Essays on Discrimination and Wage Inequality

by

Merve Sarıışık

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2022

Doctoral Committee:

Professor Tanya Rosenblat, Chair
Professor Stephen Leider
Associate Professor David Miller
Assistant Professor Ana Reynoso

Merve Sarıışık

merve@umich.edu

ORCID iD: 0000-0003-0932-1516

To my parents, H. Asuman Köksal and Birsu Sarıışık,
my brother, Mert,

and to my grandparents Aysel Köksal and Kamil Rüştü Köksal,

who all have been incredibly patient with me
and provided me with endless support.

ACKNOWLEDGEMENTS

Although my name is on the front page, this work is the result of contributions from many individuals. Firstly, I am grateful to my parents and my brother, who I love endlessly, for their continuous support. Their countless efforts allowed me the privilege to pursue this Ph.D. I am also thankful to have such incredible and loving grandparents, whose vision is at the root of most my life milestones; this one is not an exception.

I would like to thank my chair, Tanya Rosenblat, whose support, trust in me, and guidance has played a fundamental role in my time at the University of Michigan since day one. I have been lucky to have her as a "mentor" in many aspects both academically and otherwise. I would like to also thank my advisors, David Miller, Stephen Leider, and Ana Reynoso, for their continued guidance and for being incredible resources to me. I also received valuable feedback over the years from BEE Lab participants Alain Cohn, Erin Krupka, Yan Chen, Yesim Orhun, Mohamed Abbadi, Linfeng Li, Jaron Cordero, Samer Charbaji, Iman Yeckehzaare, Daphne Chang, Evgeny Kagan, Carrie Wenjin Xu, Ark Fangzhou Zhang. Laura Flak, Lauren Pulay, Julie Heintz, Cassie Kelly, and Hiba Baghdadi, from the Department of Economics, made every possible resource available for my success. I also gratefully acknowledge funding from the Graduate Research Fund of the Department of Economics for my projects, SYLFF, and the Rackham Graduate Student Research Grants that allowed me to run my experiments.

The friends I have and the new ones I made during Ph.D. also played a tremendous role during this process: Thomas Brosy and Chiara Ferrero, who kept our office alive through countless Zoom study sessions during COVID-19 pandemic and who now feel like a brother and a sister to me; Yeliz Kaçamak who became my safe haven; Doğan Sümer with whom we have built a mutually supportive friendship together during the most difficult times; Emir Murathanoğlu who I had countless walks, talks and the backgammon sessions with; Katie Gaches and John Springer who have become my chosen family, my roots, my foundation in Detroit and in the U.S.; Nick Zinis, Loren Wolfe, and Chandon Waller who made my life more joyful and the finish line possible; Baran Akpinar and Emma Heutschi who tirelessly listened to me during the hardest times during my PhD experience and are incredible friends; Barış Can Aktepe, Ecem Ezgi Coşkun, Aycan Katıtaş, Aysegül Eriç, Yonca Kumsar, Güngör Oduncu, Özgecan Kalaycı, Deniz Bayram, Neslihan İzmirli, İdil Ayata, and Nihan Aler who are my international web of support and who I kept in touch with over Whatsapp conversations.

Last but definitely not least, I would like to thank Linda Tesar, Christopher House, and their families for embracing me, providing me with incredible support, and treating me like family.

I feel very lucky and I am forever grateful for all the people who have made this possible.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

This dissertation analyzes discrimination and wage inequality in different contexts - interactions over an online peer to peer platform and discretionary disclosure of salary history under a new policy that is intended to alleviate wage inequality. We study these using online field and online lab experiments as well as collecting data through web scraping.

The first chapter studies communication among applicants and prospective employers where salary history disclosure is discretionary. Several states and cities have recently adopted Salary History Bans (SHB) that take away to right to ask about applicants' salary histories from prospective employers. The goal of SHB is to alleviate race and gender wage gap. However, given the productivity signaling value of past salary, it is natural to expect voluntary disclosure by applicants. How will SHB impact employers' inference under wage gap? We use signals that are a biased representation of uniformly distributed types to represent "past salary" in our virtual lab setting. We mimic wage inequality through two different sender groups with disproportionate endowment of either lower or higher signals, referred to as "Disadvantaged" and "Advantaged" groups respectively. We use a theoretical model to reflect employer naivete in interpreting past salary information or the lack thereof. Both theory and experimental data suggest that employers under(over)estimate types when disclosed signals are low (high) for the Disadvantaged (Advantaged) group. Further, employers overestimate types when signals are not disclosed. Our experiment reveals that the under(over)estimation of disclosed signals is of greater magnitude under voluntary disclosure. In addition, the receiver naivete induced mistakes are of greater magnitude for disclosed signals than nondisclosed signals. Our results suggest that SHB can instead worsen the existing wage discrepancies.

The second chapter is an audit study investigating discrimination against Arab/Muslim identity over Airbnb, a classic example of online marketplace that allows users to post or rent short-term housing in residential premises. We demonstrate an important consequence of Airbnb's current market design. We run an online audit study over Airbnb and display nine percentage point gap between the acceptance rate of requests from guests with distinctively Arab/Muslim identities and those of distinctively non-Arab/Muslim identities. The difference persists across different specifications.

CHAPTER I

Voluntary Disclosure under Salary History Ban: A Model and an Experimental Investigation

Abstract

Several states and cities have recently adopted Salary History Bans (SHB) that take away to right to ask about applicants' salary histories from prospective employers. The goal of SHB is to alleviate race and gender wage gap. However, given the productivity signaling value of past salary, it is natural to expect voluntary disclosure by applicants. How will SHB impact employers' inference under wage gap? We use signals that are a biased representation of uniformly distributed types to represent "past salary" in our virtual lab setting. We mimic wage inequality through two different sender groups with disproportionate endowment of either lower or higher signals, referred to as "Disadvantaged" and "Advantaged" groups respectively. We use a theoretical model to reflect employer naivete in interpreting past salary information or the lack thereof. Both theory and experimental data suggest that employers under(over)estimate types when disclosed signals are low (high) for the Disadvantaged (Advantaged) group. Further, employers overestimate types when signals are not disclosed. Our experiment reveals that the under(over)estimation of disclosed signals is of greater magnitude under voluntary disclosure. In addition, the receiver naivete induced mistakes are of greater magnitude for disclosed signals than nondisclosed signals. Our results suggest that SHB can instead worsen the existing wage discrepancies.

JEL Classification: D82, D83, C92, J38, J71
Keywords: communication, naivete, regulation, beliefs

1.1 Introduction

Half of US workers said their present employer knew their prior wage, according to Hall and Krueger (2012). Additionally, more than 80% of US workers said that if their employer knew their prior wage, they accessed this information prior to making a job offer. Since January of 2017, several states and local jurisdictions have been increasingly passing Salary History Bans (SHB hereafter), that blind prospective employers to applicants' previous salary information. Under SHB, inquiring about applicants salary history is prohibited. Bans stem from concerns for history dependence in wages for workers who were historically hurt due to entry-level mismatches, career interruptions, or discrimination. The goal of SHB is to alleviate wage inequality by disrupting that cycle.

Given the rapid spread of the SHB laws across the nation,[1] it is essential to determine its impact moving forward. Note that applicants can still voluntarily disclose salary information under SHB. Given its close relationship with marginal productivity, the decision to whether or not disclose salary is a strategic one that can alter prospective employers' inference about the applicant. Although voluntary disclosure has so far been primarily absent from the policy discussions, a recent worker compliance survey provides evidence for this kind of behavior (Agan, Cowgill and Gee, 2020)[2].

Considering salary to be a biased signal for employee productivity, sophisticated employers understand and can correct for the bias as well as the strategic disclosure decision. On the other hand, naive employers' inference is subject to systematic misperceptions. We present a theory-founded experiment to investigate the shift in nature of communication between prospective employers of heterogenous strategic sophistication adn applicants, under SHB. Carefully designed experiments that carry the fundamentals of a real-life strategic interaction into the lab have the potential to provide us with the essential intuition that would otherwise not be possible. We seek to answer the following questions: Will Salary History Ban worsen information transmission? Will Salary History Ban help or hurt applicants? Specifically, under wage inequality across two (demographic) groups, which group of applicants will by the transition into Salary History Ban?

The Model. We represent the salary history related interaction among applicants and prospective employers as a game of information transmission. Consider a a sender (applicant)

[1]As of now, the running list has 21 states and 21 localities that have enacted the ban at different levels (public, private or both).

[2]Agan, Cowgill and Gee (2020) find that employees who are paid well are very willing to disclose, even if they are the sole employee to do so. Low pay makes employees significantly less inclined to report anything at all. However, if other candidates in the applicant pool have declared, these employees are more likely to do the same.

who is endowed with a biased signal (salary) of her type (productivity). The payoff for the sender (prospective employer) increases in the receiver's belief about the sender's type, but the payoff for the receiver (prospective employer) increases with the accuracy of his belief. We distinguish between two information transmission regimes: in the mandatory disclosure regime (pre-SHB) senders always disclose the signal; in the voluntary disclosure regime (post-SHB), senders can withhold the signal or chose to disclose it voluntarily.

If all agents are rational, all information should be revealed in equilibrium (Viscusi, 1978; Grossman, 1981; Grossman and Hart, 1980; Milgrom, 1981)[3]. However, in reality, many receivers may be naive and make systematic mistakes in how they interpret information that requires debiasing and strategic skepticism. Building on Eyster and Rabin (2005) and Hagenbach and Koessler (2017), our model incorporates both naive and sophisticated receivers. Naive receivers consider all types in the type support of a given sender action to be equally likely. As will be explained in detail over the next section, naivete in our model corresponds to disregarding strategic sender motives and the conditional type distribution of any sender actions.

We incorporate two groups of senders into our model to mimic real world demographic-group-favoring wage inequality. Although both groups have the same underlying uniform type distribution, their signal distributions differ due to the structure of the bias. The mass of the signal distribution for the "disadvantaged" sender group lies on the lower range due to negative bias. In contrast, the mass of the signal distribution for the "advantaged" sender group lies on the higher range due to positive bias. When signals are disclosed, the existence of naive receivers leads to underestimation (overestimation) of the types on average for the former (latter) group. Further, naive receivers consider senders of each type equally likely to withhold information which leads to an overestimation of types on average upon nondisclosure.

Moving from mandatory disclosure to voluntary disclosure, both sender groups follow a threshold strategy; disclose if signal is greater than the threshold, do not disclose otherwise. The advantaged group has a higher threshold yet is expected to benefit less from the transition to voluntary disclosure.

The Experiment. In light of the model and the research questions, the experiment consists of three waves, a total of five treatment arms. In all treatments and in each round, senders are assigned to two numbers: one indicating their type and another indicating the error term (bias). The signal of each sender is then equal to the summation of their assigned type and

[3]An established result in information economics is that in the (perfect Bayesian) equilibrium, there will be full disclosure, as long as disclosed information is verifiable and cheap to disclose.

the error term each round. The types are uniformly distributed over the set $\{5, 6, 7, 8, 9, 10\}$ and the error terms are distributed over {-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5}. The distribution of error terms depends on the sender group. The disadvantaged sender group has a right skewed error distribution, such that the mass is on the lowest error term {-5}. On the other hand, the left skewed error distribution of the advantaged sender group has mass on the highest error term {5}. Given the sets of types and error terms, the signals range from integers 0 to 15. For example, a sender with signal 2 could have either of the following (type, error) combinations, (5, -3) (6, -4), (7, -5). Note that the disadvantaged group's signal distribution is concentrated on the lower signals (right skewed) and the advantaged group has disproportionate access to higher signals (left skewed).

The goal of the first two waves is to test the predictions of the model that will be explained further in the next section. Specifically, Wave 1 focuses on biased inference of disclosed signals and comprises of two treatments of mandatory disclosure game, one for the disadvantaged and one for the advantaged sender group, 'Treatment MD' and 'Treatment MA, respectively.'[4] Wave 2 tests the predictions of the model under voluntary disclosure for the advantaged sender group, 'Treatment VA'. [5] Finally, Wave 3 investigates welfare implications of SHB under preexisting wage gap across two demographic groups. Wave 3 has two treatments, one for the mandatory and one for the voluntary disclosure regime, each involving both sender groups ('Treatment M' and 'Treatment V'). By including both sender groups, these treatments channel a more realistic scenario where employers interact with applicants from different "demographic groups".

The theoretical predictions are largely validated in the experimental data. Receivers on average underestimate (overestimate) the types for the disadvantaged (advantaged) sender group. We observe less than full revelation in voluntary disclosure treatments. Specifically, disclosure rates of senders significantly drop for the lowest signals, reflecting a threshold equilibrium with higher threshold for the advantaged sender group. Further, receivers on average overestimate the sender type upon nondisclosure. Interestingly, experiment results indicate that the receiver naivete is greater for disclosed signals under voluntary disclosure than under mandatory disclosure. Further, we observe that the receiver bias for nondislosure is of smaller magnitude than for disclosed signals under voluntary disclosure. Consequently, we observe that the voluntary disclosure regime helps the advantaged group increase their

[4]Treatment names are comprised of initials of the identifier information transmission regime and the sender group when applicable, e.g. 'M' for Mandatory and 'D' for Disadvantaged.

[5]We expect to observe threshold sender strategy, i.e. only senders with a signal below a threshold will choose not to disclose. We choose advantaged sender group for cleaner identification: the set of not disclosed signals does not coincide with the set of messages that let us validate model's predictions for the disclosed signals.

payoffs more than the disadvantaged group.

Given the key assumption, receiver naivete, we further examine two additional questions about our model: whether other measures of cognitive ability correlate with the receiver mistakes, and whether other behavioral factors influence mistakes. After playing the sender-receiver game, we elicited subject beliefs pertaining to the behavior of the players in the opposite role, asked them to complete a risk aversion task and answer three cognitive reflection test (CRT) questions. Our results show that the elicited beliefs are consistent with the behavior of subjects in the sender-receiver game, which also facilitated a rationality check for senders. We also see that the performance in CRT is a good predictor of receiver mistakes [Note: didn't have time to include this result in the draft.]. Moreover, the risk preferences do not appear to be correlated with receivers mistakes.

Policy implications. The experimental (virtual) laboratory has significant advantages from policy implication perspective. It lets us exogenously alter the disclosure regime, unlike field settings.[6] Furthermore, our dataset allows us to directly observe receiver beliefs, which is necessary for complete investigation of the receivers' inference problem. As a result, our results can shed light on how salary history bans are likely to affect information transmission across applicants and prospective employers.

The intent of SHB is to shrink the wage gap by preventing inequality in pay from the very beginning or at least to avoid it from snowballing throughout decades long careers. By bridging the gap between labor policies, and behavioral economics, we show that these bans could instead widen wage inequality.

