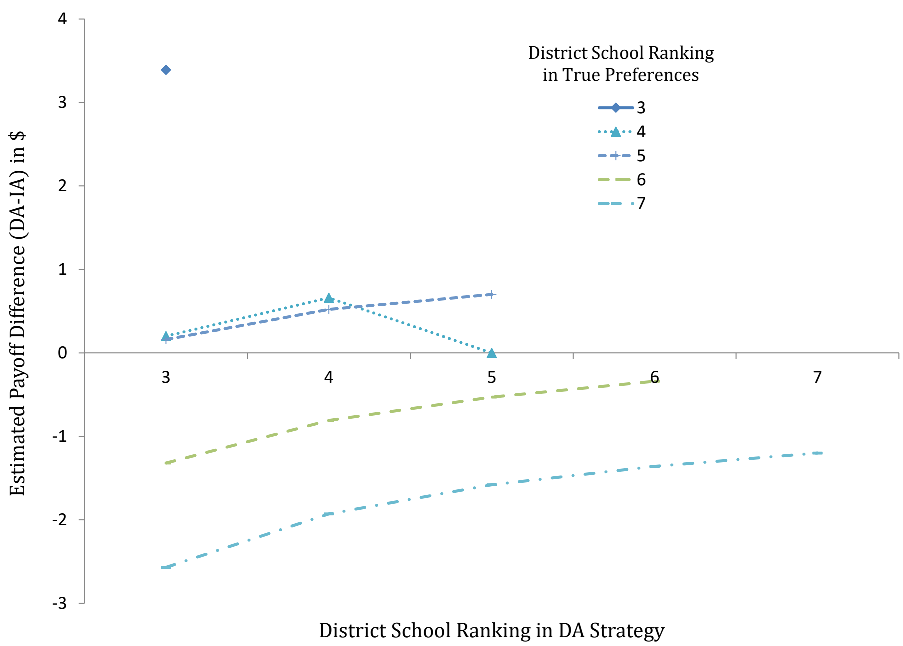
Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose the Risky strategy in an IA session of playing a District-Consistent Not-Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

Figure 3.19: Expected Payoff Difference between District-Consistent Ordered Strategy in DA and Other Strategies in IA



Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose neither the Risky strategy nor the Safe strategy in an IA session of playing a District-Consistent Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.

Figure 3.20: Expected Payoff Difference between District-Consistent Not-Ordered Strategy in DA and Other Strategies in IA



Notes: Each line traces the estimated payoff difference to participants with the same district school rank who chose neither the Risky strategy nor the Safe strategy in an IA session of playing a District-Consistent Not-Ordered x-Strategy in DA. The x-axis tracks the rank of the district school in the counter-factual strategy.