Limiting information transmission comes with the risk of sometimes worsening the problem at hand, while possibly increasing the search costs in the labor market.[7] Policy makers could instead focus on increasing transparency around pay. As Cullen and Pakzad-Hurson (2019) display, although in equilibrium, pay transparency leads to decreased wages overall[8], it could be a remedy to unequal pay. Historically disadvantaged employees would have the chance to do their research and hold a conversation around the reported salary range of the new workplace.

[6]Due to COVID-19 related limitations the experiment was conducted over Zoom with participants that were recruited from Prolific.co. We used the Zoom meeting as our virtual lab by following the regular lab protocols as closely as possible.

[7]Salary History Bans can lead to high search costs if, for example, employers interview applicants whose current salary is outside of their pay range. This could lead to lengthy interview processes followed by an offer only to learn that the prospective employer cannot afford the applicant. This would be a costly situation not only for the employer but also for the applicant.

[8]In situations where workers initially have some degree of bargaining power, pay transparency decreases that leverage.

Related Literature. A large and mature literature documents and investigates the causes of the income gaps across genders and races[9]

Our paper is related to prior literature about removing information from the hiring process—for example, banning employers from seeking candidates' credit scores, criminal records, or drug-test results (Bartik and Nelson, 2016; Ballance, Clifford and Shoag, 2020; Agan and Starr, 2018; Doleac and Hansen, 2020; Wozniak, 2015) or gender-blinding resumes (Åslund and Skans, 2012; Behaghel, Crépon and Le Barbanchon, 2015). Literature shows that some of these well intended policies may actually have the reverse impact on the labor market. Two of the most surprising results come from applications of banning employers from requesting credit score and criminal history records. Research shows that (Bartik and Nelson, 2016; Doleac and Hansen, 2020) instead of alleviating the statistical discrimination, these applications resulted in increased statistical discrimination against the demographic they hoped to help.

Given its productivity signaling value, supply side can easily override salary history bans through voluntary disclosure. This introduces the potential for unraveling (Viscusi, 1978; Grossman, 1981; Grossman and Hart, 1980; Milgrom, 1981). In many models, lack of disclosure may be viewed as a negative signal of quality, leading to full revelation. Empirically such unraveling is not always observed (Dranove and Jin, 2010; Mathios, 2000). Further, there is an experimental literature that studies information disclosure in the laboratory. The results point towards receiver naivete as the reason behind partial unraveling. (Deversi, Ispano and Schwardmann, 2018; Hagenbach, Koessler and Perez-Richet, 2014; Jin, Luca and Martin, 2015) The studies primarily focus on simple and rich language, in that, in addition to the option of nondisclosure, senders can either choose to share their true type or a set that includes their true type. They document that incomplete unraveling is due to senders with low types-signals, who anticipate and take advantage of receiver naivete.

Finally, a growing literature directly studies salary history bans. The theoretical predictions suggest that salary history ban can lead to adverse selection as job seekers with higher salary histories would lose their ability to signal this information (Meli and Spindler, 2019). This would be troublesome from a gender wage gap perspective as high-performing women are most likely to be underpaid relative to men.[10] Early empirical evidence supports this prediction; the positive short-term impact of reducing gender and racial wage gaps due to

[9]See Thilmany (2006); Solnick and Schweitzer (1999); Goldin (1990); Blau and Kahn (2020); Eckel and Grossman (2001); Juhn and McCue (2017); Blackaby, Booth and Frank (2005); Mazei et al. (2015); Moss-Racusin et al. (2012); Ponthieux and Meurs (2015); Rozada, Yeyati et al. (2018). Within this larger body, our research is primarily related to three strands of prior research.

[10]Goldin (2015) shows that MBA graduate women make 95% of what men do upon graduation, later declining to 57%-62% in the 12-15 years upon graduation. Further, more broadly, Blau and Kahn (2020) shows that gender gap has decreased more slowly at the top of the wage distribution.

SHB appears limited to employees with lower wages (Sinha, 2019; Hansen and McNichols, 2020; Bessen, Meng and Denk, 2020).[11] A few other researchers have examined the effect of salary disclosures and salary history bans using experiments in online markets (Barach and Horton, 2021) and in laboratory settings (Khanna, 2020).

This paper contributes to the salary history ban literature by developing a behavioral economic model that builds on the systematic documentation of partial unraveling due to lack of skepticism on the receiver side (Deversi, Ispano and Schwardmann, 2018; Hagenbach, Koessler and Perez-Richet, 2014; Jin, Luca and Martin, 2015). Our model's equilibrium results shed light on the expected applicant behavior and its consequences in the longer term.

Roadmap. In the next section, we present the model and the predictions. Section 3 then describes the experimental design and section 4 follows with the results. Finally section 5 concludes with a brief discussion of policy implications.

1.2 The Model

1.2.1 Setup

In order to study the impact of voluntary disclosure of biased signals, our design incorporates (i) control over productivity types, (ii) control over bias structure, (iii) common knowledge of the type and bias structure, (iv) identification of receiver misinference in the absence of voluntary disclosure.

A sender (S) and a receiver (R) engage in a game of information transmission in which S seeks to increase R's estimate while R seeks to make the most correct guess he can.

A sender (S) and a receiver (R) play an information transmission game in which S privately observes the state of the nature, t, and R's guess (g) of t determines the payoffs. S benefits from higher guesses, whereas R prefers guesses to be as close to the true state as possible. Payoffs are given by the following functions: $U_S = g$, $U_R = -(t-g)^2$. Note that U_R implies that the optimal guess is equal to the expectation of the state. The game begins by S privately observing the state t and an idiosyncratic error term e. The payoff relevant state t is drawn from a uniform distribution over the set $\{5, 6, 7, 8, 9, 10\}$ and the error term is drawn from a skewed distribution over the set {-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5}. Together, t and e add up to a signal $s = t + e$ and the set of signals is therefore $\Psi = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$. (See table 1.1)

Before R makes a guess, S sends a message m. Note that S cannot make false statements.

[11] A long list of states and localities have been enacting salary history bans on a rolling basis since 2017. As of this writing, there are 21 states and 21 localities with SHB.

Building over Hagenbach and Koessler (2017), we consider two communication regimes: *mandatory* and *voluntary* disclosure regimes. Under mandatory disclosure regime, the set of available messages to a sender with signal s is $\{s\}$. That is, S always discloses s to R. Under voluntary disclosure regime, the set of available messages expands to $\{s, \Psi\}$, i.e. S can either disclose her signal ($m = s$) or not disclose ($m = \Psi$). This representation of nondisclosure is merely for ease of exposition. It implies that the message S sends has no information pertaining to the true type.

A crucial component of our model is the distribution of error terms. There are two different sender groups, distinguished by their error distributions. One of the groups has a right skewed error distribution which means the mass is concentrated on the lower end of the set of errors, which we call as the 'disadvantaged group'. The other group, which we call as the 'advantaged group', has a left skewed error distribution, i.e. mass is on the higher end of the set of errors. Given the linear relationship between error terms and the signals, the skewness leads to disproportionate access to lower (higher) signals for the disadvantaged (advantaged) sender group. That is, we use skewed error distributions to synthetically create inequality in "salary history" across two "demographic groups of applicants". This way, we are statistically representing the pay gap through mean/median comparison, as well as through first order stochastic dominance relationship. Specifically, the skewness of the error distribution for the disadvantaged sender group can be formalized as follows $P_{\text{D}}(e = \text{-}5) = 0.5$ and $P_{\text{D}}(e = e') = \frac{0.5}{|\Psi|-1}$ for $e' \in \Psi \setminus \{\text{ -}5\}$. Similarly, the skewed error distribution of the advantaged group is then $P_{\text{A}}(e = 5) = 0.5$ and $P_{\text{A}}(e = e') = \frac{0.5}{|\Psi|-1}$ for $e' \in \Psi \setminus \{5\}$.

Main component of our analysis is R's lack of sophistication. When $R$is fully naive, he considers each type within a mesaage's type support to have equal likelihood. This implies the naivete can be in one of the two forms. Upon $m = s$, statistical naivete, in which R disregards the skewed error distribution and upon $m = \Psi$, strategic naivete, in which R disregards the strategic motive behind nondisclosure.[12] Following seminal papers in the literature (Deversi, Ispano and Schwardmann, 2018; Hagenbach and Koessler, 2017; Milgrom and Roberts, 1986), we model R's sophistication bimodally. Specifically, we assume that $\chi \in (0, 1)$ is the probability that R is fully naive and $(1 - \chi)$ is the probability that R is fully sophisticated.

1.2.2 Empirical Predictions from the Model

In this section we present the empirical predictions we are testing with the experimental data. Proofs are presented in Appendix A.1.1.

[12] Strategic naivete is in the spirit of "cursed equilibrium" of Eyster and Rabin (2005).

For the disclosed signals, naive R chooses actions as if the distribution of the error term was uniform over the type support of the disclosed signal. This leads $g_{\chi}(s)$ and $g(s)$ to differ when the signal is exactly less (greater) than a type by 5 for the disadvantaged (advantaged) group due to the skewness in the error distributions. For example, $g_{\chi}(2)$ is equal to 6 for signal 2 regardless of the group S belongs to. However, for the disadvantaged sender group, $g(2)$ or the correct posterior upon disclosed signal 2 is equal to

$$\begin{aligned}\mathbb{E}\big(t|m=2\big) &= \sum_{t\in T} t\frac{P\big(m=2|t\big)P\big(t\big)}{P\big(m=2\big)} \\ &= 5\frac{P\big(e=-3\big)P\big(t=5\big)}{P(m=2)} + 6\frac{P\big(e=-4\big)P\big(t=6\big)}{P(m=2)} \\ &\qquad + 7\frac{P\big(e=-5\big)P\big(t=7\big)}{P(m=2)} = 6.75.\end{aligned}$$

Prediction 1 *Under mandatory disclosure regime, R's expected guess is below the expected type for the disadvantaged sender group. Specifically,*

1. *R's expected guess is less than the expected type for the signal region* $\{1, 2, 3, 4, 5\}$*;*
2. *R's expected guess matches the expected type for the signal region* $\{6, 7, 8, 9, 10, 11, 12, 13, 14\}$*.*

Prediction 2 *Under mandatory disclosure regime, R's expected guess exceeds expected type for the advantaged sender group. Specifically,*

1. *R's expected guess exceeds the expected type for the signal region* $\{10, 11, 12, 13, 14\}$*;*
2. *R's expected guess matches the expected type for the signal region* $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$*.*

When the information transmission regime is voluntary disclosure, it is optimal for S to disclose if and only if her signal is sufficiently high. S with marginal signal s^* is indifferent between disclosing and not disclosing. Note that disclosing s^* decreases naive R's guess and increases sophisticated R's guess relative to not disclosing. Upon nondisclosure, naive R chooses the prior mean ($g_{\chi} = 7.5$) and sophisitcated R chooses the average silent type ($g_{(1-\chi)} = \mathbb{E}(t \mid s \leq s^*)$). Then, S with higher signals indeed find it optimal to disclose and the others to remain silent.

Prediction 3 *Under voluntary disclosure regime,*

1. *for any* $\chi \in (0, 1)$*, there exists a unique cutoff signal* $s^*(\chi) \in \{0, 1, 2, 3, 4\}$*, such that senders with signal* $s \leq s^*$ *do not disclose* ($m = \Psi$) *and senders with* $s > s^*(\chi)$ *disclose* ($m = s$)*;* $s^*(\chi)$ *is weakly increasing in* χ*.*

2. *R's expected guess:*

   (a) *exceeds expected (non-disclosing) S type upon nondisclosure ($m = \Psi$),*

   (b) *follows the patterns described in* ***Prediction 1*** *and* ***Prediction 2*** *upon nondisclosure ($m =$).*

Based on equilibrium, following predictions compare players' behavior.

Prediction 4 (Disclosure Behavior Comparison) *For any $\chi \in (0, 1)$, the disadvantaged sender group discloses a weakly greater range of signals than the advantaged sender group, i.e. $s^*_{dis.} \leq s^*_{adv.}$*

Prediction 5 (Guess Comparison) *For any given $\chi \in (0, 1)$,*

1. *R's expected guess for both sender groups is higher under voluntary disclosure than under mandatory disclosure;*

2. *R's expected guess increase is greater for the disadvantaged sender group than for the advantaged sender group.*

1.3 Experiment

1.3.1 Experimental Procedure

The experiment was programmed in oTree (Chen, Schonger and Wickens, 2016). We recruited experiment participants over research subject recruitment website Prolific (prolific.co). We conducted our sessions online between June 2021 and August 2021. We limited participants to U.S. residents with web-cameras.[13] A Zoom meeting was used to imitate the conditions of the lab, including one-by-one check-in process, going through instructions and answering any questions subjects had via Zoom chat box. We put subjects in Zoom breakout rooms based on their (fixed) roles in the sender-receiver game.

One session lasted about 60 minutes and the average earnings (including a \$6.50 fixed reward) were \$11.43, with minimum earnings of \$8.37 and maximum earnings of \$15.25. A total of 500 subjects participated in 38 sessions.[14] Each sessions had 8 to 20 participants.

[13] We required participants to keep their web-cameras on throughout the session to ensure they were paying attention.

[14] We conducted 7 sessions of Treatment MA, 6 sessions of Treatment MD, 5 sessions of Treatment VA, 9 sessions of wave Treatment M, and 11 sessions of Treatment V. Overall, we had a total of 68 participants for Treatment MD and Treatment MA each, 76 participants for Treatment VA, 132 participants for Treatment M, and 156 participants for Treatment V.

half the participants played in the role of senders and half in the role of receivers.

1.3.2 Design

The experiment consists of 15 rounds of the sender-receiver game, followed by out-of-sample belief elicitation, a lottery task (Dohmen and Falk, 2011) to measure risk attitudes, a three-item cognitive reflection test (Frederick, 2005) to measure cognitive ability, and finally a short demographic survey.[15] Demo of the experiment can be reached here.

1.3.2.1 The Sender-Receiver Game

For ease of exposition, we introduced the game as "Guessing the Secret Number" game, where "secret number" referred to the state of the nature/type of S. In each treatment, we split participants randomly in half and assigned them the role of the sender or the role of the receiver. We referred to senders and receivers as "Observers" and "Guessers" of the secret number, respectively. Subjects then remained in their assigned roles for all 15 rounds. We then matched them randomly and anonymously with a subject in the other role each round. This ensured our results were not contaminated by repeated game or reputation effects.

Each round, sender's type was drawn from the set $\{5, 6, 7, 8, 9, 10\}$ and sender's error term was drawn from the set $\{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$. Although each type was equally likely to be selected, the likelihoods of error terms depended on the sender group. We called the error term, the "color number" and user sender group names that had color associations with no semantic association within the game. Specifically, we called the disadvantaged sender group, the "Sky color group" and the advantaged sender group, the "Forest color group". Using colorful wheels that emphasized the different probabilities of each error term, we explained that for the advantaged (disadvantaged) sender group, half the time the error term was equal to 5 (-5) and half the time and to one of the remaining error terms, each with equal likelihood. Figure 1.1 displays the related part of the instructions used in the experiment. The type and error term selection process was common knowledge among participants.

Types and error terms were then summed up to generate sender's signal, a whole number between (and including) 0 and 15. Under mandatory disclosure regime, the signals were automatically shared with the receiver, whereas under voluntary disclosure regime senders decided to share or withhold the signal. After observing the sender's signal or that the sender chose to not disclose her signal, the receiver submitted his guess about the type of

[15]A central result in the literature (Bergman et al., 2010; Oechssler, Roider and Schmitz, 2009) is that individuals with low cognitive abilities tend to be significantly more affected by behavioral biases.

the sender in 0.5 increments, i.e. $g \in \{5, 5.5, , 6, 6.5, 7, 7.5,\ 8, 8.5, 9, 9.5, 10\}$. The incentives were designed to reflect the utility functions in our model. The sender's earnings grew in receiver guesses, whereas the receiver was rewarded for accuracy of his guesses. We asked participants to answer comprehension questions prior to the start of the game. Participants could only start playing the sender-receiver game once everybody correctly answered all of their comprehension questions. Appendix figure A.1 and A.2 display the screenshots of the sender (Observer) and the receiver (Guesser) decision screens, respectively.

Receiver point payoffs each round was equal to $p_R = 110 - 15|t-g|^{1.4}$. This functional form increases strictly as the guess gets closer to the sender type. Sender point payoffs depend only on the receiver's guess and increase strictly with the receiver guesses $p_S = 110 - 15|10-g|^{1.4}$, since the maximum guess is 10. The small number of receiver actions and sender types allowed us to represent payoffs for each scenario in the game using a table. This way, participants did not need to use the functional forms for their payoffs.

The payoffs could vary substantially across combinations of types and guesses. In order to let performance play a greater role than luck, we allowed subjects to accrue points in all rounds towards their final payment.[16] The total points were then converted to dollar amounts, using conversion rates 125 : 1 (125 points = $1) for senders and 215 : 1 (215 points = $1) for receivers. We display a copy of instructions in Appendix A.1.3 and the payoff tables in Appendix figure A.5.

Treatment Variation. The experiment consisted of 5 treatments that varied in two dimensions: the information transmission regime and the sender group. As demonstrated in Table 1.2, we split these treatments into 3 "Treatment Waves" based on the purpose they serve. Wave 1 consisted of mandatory disclosure regime treatments, one for each sender group (Treatment MD and Treatment MA). These treatments provided us with clean identification regions to vindicate our model's predictions for disclosed messages. Wave 2 consisted of voluntary disclosure regime treatment for the advantaged sender group only (Treatment VA). This treatment was used to test our model's prediction for nondisclosed signals. We run this treatment only for the sender group that provided us with the cleanest identification. Finally, Wave 3 had two treatments, one under mandatory disclosure regime the other under voluntary disclosure regime, each with both groups of senders (Treatment M and Treatment V, respectively). These treatments facilitated a between subject design for welfare analysis of transition into Salary History Ban, under wage (signal) inequality across two sender groups.

[16] Note that we only told participants the total points they accrued at the very end of the game in order to alleviate the possibility of wealth and portfolio effects.

Out-of-Sample Beliefs. After the main part of the experiment, we conducted incentivized belief elicitation for both senders and receivers. We collected senders' beliefs for the average receiver guess in the pilot session after each message from 0 to 15, and receivers' belief for the sender type in different signal regions. Further, we asked receivers for the average generated (regardless of disclosure) signal to elicit their understanding of the signal distribution.

In Voluntary Disclosure treatments, we also collected sender's beliefs of the receiver guess distribution upon nondisclosure in the pilot session. Similarly, we collected receivers' beliefs for the type distribution upon nondisclosure in the pilot experiment. We incentivized subjects to induce proximity to the empirical distribution of the variable in question.

1.3.2.2 Two Additional Tasks

Risk Preferences. In order to elicit participants' risk preferences, we used the method introduced by Holt and Laury (2002). For this measure, we asked subjects to make 15 choices between a fixed payoff or a 50-50 lottery. Fixed payoff options ranged from 2.5 to 37.5 points increasing in 2.5 increments, whereas 50-50 lottery was fixed to always be between 40 and 0 points. One decision was then selected at random and the corresponding points were converted to US dollars for payment.[17] We counted the number of times each subject preferred the fixed payoff to create a measure of risk aversion.

Cognitive Reflection Test. We used cognitive reflection test (CRT) to generate a measure of reflective thinking (Frederick, 2005). The test had 3 seemingly straightforward questions each with an intuitive yet incorrect answer, and a correct answer that required a few moments of consideration. CRT performance has been shown to have strong correlation with heuristics-and-biases (Oechssler, Roider and Schmitz, 2009), bergman2010anchoring. Participants were paid for each correct answer.

1.3.2.3 Follow-up Survey

Upon completion of above mentioned tasks, we conducted a short survey to collect the following information: age, gender, education level, and whether English was their first language.

[17] Conversion rates of points were given presented to participants prior to their decision page.

1.4 Results

We first present wave 1 and wave 2 treatments to investigate the validity of our model. We then move to wave 3 treatments. In this section, we focus on the comparison of wave 3 treatments and also present their model validity checks for Appendix A.1.2.

1.4.1 Wave 1

1.4.1.1 Wave 1: Treatment MD

According to theory, receivers underestimate the sender type in mandatory disclosure treatment when the sender group has right skewed error distribution (disadvantaged). More specifically, we expect to see the underestimation to occur for the signals where a rational and naive receivers' guess would deviate, i.e., signals 1-5.

Table 1.3 shows benchmark rational and naive receiver guesses for signal regions. We see that the observed mean receiver guess differs from the rational benchmark in the direction our model predicts. Further, we see that receiver guesses are not significantly different than the rational benchmark for the signal region 6-14, where benchmark for rational and naive receivers overlap.

At an individual level, 59% of receiver subjects have higher absolute average deviation from the rational benchmark in the shaded region than in the signal region 6-14 (Wilcoxon signed-rank test: $p = 0.0216$).

1.4.1.2 Wave 1: Treatment MA

Next, we move to mandatory disclosure treatment for the (advantaged) sender group with left skewed error distribution. Our model predicts that the presence of naive receivers will lead to underestimation of the sender type. Once again, we focus on the signal region where the benchmark predictions of rational and naive receiver differ (signals 10-14). Table 1.4 indeed shows that receiver guesses in the experiment are greater than the average sender type in this region. Moreover, we see that the receiver guess for the signal region 1-9 also lines up with our model's prediction, i.e., it is not significantly different than what the rational benchmark suggests.

At an individual level, 61% of receiver subjects have higher absolute average deviation from the rational benchmark in the shaded region than in the signal region 1-9 (Wilcoxon signed-rank test: $p = 0.0135$).

1.4.2 Wave 2

1.4.2.1 Wave 2: Treatment VA

We move on the voluntary disclosure treatment for the (advantaged) sender group with left skewed error distribution to test out model's prediction of receiver behavior upon nondisclosure.

First, we start with sender behavior. Disclosure is no longer the default option and we expect partial unraveling of signals. Figure 1.2 shows that, in line with our theoretical predictions, senders follow a threshold strategy. The disclosure rate is close to zero for the signals 0-3, over 50% for signal 4 and over 60% for the higher signals. Note that after the initial 7 rounds, the disclosure rates of highest signals increase considerably; e.g. for signal 14 it increases from 83.9 percent to 94.1 percent and for signal 15 it increases from 86.4 percent to 97.3 percent. Overall, senders behavior gets closer to our model's assumption of sender rationality. Tables 1.5 and 1.6 display that the disclosure rates increase for high signals and decrease for low signals as senders gain experience in the game. There is also sharp increase in the disclosure rate of signal 4 and 5. We interpret this as a deliberate shift in cutoff strategy in response to sender behavior in the first rounds. Focusing on the first 5 rounds, we see that the disclosure rate jumps from below 50% to above 50% across signals 5 and 6. Notice that this jump happens across signals 3 and 4 on the later rounds. This implies a shift in disclosure cutoff with experience. Considering receivers become more rational with experience (smaller χ), this behavior is in line with our prediction that $s^*(\chi)$ is increasing in χ.

We then move to receivers' guesses. Table 1.7 separates the nondisclosure region (signals 0-3) and the shaded signal region (10-14) from the rest of the signals. These are the regions where we expect to see deviation from rational benchmark. More specifically, our model predicts overestimation of receiver type for both regions. We see that indeed the receiver subjects in Treatment VA behaved inline with our theoretical predictions.

At an individual level, of the receivers who faced with both nondisclosure and disclosed signals 4-9 (37%), 63% have higher absolute average deviation from the rational benchmark in the nondisclosure signal region 0-3 than in the signal region 4-9 (Wilcoxon signed-rank test: $p = 0.0153$). Similarly, of the receivers who have faced with both (68%), 71% of receiver subjects have higher absolute average deviation from the rational benchmark in the shaded signal region 10-14 than in the signal region 4-9 (Wilcoxon signed-rank test: $p = 0.0167$).

Interestingly, of the receiver subjects who encountered both nondisclosure and signals from the identification region (48%), 67% have higher absolute average deviation from the rational benchmark in the shaded signal region 10-14 than in the nondisclosure signal region

(Wilcoxon signed-rank test: $p = 0.003$). This implies that, although we observe receiver behavior aligned with our model's predictions, the share of naive receivers are greater for the identification region signals (10-14) than for the nondisclosure region.

We also examine absolute deviation from Bayesian benchmark for the shaded identification region across Treatment MA and Treatment VA. Regressing absolute deviation on a dummy variable representing voluntary disclosure, we find that voluntary disclosure increases deviations by 0.19 ($p = 0.0292$).

Further, we compare receiver guesses for both the disclosed signals and nondisclosure. Figure 1.3 gives us a general sense of equivalence between the average guess upon nondisclosure and upon disclosed signals of 3 and 4. Average receiver guess upon nondisclosure is equal to 6.6, which is greater than or equal to the average guess for disclosed signals when $s \leq 4$ and significantly less than the average guess when $s > 4$ (Wilcoxon signed-rank tests for equality when disclosed message ≤ 4, ≤ 5: p=0.1026, p=0.0311, respectively). When we bundle signals accordingly into groups of three as in figure 1.4, we see that, not disclosing, indifference, and disclosing are the best responses to receivers' behavior for the sufficiently low signals (0-3), signal 4, and sufficiently high signals (5-15), respectively.

1.4.2.2 Heterogeneity in Risk Attitudes and CRT Performance

Before moving on to the main results, we put our model's predictions to another test. Risk preferences can impact receiver guesses when there is uncertainty about the underlying state, and risk aversion can push guesses of receivers towards types for which there is less variation. This would also imply overestimation of nondisclosed signals, as well as over-(under-)estimation of the disclosed identification region signals for advantaged (disadvantaged) sender group with left (right) skewed distribution.

After playing the main game, subjects completed the risk preferences task. We generated a risk measure by dividing the times a subject chose the fixed payoff over the 50-50 chance lottery. This gave us a distribution of subjects' risk aversion measure, that was between 0 an 1.

In addition, we asked subjects to answer 3 questions know as cognitive reflection test (CRT). We let them collect 1 point for each correct answer. CRT performance has been shown to have strong correlation with heuristics-and-biases (Bergman et al., 2010; Oechssler, Roider and Schmitz, 2009). We consider better performance in CRT a measure of sophistication in our context.

Table 1.8 display the summary statistic for our CRT and risk measures. Using the median values, we split participants into two categories for both measures: high and low. Pooling receiver guesses from Treatment MA, MD and VA, we run two regressions as displayed in

Table 1.9. We define mistakes as absolute deviation from Bayesian benchmark. We see that risk aversion does not correlate with receiver's deviation. Further, column 2 results indicate that CRT performance is negatively correlated with deviation from Bayesian benchmark.

1.4.3 Wave 3

1.4.3.1 Wave 3: Treatment M versus Treatment V

Next, we move on to the last wave of our design. We leave the above model verification type analyses for wave 3 to Appendix A.1.2 and focus on comparison of Treatment M and Treatment V here. We start with sender behavior in Treatment V.

Figure 1.5 shows that, in line with our theoretical predictions, the (disadvantaged) sender group with right skewed error distribution follows a threshold strategy. The disclosure rate is almost zero for the signals 0-2, over 62% for signal 3 and over 60% for the rest of the signals.

Similarly, figure 1.6 shows that, the (advantaged) sender group with left skewed error distribution also follows a threshold strategy. The disclosure rate is almost zero for the signals 0-2, less than 30% for the signal 3, and above 56% for the higher signals. Further, the disclosure threshold for the (advantaged) sender group with left skewed error distribution is greater than that of other sender group, as predicted.

Table 1.10 displays our main results. Our model predicts that moving from mandatory to voluntary disclosure will increase receiver guesses. Further, we expect this increase to be of greater magnitude for the (disadvantaged) sender group with right skewed error distribution. Given the under/over estimation for disclosed signals, this implies narrowing of the receiver guess gap between two groups. Column 3 of Table 1.10 repeat the regression in column 2, but include only theory-conforming senders' observations. We define theory-conforming disclosure behavior based on empirically observed thresholds of signal 2 for the disadvantaged sender group and signal 3 for the advantaged sender group.

We see that shift to voluntary disclosure increased receiver guesses as predicted. However, surprisingly, we see that this increase is of smaller magnitude for the disadvantaged sender group with right skewed errors. This result, although not in line with our theoretical predictions, is a corollary of the observations we made during analyses of wave 1 and wave 2 treatments (see Appendix A.1.2 for the same analyses of wave 3 treatments). That is, the shift to voluntary disclosure increases the deviation from Bayesian benchmark for the disclosed signals. Further, as stated in the previous section, we observe greater average absolute deviation for nondisclosure than disclosed identification region signals at an individual level.

To investigate this further, columns 4-7 of table 1.10 display the treatment effect on guesses for disclosed signals separately for each sender group. Columns 4-5 include only the (advantaged) sender group with left skewed error distribution, such that column 4 focuses on signals 0-9, where naive and sophisticated receiver guesses coincide and column 5 focuses on signals 10-14, the model identification region. Similarly, column 6 focuses on the identification region for (disadvantaged) sender group with right skewed error distribution and column 7 on the signal region 6-15. We see that, indeed the receiver guesses are higher for the identification region of the advantaged senders. That is, the overestimation of the sender type, due to receiver (statistical) naivete, is of greater magnitude when the disclosure is voluntary. On column 7, we can also see that the underestimation of disadvantaged sender type, due to receiver (statistical) naivete is of greater magnitude. However, that is a noisy estimate so we cannot reject null treatment effect.

Figure 1.7 and 1.8 further visually compare receiver behavior upon nondisclosure and disclosed signals by the same sender group across Treatment M and Treatment V. We see that the average guess upon nondisclosure for the disadvantaged and advantaged sender groups are 6.56 and 6.6 respectively.

1.5 Discussion

We investigate how prospective employers' (receivers) inference will be impacted by the Salary History Bans under wage discrepancies. We suggest that the under Salary History Ban applicants (senders) may choose to voluntarily disclose their salaries to differentiate themselves. We use a behavioral model to incorporate the potential flaw in prospective employers' inference of applicants' productivity type when salary history has a systematically biased relationship with applicants productivity levels.

The experimental results from the first two waves largely confirm our hypotheses based on the partially naive receivers. First, receivers underestimate the (disadvantaged) sender types when the bias in signals is right skewed, i.e. mass on the negative values of bias. Similarly, receivers overestimate the (advantaged) sender types when the bias in signals is left skewed, i.e. mass on the positive values of bias. Further, receivers overestimate sender type when the signal is not disclosed. Finally, we show that senders with right skewed bias disclose a greater range of signals, i.e. the disclosure threshold is lower relative to the other sender group.

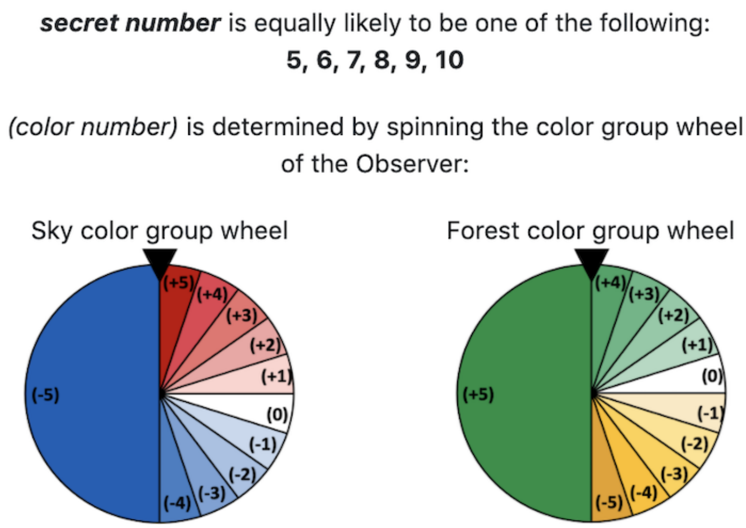
Our model, predicts that the shift to voluntary disclosure regime will decrease the payoff gap across two sender groups. This result stems from the expectation that both groups' payoffs will increase with the shift, but the disadvantaged sender group will have a greater increase. In contrast, we observe that both groups benefit from the shift to voluntary disclosure yet the disadvantaged group experiences a smaller payoff increase. This result appears to be due to an increase in receiver naivete for the disclosed identification region signals under voluntary disclosure relative to mandatory disclosure. Further, we find that the receiver mistakes are of lower magnitude for the nondisclosed signal regions relative to the disclosed identification region signals under voluntary disclosure regime.

Our results suggest that when salary history disclosure is at senders' discretion, the biases can shift in a way that would lead to unintended consequences such as widening the pre-existing wage discrepancies due to employers' naivete. This result, although may seem surprising at first glance, is not new. Previous literature in regulating information transmission has already displayed similar results in contexts like ban-the-box.

If the goal is to decrease wage discrepancies, a better regulation may focus on the demand side, such as pay transparency regulation. As Cullen and Pakzad-Hurson (2019) display, although in equilibrium, pay transparency leads to decreased wages overall, it could be a remedy to unequal pay. Historically disadvantaged employees would then have the chance to do their research and hold a conversation around the reported salary range of the new workplace.

1.6 Figures & Tables

Figure 1.1: Depiction of Type and Error Term Structures in the Experiment



Notes: Type and error term were "secret number" and "color number" in the experiment for ease of exposition. We gave sender groups names without any semantic meaning; "Forest color group" and 'Sky color group". Connecting the error term to colorful wheels for each sender group let us explain the likelihoods of error terms with ease.

Figure 1.2: Disclosure rates in Treatment VA

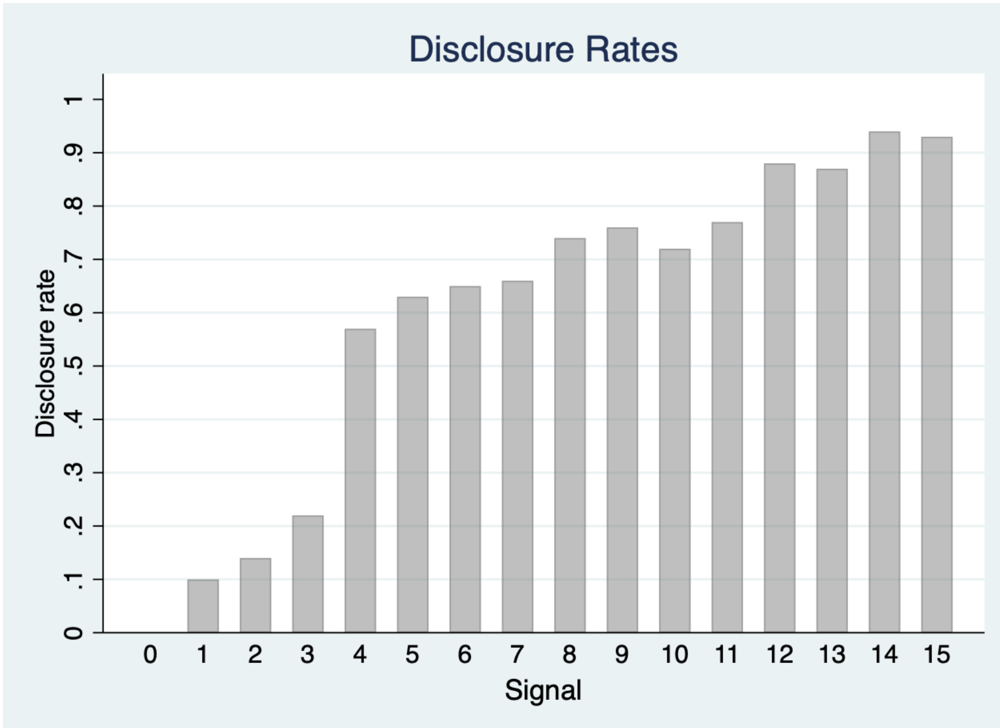
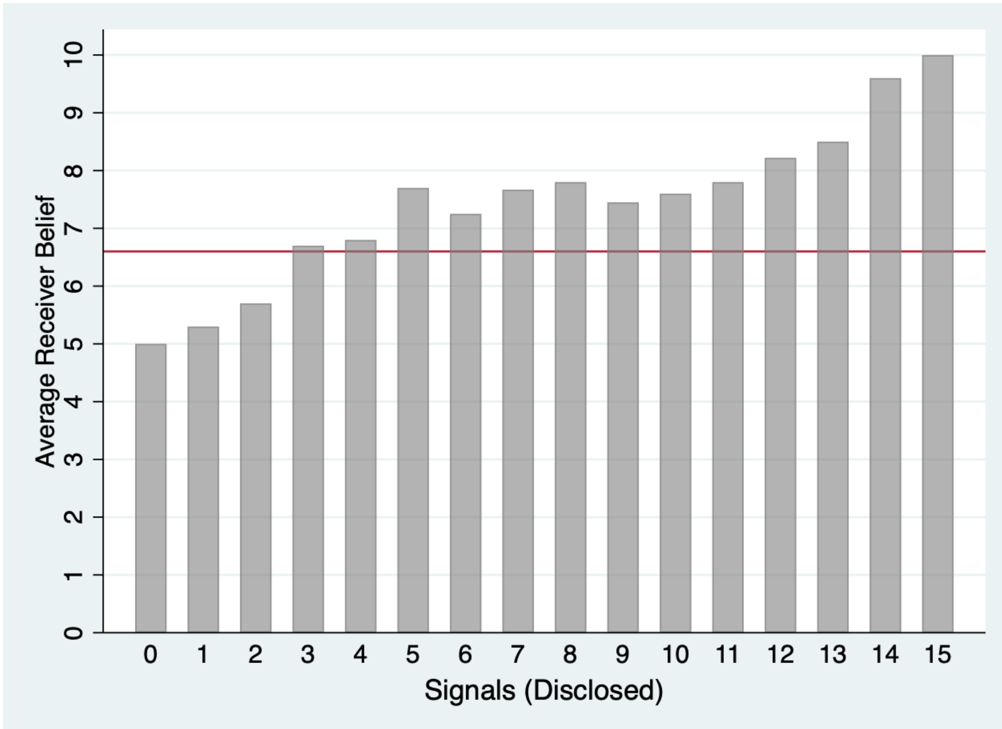
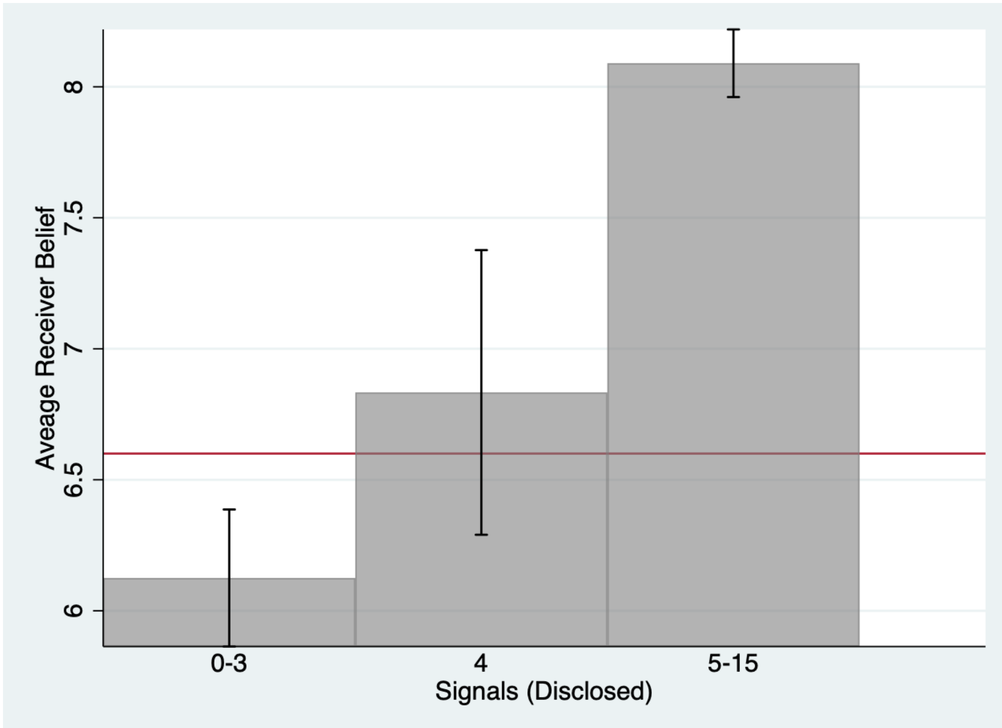


Figure 1.3: Average Receiver Beliefs in Treatment VA



Notes: Bars indicate the average receiver guess for each disclosed signal. Red horizontal line represents the average receiver guess upon nondisclosure at 6.6.

Figure 1.4: Average Receiver Beliefs in Treatment VA



Notes: Bars indicate the average receiver guess for disclosed signals, bundled in reference to figure 1.3. Vertical lines represent confidence intervals for each bar. Red horizontal line stands for the average receiver guess at 6.6.

Figure 1.5: Disclosure rates of (Disadvantaged) Senders with Right Skewed Errors in Treatment V

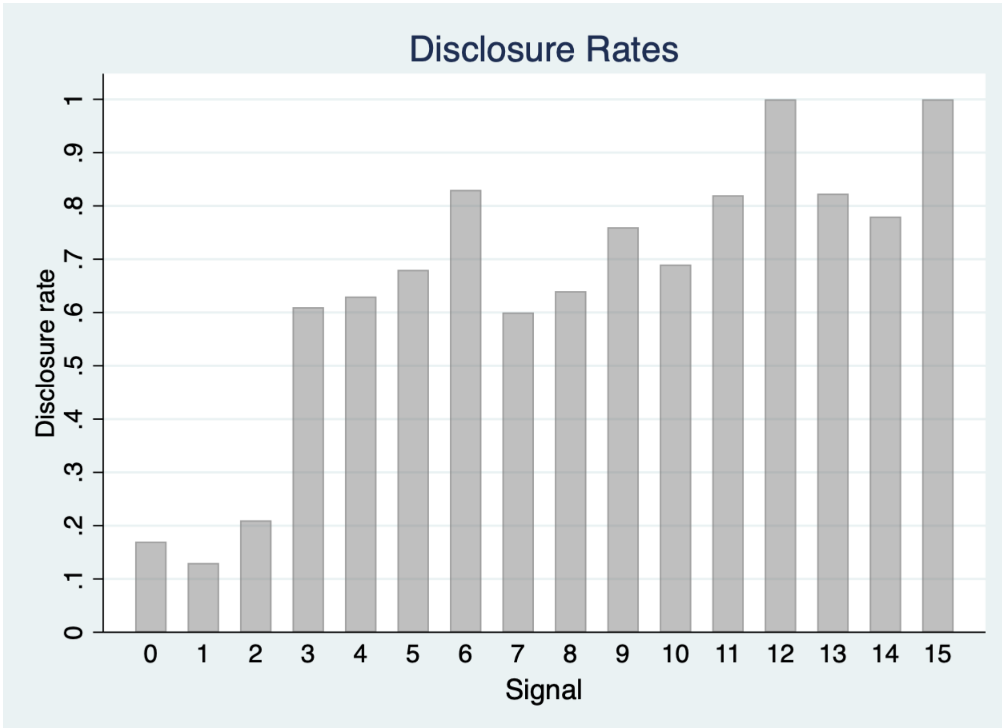


Figure 1.6: Disclosure rates in Treatment V for (Advantaged) Senders with Left Skewed Errors

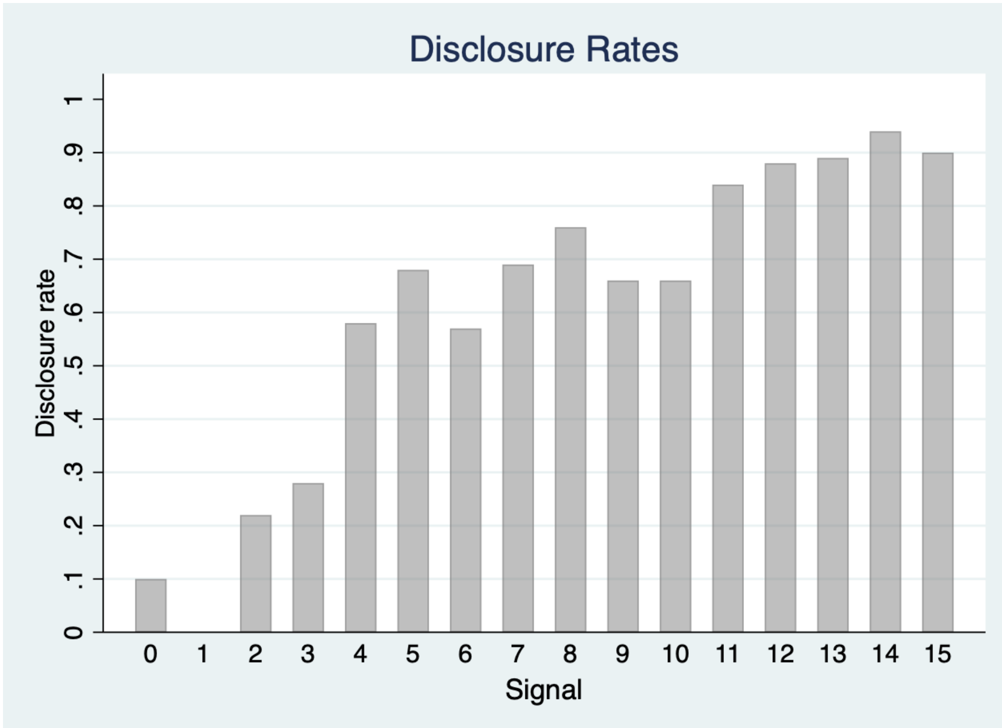
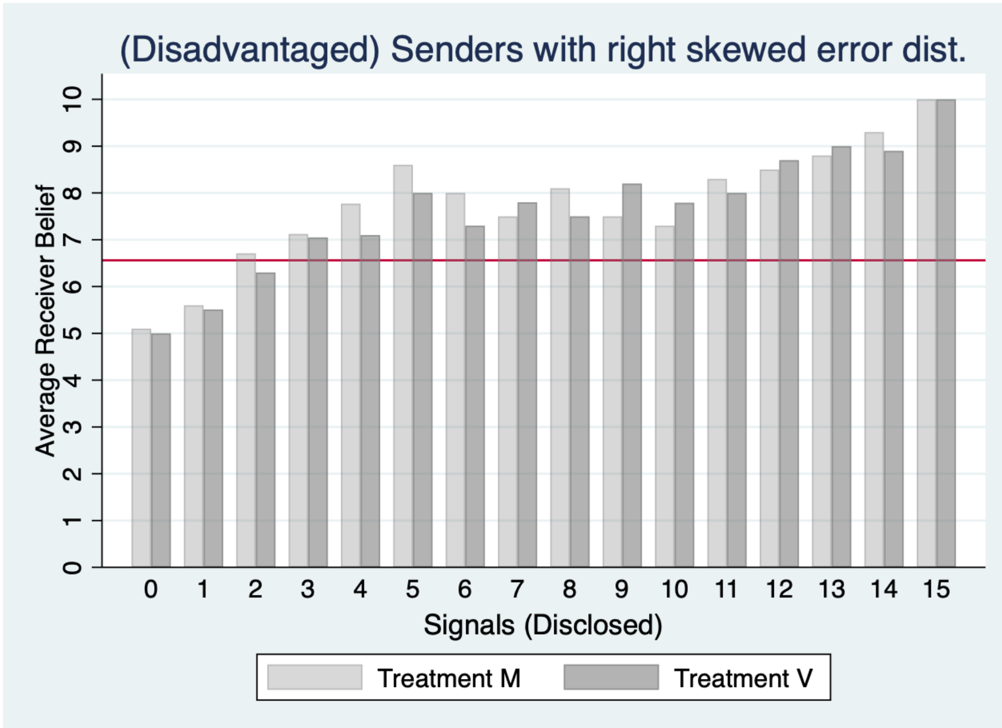
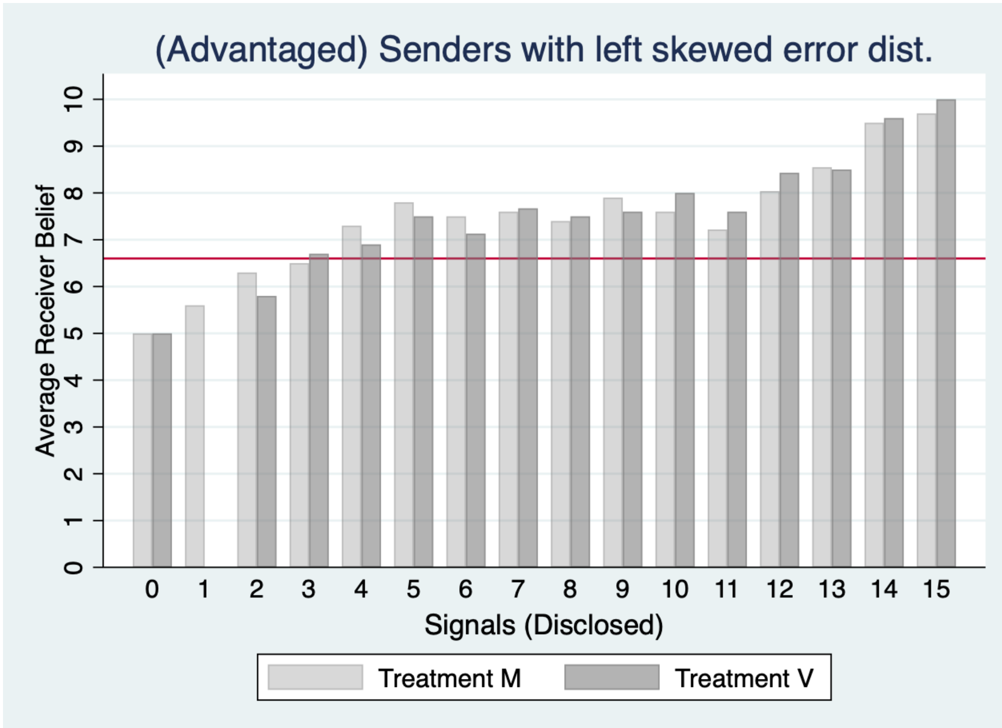


Figure 1.7: Average Receiver Guesses in Wave 3 for (Disadvantaged) Senders with Right Skewed Errors



Notes: Average receiver guess for disclosed signals of the (disadvantaged) sender group with right skewed error distribution. Red line indicates the average receiver guess upon nondisclosure in treatment V.

Figure 1.8: Average Receiver Guesses in Wave 3 for (Advantaged) Senders with Left Skewed Errors



Notes: Average receiver guess for disclosed signals of the (advantaged) sender group with right skewed error distribution. Red line indicates the average receiver guess upon nondisclosure.

Table 1.1: Signals and type support of each signal.

Signals	**Types** (and corresponding Error terms)
0	**5**(-5)
1	**5**(-4), **6**(-5)
2	**5**(-3), **6**(-4), **7**(-5)
3	**5**(-2), **6**(-3), **7**(-4), **8**(-5)
4	**5**(-1), **6**(-2), **7**(-3), **8**(-4), **9**(-5)
5	**5**(0), **6**(-1), **7**(-2), **8**(-3), **9**(-4), **10**(-5)
6	**5**(1), **6**(0), **7**(-1), **8**(-2), **9**(-3), **10**(-4)
7	**5**(2), **6**(1), **7**(0), **8**(-1), **9**(-2), **10**(-3)
8	**5**(3), **6**(2), **7**(1), **8**(0), **9**(-1), **10**(-2)
9	**5**(4), **6**(3), **7**(1), **8**(1), **9**(0), **10**(-1)
10	**5**(5), **6**(4), **7**(3), **8**(2), **9**(1), **10**(0)
11	**6**(5), **7**(4), **8**(3), **9**(2), **10**(1)
12	**7**(5), **8**(4), **9**(3), **10**(2)
13	**8**(5), **9**(4), **10**(3)
14	**9**(5), **10**(4)
15	**10**(5)

Table 1.2: Treatments

	Disadvantaged Senders	Advantaged Senders	All Senders
Mandatory Disclosure	Treatment MD	Treatment MA	Treatment M
Voluntary Disclosure		Treatment VA	Treatment V

Notes: Treatments were conducted in waves of three. Wave 1 was composed of Treatment MD and Treatment MA, Wave 2 was Treatment VA, and Wave 3 consisted of treatments with both sender groups; Treatment M and Treatment V. Wave 1 and Wave 2 treatments were used to test our model, whereas Wave 3 provided welfare analysis for Salary History Ban.

Table 1.3: Receiver Naivete by Signal Regions in Treatment MD

Signals	Bayesian Belief (Benchmark)	Naive Belief (Benchmark)	Mean Belief (Observed)	Wilcoxon signed-rank (p-value)	Observations
0	5	5	5		
1-5	7.62	6.58	7.19	0.0000, 0.0000	245
6-14	7.84	7.84	7.72	0.263	150
15	10	10	9.95		

Notes: Wilcoxon signed-rank test results compare the benchmarks to the observed values.

Table 1.4: Receiver Naivete by Signal Regions in Treatment MA

Signals	Bayesian Belief (Benchmark)	Naive Belief (Benchmark)	Mean Belief (Observed)	Wilcoxon signed-rank (p-value)	Observations
0	5	5	5		
1-9	7.16	7.16	7.13	0.2394	160
10-14	7.39	8.42	7.83	0.0000, 0.0000	243
15	10	10	9.95		

Notes: Wilcoxon signed-rank test results compare the benchmarks to the observed values.

Table 1.5: Evolution of Disclosure Rates by Signal (Signals 0-7)

Signal:	$s = 0$	$s = 1$	$s = 2$	$s = 3$	$s = 4$	$s = 5$	$s = 6$	$s = 7$
Rounds 1-5	0%	14%	18%	23%	35%	43%	61%	64%
Rounds 6-10	0%	10%	15%	21%	56%	60%	63%	62%
Rounds 11-15	0%	6%	9%	21%	66%	75%	70%	68%

Table 1.6: Evolution of Disclosure Rates by Signal (Signals 8-15)

Signal:	$s = 8$	$s = 9$	$s = 10$	$s = 11$	$s = 12$	$s = 13$	$s = 14$	$s = 15$
Rounds 1-5	67%	73%	69%	73%	79%	82%	87%	85%
Rounds 6-10	72%	76%	71%	75%	85%	93%	92%	94%
Rounds 11-15	83%	79%	76%	83%	97%	96%	95%	97%

Table 1.7: Receiver Naivete by Signal Regions in Treatment VA

Signals	Bayesian Belief (Benchmark)	Naive Belief (Benchmark)	Mean Belief (Observed)	Sign-rank Test (p-value)	Observations
0-3	6	7.5	6.7	0.000, 0.0000	47
4-9	7.42	7.42	7.51	0.2394	162
10-14	7.39	8.42	8.02	0.0000, 0.0000	227
15	10	10	9.95		

Notes: Wilcoxon signed-rank test results compare the benchmarks to the observed values.

Table 1.8: Receiver Naivete by Signal Regions in Treatment MA

	Mean	Median	Std. Dev.	Min	Max
CRT	1.624	2	1.243	0	3
Risk	0.532	0.533	0.219	0	1

Table 1.9: Effect of CRT performance and Risk preferences on Receivers' Deviation from Rational Benchmark.

	(1)	(2)
Dependent Variable:	Mistake	
Voluntary disc.	0.516**	0.389**
	(0.1700)	(0.1364)
High CRT		-0.210**
		(0.08386)
High Risk Aversion	0.119	
	(0.08027)	
High CRT X Vol. disc.		-0.3152*
		(0.1644094)
High Risk Aversion X Vol. disc.	-0.213	
	(0.1776)	
Round	0.002	0.001
	(0.0048)	(0.0048)
Constant	1.375***	1.581***
	(0.1558)	(0.1927)
Signal dummies	Yes	Yes
R^2	0.195	0.2241
Observations	1590	1590

Notes: Column 1: OLS regression of the effect of CRT on receivers' deviation from rational benchmark. Column 2: OLS regression of Risk measures on receivers' deviation from rational benchmark. Robust standard errors clustered at the subject level in parenthesis; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 1.10:
Voluntary Disclosure Regime Effect on Receivers' Guesses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable:	Guess	Guess	Guess s				
Voluntary disclosure	0.093**	0.199*	0.217**	0.079	0.187**	-0.063	-0.101
	(0.0392)	(0.1054)	(0.1018)	(0.058)	(0.081)	(0.045)	(0.077)
Disadvantaged sender		-0.402**	-0.434**				
		(0.1827)	(0.1808)				
Vol. disc. X Disad. sender		-0.0694**	-0.096**				
		(0.0316)	(0.0401)				
Round	0.0059	0.0062	0.0081	0.0032	0.0020	0.0093	0.0038
	(0.0064)	(0.006)	(0.0064)	(0.0073)	(0.0068)	(0.0081)	(0.0084)
Constant	7.385***	7.721***	7.719***	7.623***	7.743***	7.217***	7.367***
	(0.145)	(0.153)	(0.224)	(0.108)	(0.125)	(0.076)	(0.137)
Signal dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Signal region	All	All	All	0-9	10-14	1-5	6-15
Included senders	All	All	Theory-conforming	Left (Adv.)	Left (Adv.)	Right (Dis.)	Right (Dis.)
R^2	0.436	0.447	0.475	0.336	0.460	0.358	0.354
Observations	1778	1778	1536	264	445	281	265

Notes: robust standard errors clustered at the subject level in parenthesis; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

CHAPTER II

Identity Discrimination in the Online Marketplace

Abstract

The number of online marketplaces and sharing economy platforms are rapidly increasing. Airbnb, is a classic example of online marketplace that allows users to post or rent short-term housing in residential premises. Through an online audit study, we demonstrate an important consequence of Airbnb's current market design. There is a nine percentage point gap between the acceptance rate of requests from guests with distinctively Arab/Muslim identities and those of distinctively non-Arab/Muslim identities. The difference persists across different specifications.

JEL Classification: D40, C93, J71
Keywords: discrimination, marketplace

2.1 Introduction

According to the Arab American Institute, considerably large immigration from Arab countries to the US began in the 1880's. The current estimates suggest that currently 3.7 million Americans have Arab roots. Although Arabs and Muslims constitute a substantial share of Americans, there is notable lack of research on these demographic. One of the most natural starting points is the treatment of Arabs and Muslims, specifically an investigation of discriminatory behavior.

Although there are regulations against discriminatory behavior in the area of housing and rental accommodations, the rise of the internet has changed these markets significantly. Airbnb is the leading online platform which provides users with a peer-to-peer short term rentals. The well-established anti-discrimination laws in the rental accommodation market is not likely to affect faster pace and the smaller "landlords" (hosts) of Airbnb. In offline markets, information disclosure tends to be almost automatic. In principle, the rise of online marketplaces makes it easy to envision a platform which naturally limits opportunities of discrimination by deciding the level of personal information shared across users. Examples for such online platforms are eBay or Expedia, which choose to withhold some user information from market participants. Airbnb, on the other hand, uses a different model that is prone to discriminatory behavior. Consider a guest seeking to book an apartment or room via Airbnb. Although ideally the name and the photo of an interested guest should not play a role in host's decision to accept guest's booking request, Airbnb provides this information to the hosts prior to booking confirmation. In this study, we conduct a field experiment to identify discrimination on Airbnb. Specifically, we are investigating discriminatory behavior towards Muslim/Arab guests by the hosts over Airbnb.

The Experiment. In order to answer our research question, we follow the audit study methodology and create Airbnb guest accounts that differ by name and a profile photo. The profile photos display a silhouette in front of a sunset scenery without a visible face to avoid potential confounds due to facial characteristics. The guest accounts comprise of a treatment group with female Arab/Muslim sounding name and male Arab/Muslim sounding names, and a control group with female European origin sounding names and male European origin sounding names. Using these accounts, we inquire about the availability of 537 listings from Airbnb hosts across 12 states.

Overall, the results are concerning. We find discrimination against Arab/Muslim guests. The guests with European origins received an affirmative response from the hosts 84% of the time, compared to 75% for Arab/Muslim guests. We continue our investigation by controlling for listing and guest characteristics, which we access scraping publicly available data over Airbnb. We find that the 9 percentage point difference roughly stays the same across specifications that capture how professional the host is, the host's proximity to the guest within the unit, and whether the host is (approximately) under the age of 30 or not.

Implications. The results highlight an online platform's role in preventing discrimination or facilitating it. Airbnb has several design options to reduce discrimination, such as concealing guest and host names and let them use automatic salutation as it has been the case with eBay. Another natural next step would be to increase the use of 'Instant Book' option, which does not require hosts to manually approve a guest's request to rent at a date that is listed available. In fact, this option has been in use since 2016. Further, since this experiment, Airbnb has already made some changes: Airbnb now has a small team responsible for monitoring discriminatory behavior on the platform, users must now agree to Airbnb's 'Community Commitment and Nondiscrimination' policy, and they announced 'Project Lighthouse' on June 2020 which conducts audit studies of this type to investigate discriminatory behavior. Although it is noteworthy that Airbnb listened to researchers' and users' concerns regarding discriminatory behavior, it is not clear whether the actions they chose to take are enough to eradicate it.

Related Literature. Our paper uses the well-established audit study methodology to investigate discrimination on Airbnb. The seminal paper in this literature is Bertrand and Mullainathan (2004) in which authors send made up resumes to employers. The resumes differ in their perception of the name: African-American or White. In this clean yet effective field experiment, they document the striking result that White sounding names receive 50 percent more callback than the African-American sounding names. They also document that this racial gap persists across industry, employer size, and occupation. Researchers in many different fields continue to increasingly use this method, and our paper most closely follows Edelman, Luca and Svirsky (2017). They conduct the audit study over Airbnb to investigate discrimination against African-American guests. They display that guests with White sounding names are 16 percent more likely to receive an affirmative response than guests with African-American sounding names.

2.2 Background on Airbnb

Airbnb is an online marketplace for short-term housing and rental accommodations. It was founded in 2008 and since then expanded into an online service that connects hosts who want to rent out their apartment, room, sometimes even a castle to guests on the platform.

According to Airbnb's website as of August 2022, they have 6 million listings, over 4 million hosts, and more than 1 billion guest arrivals worldwide.[1]

Hosts display their listings by providing photos, and information about the price, number of bathrooms and bedrooms. Airbnb also displays information on hosts and guests, including their profile photos, their personal introductions as well as the previous reviews. Finally, note that although hosts are encouraged to update their calendars regularly, Airbnb still recommends guests to message hosts to confirm their listings' availability.

2.3 The Experiment

2.3.1 Experimental Procedure

We collected data from diverse geographic regions based on states' political leanings. We used the list called "Party Identification by State" by Gallup Organization, which distinguishes states on a scale of 'most democratic' to 'most republican' with 'lean democratic', 'competitive', 'lean republican' in between based on their survey results. Specifically, our data comes from the following states: Illinois, New Jersey, New Mexico, Oregon, Florida, Arizona, Missouri, South Carolina, Oklahoma, Mississippi, Montana, and Utah. Airbnb limits search per geographical location to 306 listings, so we were able to determine a total of 3672 listings to include in our experiment.

Using the guest accounts we have created, we were able to send 694 messages to hosts between March 25, 2017 and April 7, 2017.[2] The messages inquired about the availability of a specific weekend we pre-determined. We made sure those dates did not overlap with a holiday or big event in the regions we reached out for. When a host responded to our inquiry, we responded within 24 hours and clarified that we were just looking for options at that stage. This way we minimized the likelihood that our intervention would be costly for the hosts if they were to consider the listing reserved upon our inquiries.

2.3.2 Hypothesis

Bertrand and Mullainathan (2004) define discrimination as when members of a minority group are treated differently (less favorably) than members of a majority group with identical characteristics. In our setting, this definition corresponds to different acceptance probabilities across two groups of guests who are identical except for their religious/ethnic identity.

[1] See https://news.airbnb.com/about-us/

[2] Our plan was to collect 3,000 responses. However, we could not match our goal due to Airbnb starting to shut down our guest accounts.

2.3.3 Treatment Groups

We generated 2 main treatment groups, Airbnb guest accounts with Arab/Muslim sounding names and Airbnb guest accounts with non-Arab/Muslim sounding names. Each of these groups had an equal number of female and male sounding names. In total we created 20 guest accounts: 5 accounts with female Arab/Muslim sounding names, 5 accounts with male Arab/Muslim sounding names, 5 accounts with female non-Arab/Muslim sounding names, and 5 accounts with male non-Arab/Muslim sounding names. We selected the Arab-Muslim sounding names from 'the most common Arab/Muslim names' lists and the non-Arab/Muslim sounding names from the White sounding names used in Bertrand and Mullainathan (2004). We validated our selection by conducting a perception survey among Econ 102 students. We asked students to categorize each name by gender, religious affiliation, and race.[3] Following the norm elicitation protocol of Krupka and Weber (2013), we incentivized participants with a chance of winning a Starbucks gift card, conditional on their responses matching the responses of the majority. The survey results for the names used in our experiment are presented in Appendix table B.1.

Our guest accounts were identical except for the names and the profile photos. To avoid confounds that would result from profile photos with visible facial structures, we used photos with a silhouette that relates to the perception of the account name. We tried to standardize the photos as much as possible by selecting sunset photos that only include a silhouette of the upper body from behind. We used the same photo for all male accounts, whereas female accounts differed slightly. The accounts with female Arab/Muslim sounding names had a woman's silhouette who wore a hijab. All of our accounts joined the Airbnb within the same month and they all had a verified e-mail address and a verified phone number. It should be noted neither the e-mail address nor the phone number is visible to hosts before a reservation is made. Without a confirmed reservation, users can only see each others' verifications over Airbnb. Example Airbnb profiles for each treatment group can be seen on Appendix figure 2.1.

We randomly matched our 20 guest accounts with the prospective hosts, who were determined based on their availability through listing search in the states we listed above. We then sent messages that inquired about the availability of the corresponding listing. Figure 2.2 in Appendix shows an inquiry sent from one of our guests to an Airbnb host. We chose the dates of listing availability 4 weeks into the future from the time we sent the inquiries.

Hosts responded to our inquiries in various ways. We have encoded each response into one of the following groups: 'No', 'No unless more information is provided', 'Not sure, check

[3]The survey was granted a separate IRB approval.

back later', 'Request for more information', ' Yes but in a different property', 'Yes with questions', and 'Yes'. In our analyses below, we coded host responses to 1 (acceptance) for the answers 'Yes' and 'Yes with questions'. It should also be noted that we had to eliminate 157 inquiries for which we could not record a response due to Airbnb blocking our accounts. Therefore, our analyses will be based on the 537 observations for which we were able to successfully categorize the hosts' responses.

2.3.4 Observational Data

In addition to the experimental data, we were able to collect observational data by means of web-scraping. We used this method to collect information about the listings and the hosts. Specifically, the information we collected about the listings comprised of we the price, number of bedrooms, number of bathrooms, number of reviews, whether the property is shared or separate from where host lives, and whether the listing is instant bookable. As for the hosts, we collected whether they have multiple listings, whether they have a superhost badge, whether their ID's are verified, and their profile photos. Using their profile photos, we assessed what race each host was most likely to be (White with European origin, White with Middle East/North Africa origin, Asian, Hispanic, African-American, unknown), what their gender was (male, female, two people of the same gender, two people of different genders, unknown), and their age bracket (young: up to 30 years old, middle aged 1: 30-45 years old, middle aged 2: 45-60 years old, old: 60 years old and up, unknown). Each image was assessed by two different research assistants separately and were all cross-checked. Table 2.1 displays the summary statistics and balance tests for the variables used in our analyses.

2.4 Results

Table 2.2 displays our main results. Looking at the first column, we see that the inquiries sent from guest accounts with non-Arab/Muslim sounding names were accepted 84% of the time, whereas Arab/Muslim sounding names received acceptance 75% of the time, roughly a 9 percentage points difference. Moving onto the next two columns, we see that the difference stays at around 9 percentage points when we introduce controls to our regression. Note that the we used a median split for the host reviews. As we move further with our analyses, we will be using this information as a proxy for host experience.

We will now further investigate the persistence of this difference using various specifications in the next section.

2.4.1 Specifications based on Host and Listing Characteristics

Using the additional observational data at hand, we will now investigate potential reasons for the acceptance rate differences.

One could expect that hosts feel more comfortable renting out their property to guests of the same race, i.e. homophily. If this were the main reason behind the acceptance rate differences, we would expect to see higher acceptance rates of Arab/Muslim guests by Arab/Muslim hosts. Table 2.3 displays results for specifications that include not only guest identity indicator, but also a host identity indicator, and their interaction. Column 1 includes all of the observations, whereas columns 2, 3, and 4 repeat the same regression for the male, female, and other hosts (i.e. gender was not identifiable or profile photo had multiple people with mixed gender), respectively. We see that the interaction term is not significantly different than zero in any of the specifications, suggesting that homophily did not play a role. Further, we see that there are behavior differences across male and female hosts. Firstly, although both groups have a negative coefficient in front of the treatment variable, female hosts' coefficient is double the magnitude of male hosts' coefficient. Further, we notice that the interaction term results in a negative coefficient for male hosts and a positive coefficient for female hosts. However, none of these coefficients are significantly different than zero.

We will now move on to investigate whether host-guest physical proximity may have played a role in the discriminatory behavior we have identified against Arab/Muslim guests. One could imagine hosts having preferences against cultural practices of Arabs and Muslims. We group rental listings with respect to whether they were a 'shared property' (e.g. private room in host's apartment) or not (e.g. entire apartment). Column 1 of Table 2.4 shows that the discrimination persists regardless of the unit type.

Next, we will focus on the host professionality type. Although some hosts use Airbnb occasionally to make extra money while for example traveling, others use it more professionally and rent out units regularly. We use two proxied for host professionality: whether or not host ha multiple listings over Airbnb and whether hosts have over 107 reviews (median split). Columns 2 and 3 focus on specifications that control for these proxies respectively. We see that the discriminatory behavior persists at more or less the same level.

Finally, we will consider a specification that controls for the age of the host. Column 4 of Table 2.4 shows that although the significance level drops by 5 percentage points, the discriminatory behavior stays relatively put around the same level.

2.5 Discussion

Online platforms are increasingly dominating the marketplaces and giving us the opportunity to rethink what was not working in the original systems. Discrimination against minority groups has long been documented in many strands of economic transactions and therefore constitutes a natural candidate to focus on in this process. In addition to Edelman, Luca and Svirsky (2017) we have documented that the discrimination on Airbnb exists, not only against African-American guests but also Arab/Muslim identifying guests. As a pioneer in online marketplaces, we believe Airbnb has a great opportunity to become a blueprint for the younger platforms in how to eliminate discrimination with a few substantial changes. We will now talk about the actions Airbnb has taken so far and the actions we believe Airbnb should further take.

2.5.1 Policy Implications

Since this audit study, Airbnb has taken several actions against discrimination. Although small, they now have an 'anti-discrimination' team. One of the changes they brought to the platform was the 'Community Commitment and Nondiscrimination Policy' which is a preemptive official contract they now require users to agree to. In addition, they have announced 'Project Lighthouse' in June 2020, which will run audit studies of similar kind with the help of an independent partner.

Although these are promising steps, we believe there are other more effective actions that Airbnb could implement which would solidify their commitment to the issue. They could conceal the names and photos of the users until after the confirmation of a reservation. In the same vein, they could incentivize their hosts to choose 'Instant Book' option to spread it even further across the platform. Instant Book eliminates the need to get manual approval from a host for the dates a listing is listed as available.

2.5.2 Future Work

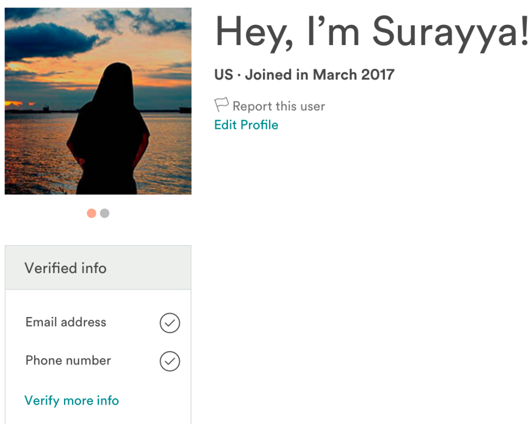
Our paper comes with its own limitations and there are some natural next steps we would like to take for robustness purposes. Firstly, we want to incorporate host names into the racial background identification process along with their profile photos. Next, we would like to incorporate geographical data into our analyses. Further, we would like to make use of the host comments in validating our results that rely on relatively empty guest profiles.

The results suggest that there is room for further exploration. Considering the lack of research regarding behavior towards Arab/Muslim identity, we would like to pursue an extension in which we differentiate between first generation Arab/Muslim identity and a later generation Arab/Muslim identity.

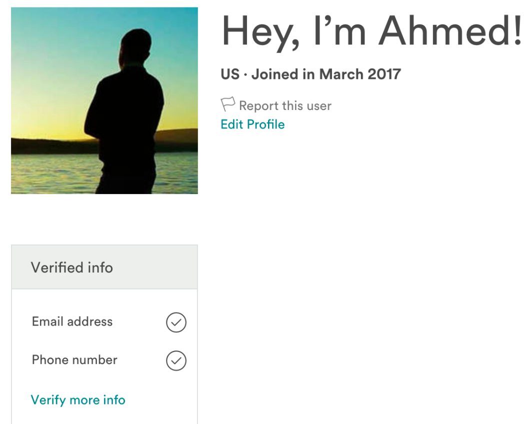
2.6 Figures & Tables

Figure 2.1: Guest Profiles

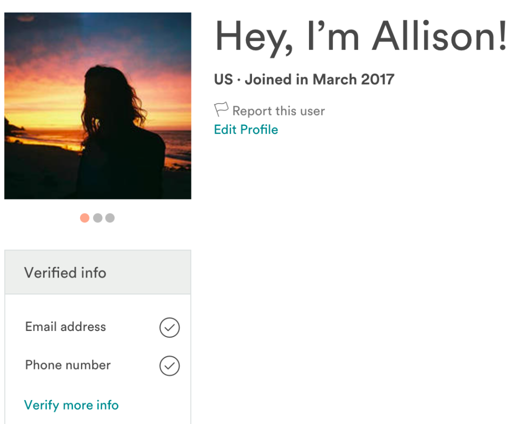
A: Muslim/Arab Female



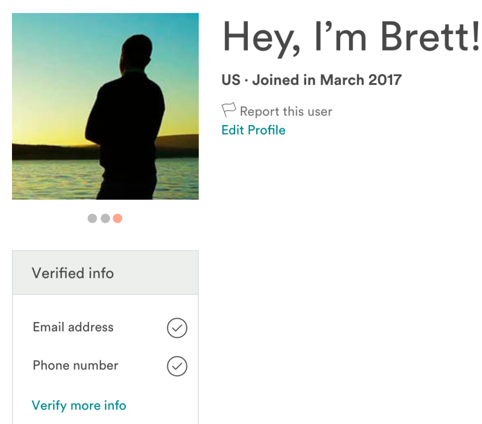
B: Muslim/Arab Male



C: non-Muslim/Arab Female



D: non-Muslim/Arab male



Notes: We standardized photos as much as possible; they are all sunset photos and only include a silhouette of the the upper body from behind. All of our accounts have joined Airbnb within the same month and they all had a verified e-mail address as well as a verified phone number.

Figure 2.2: Inquiry Sent from a Guest Account

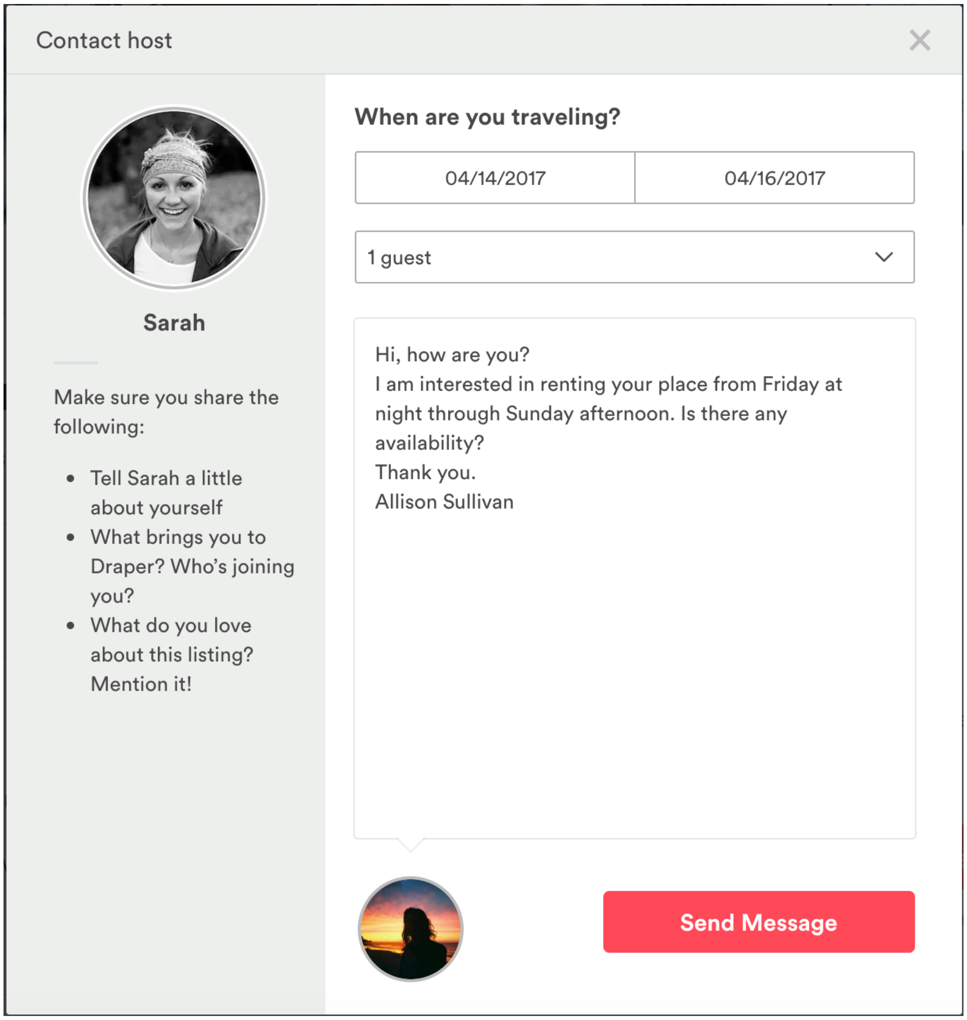


Table 2.1: Summary Statistic and Balance of Treatment Samples

	Obs.	Mean	Std. Dev.	25th %ile	75th %ile	Mean (1)	Mean (2)	p-value
Host is white, European origins	537	0.683	0.466	0	1	0.696	0.669	0.511
Host is white, ME or NA origins	537	0.047	0.211	0	0	0.063	0.028	0.055
Host is Female	537	0.304	0.460	0	1	0.301	0.307	0.879
Host is Male	537	0.246	0.431	0	0	0.248	0.243	0.889
Price ($)	527	105.220	91.701	55	124	103.881	106.7149	0.724
Number of bedrooms	537	1.449	1.0284	1	2	1.483	1.410	0.4177
Number of bathrooms	526	1.486	4.778	1	1.5	1.303	1.689	0.356
Number of reviews	536	107.674	232.755	17	108	84.045	134.598	0.012
Host has multiple listings	537	0.184	0.388	0	0	0.192	0.175	0.613
Listing is instant bookable	537	0.454	0.498	0	1	0.455	0.454	0.993
Host has superhost badge	537	0.428	0.495	0	1	0.416	0.442	0.542
Host has verified ID	537	0.745	0.436	0	1	0.738	0.753	0.687

Notes: (1): Arab/Muslim Accounts, (2):non-Arab/Muslim Accounts.

Table 2.2: The Effect of Identity on Acceptance Rates

	Dependent Variable: 1(Host Accepts)		
Guest is Arab/Muslim	-0.0889**	-0.0895**	-0.0828**
	(0.0380)	(0.0381)	(0.0383)
Host is white, ME or NA origins		0.00408	-0.00338
		(0.0861)	(0.0924)
Host is Male		0.0836**	0.0778**
		(0.0366)	(0.0381)
Host has multiple listings			0.0445
			(0.0502)
Shared Property			0.0617
			(0.0401)
Host has 107+ Reviews			0.0561
			(0.0423)
ln(Price)			0.00505
			(0.0323)
Constant	0.841***	0.820***	0.751***
	(0.0265)	(0.0282)	(0.157)
R-squared	0.012	0.020	0.030
Observations	537	537	527

Notes: A host's response is coded as 1 if, in his reply was either a direct 'Yes' or 'Yes with questions'.
ME: Middle East, NA: North Africa.
Standard errors are clustered by (guest name)*(state) in parenthesis; *** p<0.01, ** p<0.05, * p<0.1

Table 2.3: Acceptance Gap by Host Identity

	(1)	(2)	(3)	(4)
	Dependent Variable: 1(Host Accepts)			
Guest is Arab/Muslim	-0.0980**	-0.0567	-0.111	-0.110**
	(0.0387)	(0.0623)	(0.0781)	(0.0513)
Host is white, ME or NA origins	-0.130	0.121***	-0.327	-0.338
	(0.181)	(0.0434)	(0.360)	(0.363)
Host is white, ME or NA origins * Guest is Arab/Muslim	0.217	-0.0544	0.411	0.360
	(0.204)	(0.124)	(0.390)	(0.425)
Constant	0.844***	0.879***	0.827***	0.838***
	(0.0264)	(0.0434)	(0.0486)	(0.0359)
R-squared	0.015	0.011	0.021	0.021
Observations	537	132	163	242

Notes: (1) All Hosts (2) Male Hosts (3) Female Hosts (4) Other Hosts.
We have some missing property variables as some of the listings were removed between the time we scraped for the host and property ids (to start the experiment) and the time we scraped host and property characteristics using those id's.
(1): Arab/Muslim Accounts, (2):non-Arab/Muslim Accounts.

Table 2.4: Are Host and Listing Characteristics Responsible for the Gap?

	Dependent Variable: 1(Host Accepts)			
Guest is Arab/Muslim	-0.100**	-0.0955**	-0.107**	-0.0794*
	(0.0489)	(0.0415)	(0.0450)	(0.0414)
Shared Accomm.	0.0424			
	(0.0491)			
Shared Accomm. * Guest is Arab/Muslim	0.0414			
	(0.0689)			
Host has multiple listings		0.0279		
		(0.0702)		
Host has multiple listings * Guest is Arab/Muslim		0.0319		
		(0.101)		
Host has 107+ Reviews			0.0229	
			(0.0526)	
Host has 107+ Reviews * Guest is Arab/Muslim			0.0832	
			(0.0896)	
Host is young				-0.0657
				(0.0673)
Host is young * Guest is Arab/Muslim				-0.0540
				(0.0963)
Constant	0.824***	0.836***	0.834***	0.853***
	(0.0353)	(0.0292)	(0.0321)	(0.0279)
R-squared	0.017	0.014	0.019	0.021
Observations	529	537	537	537

Notes: Standard errors are clustered by (guest name)*(state) in parenthesis; *** p<0.01, ** p<0.05, * p<0.1

APPENDIX A

Chapter I Supporting Material

A.1 Appendix

A.1.1 Proofs

Proof of Prediction 1 and 2

When the receiver is strategic, he takes the action that coincides with the prior mean, $\mathbb{E}[t|s]$, as that is the optimizer for his utility function.

When the receiver is naive, he thinks that the distribution of types is uniform over the type support of the signal (truncated prior) and takes the action that is the average of the type support for the disclosed signal, $\mathbb{E}_\chi[t|s]$.

For a given $\chi \in (0,1)$, this leads to different guesses by two types of receivers for signal regions where the correct posterior is not uniform across the type support; i.e. signals $\{1,2,3,4,5\}$ and signals $\{10,11,12,13,14\}$ for the Disadvantaged and Advantaged senders respectively. For these signals, the correct posterior has mass greater on the highest (lowest) type in the type support of each signal for the Disadvantaged (Advantaged) sender group. Therefore, the existence of naive receiver leads to an expected guess that is below (above) the expected type for the Disadvantaged (Advantaged) sender group.

Proof of Prediction 3

In the presence of naive receivers, first note that, regardless of receiver's action upon nondisclosure, sender's best response is a threshold disclosure strategy: $m = \Psi$ if $s < s^*$ and $m = s$ if $s \geq s^*$.

Upon nondisclosure, when the receiver is strategic, his action is equal to $\mathbb{E}[t|s < s^*]$. Whereas, when the receiver is naive, the response to nondisclosure follows the form $\mathbb{E}[t]$

as nondisclosure is available to sender with any signal. So the sender prefers to disclose whenever

$$\chi\,\mathbb{E}_{\chi}[t|s] + (1-\chi)\,\mathbb{E}[t|s] > \chi\,\mathbb{E}[t] + (1-\chi)E[t|s < s^{*}] \tag{A.1}$$

Note that disclosing is not soptimal for $s = 0$. Similarly, nondisclosure is not optimal for signals $s \geq 5$ (i.e., if no signal were disclosed, the expected receiver would be equal to $\mathbb{E}[t] = 7.5$, for any threshold $s^{*} < 15$, the expected receiver guess is strictly less than 7.5) for either sender group. The payoff sender gets from disclosure is strictly increasing. Therefore, at most one signal $s^{*} < 5$, can be indifferent between disclosing and not disclosing in equilibrium.

For the randomizing signal, s^{*}, equation A.1 should hold in equality. When s^{*} goes to 0, the LHS of equation goes to 5 and the RHS goes to $5+\chi 2.5$ for both sender groups. When s^{*} goes to 5 the LHS goes to $9-\chi 1.5$ and the RHS goes to $7.26-\chi 0.24$ for the (Disadvantaged) sender group with right skewed error distribution. When s^{*} goes to 5 the LHS goes to 7.5 and the RHS goes to $6.6-\chi 0.9$ for the (Advantaged) sender group with left skewed error distribution. Hence, there exists a solution to $s^{*} \in (0, 5)$ for both sender groups.

Next, notice that, for any $\chi \in (0, 1)$, the LHS increases by a greater magnitude than RHS in s for both sender groups. Hence, the solution s^{*} is unique.

Proof of Prediction 4

Note that the naive receiver does not differentiate between two sender groups. This results follow from the relative relationship of the $\mathbb{E}[t|s']$ and $\mathbb{E}[t|s < s']$ that's increasing in s' for each sender group. Notice that the LHS of equation A.1 increases in s steadily for the (Advantaged) sender group with left skewed error distribution, by 0.5. In contrast, both the LHS and the RHS of the equation increase Notice that the equality between the LHS and the RHS is satisfied sooner and more frequently for the (Advantaged) sender group with left skewed error distribution.

Specifically, Table A.1 and display that for every $\chi \in (0, 1)$, the (Advantaged) sender group with left skewed error distribution withholds a (weakly) greater number of signals than the other sender group.

Proof of Prediction 5

Under voluntary disclosure, senders with low signals are able to enjoy the benefits of nondisclosure, which results in higher receiver expected guess. Such senders in the (Disadvantaged) group with right skewed error distribution were subject to underestimation in mandatory disclosure regime, whereas those in the other (Advantaged) group were getting the Bayesian expected guess. Further, such senders make up a greater fraction of their group for the (Disadvantaged) group with right skewed errors compared to the other (Advantaged) group. Therefore, the (Disadvantaged) sender group gets to increase the expected receiver guess more than the other (Advantaged) by the transition into voluntary disclosure regime.

A.1.2 Identification in Treatment M and Treatment V

Identification in Treatment M

Table A.2 shows benchmark rational and naive receiver guesses for signal regions. We see that the observed mean receiver guess differs from the rational benchmark in the direction our model predicts for the signal region 1-5. Further, our experimental results also verify that the receive guesses are not significantly different than the rational benchmark for the signal region 6-14, for which the rational and the naive benchmarks overlap.

At an individual level, 63% of receiver subjects have higher absolute average deviation from the rational benchmark in the shaded region than in the signal region 6-14 (Wilcoxon signed-rank test: $p = 0.0241$)

Similarly, table A.3 shows benchmark rational and naive receiver guesses for signal regions. We see that the observed mean receiver guess differs from the rational benchmark in the direction our model predicts for the signal region 10-14. Further, our experimental results also verify that the receive guesses are not significantly different than the rational benchmark for the signal region 1-9, for which the rational and the naive benchmarks overlap.

At an individual level, 65% of receiver subjects have higher absolute average deviation from the rational benchmark in the shaded region than in the signal region 1-9 (Wilcoxon signed-rank test: $p = 0.0311$)

Identification in Treatment V

Table A.4 separates the nondisclosure region (signals 0-2) and the shaded region (signals 3-5) from the rest of the signals. These are the regions where we expect to see deviation from the rational benchmark. More specifically, our model predicts overestimation and underestimation of expected receiver type in the non disclosure region and in the shaded signal region, respectively. We see that indeed the receiver subjects in Treatment V behaved inline with our theoretical predictions.

At an individual level, of the receivers who have faced with nondisclosure and disclosed signals 6-14 (68%) 59% of receiver subjects have higher absolute average deviation from the rational benchmark in the nondisclosure signal region 0-2 than in the signal region 6-

14 (Wilcoxon signed-rank test: $p = 0.0309$). Similarly, of the receivers who have faced both (72%), 69% have higher absolute average deviation from the rational benchmark in the shaded signal region 3-5 than in the signal region 6-14 (Wilcoxon signed-rank test: $p = 0.0098$).

Interestingly, of the receiver subjects who were faced with nondisclosure and signals from the identification region (49%), 57% have higher absolute average deviation from the rational benchmark in the shaded signal region 3-5 than in the nondisclosure signal region (Wilcoxon signed-rank test: $p = 0.0204$). This implies that, although we observe receiver behavior aligned with our model's predictions, the share of naive receivers are greater for the identification region signals (disclosed) than for the nondisclosure region.

Further, we compare absolute deviation from Bayesian benchmark for the shaded identification region across Treatment M and Treatment V, for the (disadvantaged) senders with right skewed error distribution. Regressing absolute deviation on a dummy variable representing voluntary disclosure, we find that voluntary disclosure increases deviations by 0.22 ($p = 0.0196$).

Similarly, table A.5 separates the nondisclosure region (signals 0-3) and the shaded region (signals 10-14) from the rest of the signals. These are the regions where we expect to see deviation from the rational benchmark. More specifically, our model predicts overestimation of expected receiver type in both regions. We see that the experimental data verifies this prediction.

At an individual level, of the receivers who have faced with nondisclosure region and disclosed signals 4-9 (45%) 60% of receiver subjects have higher absolute average deviation from the rational benchmark in the nondisclosure signal region 0-3 than in the signal region 4-9 (Wilcoxon signed-rank test: $p = 0.009$). Similarly, of the receivers who have faced both (67%), 78% have higher absolute average deviation from the rational benchmark in the shaded signal region 10-14 than in the signal region 4-9 (Wilcoxon signed-rank test: $p = 0.0118$).

Interestingly, of the receiver subjects who were faced with nondisclosure and signals from the identification region 10-14 (35%), 63% have higher absolute average deviation from the rational benchmark in the shaded signal region 10-14 than in the nondisclosure signal region (Wilcoxon signed-rank test: $p = 0.0361$). This implies that, although we observe receiver behavior aligned with our model's predictions, the share of naive receivers are greater for the identification region signals (disclosed) than for the nondisclosure region.

Further, we compare absolute deviation from Bayesian benchmark for the shaded identification region across Treatment M and Treatment V, for the (advantaged) senders with left skewed error distribution. Regressing absolute deviation on a dummy variable representing voluntary disclosure, we find that voluntary disclosure increases deviations by 0.25 ($p = 0.0294$).

A.1.3 Instructions

Below are the instructions from wave 3 treatments. The parts that differ across W3.1 and W3.2 are marked in parenthesis with the corresponding information transmission regime name as *Mandatory Disclosure* or *Voluntary Disclosure*.

Instructions. We will start with "Guessing the Secret Number" Game. You will play 15 rounds of this game.

This game involves two roles: an Observer and a Guesser. Roles are randomly determined before the game begins and will be fixed throughout the entire game. You are equally likely to be assigned to either role. You will see your assigned role when the game begins.

Each round, the computer will randomly pair you with a participant in the other role. You will NOT know which participant you are paired up with.

Note also that, before the game begins, participants are put into 2 different Zoom breakout rooms based on their randomly assigned roles. Therefore, you will never be paired with a participant from your Zoom breakout room.

(Mandatory Disclosure) Each round will proceed as follows:

- For each Observer-Guesser pair the computer will send a "secret number" and a "color number" to the Observer.
  - The sum of the secret number and the color number generates a "message".
- The Guesser will only see the message. Note that the Guesser will never see the secret number or the color number.
- The Guesser will then guess the value of the secret number, concluding the round.

(Voluntary Disclosure) Each round will proceed as follows:

- For each Observer-Guesser pair the computer will send a "secret number" and a "color number" to the Observer.

- The sum of the secret number and the color number generates a ”message”.

- The Observer will then decide whether or not to send the message to the Guesser. Note that the Guesser will never see the secret number or the color number.
  - If message is sent: the Guesser will only see the message.
  - If message is not sent: the Guesser will only see that the Observer did not send a message.
- The Guesser will then guess the value of the secret number, concluding the round.

Secret Numbers and Color Numbers

The secret number is one of the following: 5, 6, 7, 8, 9, 10. The color number can be zero, negative or positive. To differentiate it, we put parentheses around the (color number). Specifically, it is one of the following: (-5), (-4), (-3), (-2), (-1), (0), (+1), (+2), (+3), (+4), (+5).

Messages

Message = secret number + (color number)

Table below (see figure **??**) shows the list of all possible Messages. It also shows the secret numbers and the corresponding (color numbers) that could generate each Message.

For example:
Message 1: secret number could be either be 5 or 6
Message 7: secret number could be 5, 6, 7, 8, 9, or 10
Message 12: secret number could be 7, 8, 9, or 10

Above table will also be available on your decision screens.

How is the secret number determined each round?

Each round, the computer randomly picks the secret number such that 5, 6, 7, 8, 9, 10 are each equally likely to be selected.

How is the (color number) determined each round?

The computer will spin a wheel filled with (color numbers): (-5), (-4), (-3), (-2), (-1), (0), (+1), (+2), (+3), (+4), (+5).

Each Observer belongs to one of the two color groups. Each color group has its own wheel:

Sky color group

Forest color group

Note that the color groups will be fixed throughout the experiment and each round the Guessers will see the color group of the Observer they are matched with.

Each round and for every Observer, the computer will spin the corresponding wheel to determine the (color number). (See figure A.4)

Sky color group wheel

- half of the wheel is filled with: (-5)
- other half of the wheel is filled with: (-4), (-3), (-2), (-1), (0), (+1), (+2), (+3), (+4), (+5)

Forest color group wheel

- half of the wheel is filled with: (+5)
- other half of the half of the wheel is filled with: (-5), (-4), (-3), (-2), (-1), (0), (+1), (+2), (+3), (+4)

Above information will also be available on your decision screens.

Guesses

The Guessers can enter their GUESS of the secret number in 0.5 increments: 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 10.

How are payoffs determined?

The payoffs depend on the GUESS such that:

- the Observer earns <u>more when the GUESS is higher</u>
- the Guesser earns <u>more when the GUESS is closer to the secret number</u>

The payoffs of this game are defined in terms of points as displayed in the tables below (Figure A.5).

For example, suppose GUESS is 6.5

- the Observer gets 23 points

  (regardless of the secret number)

- if the secret number is 6, the Guesser gets 104 points
- if the secret number is 10, the Guesser gets 23 points

Payoff tables will also be available on your decision screens.

Your final earning from this part ("Guessing the Secret Number") will be calculated by converting the TOTAL points you collect throughout the 15 rounds into US dollars. The rate at which points are worth US dollars will be shown on the screen before the game begins.

Note that when the game begins, you will have up to 2.5 minutes in the first 10 rounds and up to 1.5 minutes in the last 5 rounds to proceed.

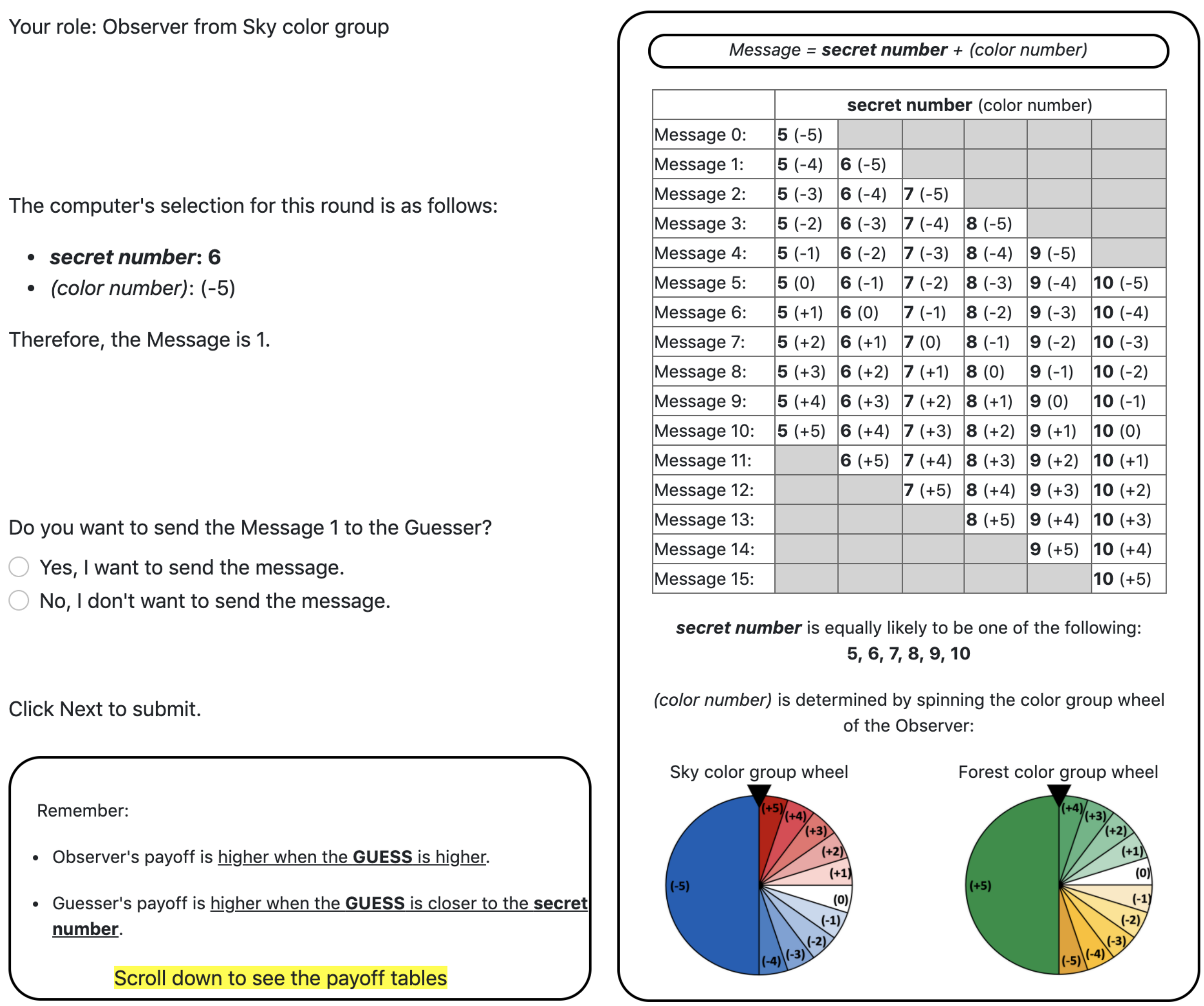
(Mandatory Disclosure) In the case that the Guesser doesn't submit their choice within the allotted time, the computer will submit a random choice.

(Voluntary Disclosure) In the case that you don't submit your choice within the allotted time, the computer will submit a random choice.

Before we begin, let's briefly check our understanding with some examples...

A.1.4 Appendix: Figures & Tables

Figure A.1: Sender decision screen from Treatment V



Notes: The figure shows an example sender decision screen for the Disadvantaged sender group in Treatment V. Here, the senders were provided with the two options in random order.

Figure A.2: Receiver decision screen from Treatment V

Your role: Guesser

The Observer from Sky color group -

did Not send a Message.

Please enter your **GUESS** for the **secret number**:

- ○ 5
- ○ 5.5
- ○ 6
- ○ 6.5
- ○ 7
- ○ 7.5
- ○ 8
- ○ 8.5
- ○ 9
- ○ 9.5
- ○ 10

Click Next to submit.

Remember:

- Observer's payoff is higher when the **GUESS** is higher.
- Guesser's payoff is higher when the **GUESS** is closer to the **secret number**.

Scroll down to see the payoff tables

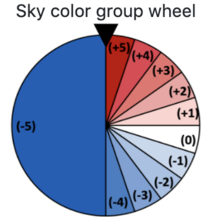
Message = ***secret number*** + *(color number)*

	secret number (color number)					
Message 0:	**5** (-5)					
Message 1:	**5** (-4)	**6** (-5)				
Message 2:	**5** (-3)	**6** (-4)	**7** (-5)			
Message 3:	**5** (-2)	**6** (-3)	**7** (-4)	**8** (-5)		
Message 4:	**5** (-1)	**6** (-2)	**7** (-3)	**8** (-4)	**9** (-5)	
Message 5:	**5** (0)	**6** (-1)	**7** (-2)	**8** (-3)	**9** (-4)	**10** (-5)
Message 6:	**5** (+1)	**6** (0)	**7** (-1)	**8** (-2)	**9** (-3)	**10** (-4)
Message 7:	**5** (+2)	**6** (+1)	**7** (0)	**8** (-1)	**9** (-2)	**10** (-3)
Message 8:	**5** (+3)	**6** (+2)	**7** (+1)	**8** (0)	**9** (-1)	**10** (-2)
Message 9:	**5** (+4)	**6** (+3)	**7** (+2)	**8** (+1)	**9** (0)	**10** (-1)
Message 10:	**5** (+5)	**6** (+4)	**7** (+3)	**8** (+2)	**9** (+1)	**10** (0)
Message 11:		**6** (+5)	**7** (+4)	**8** (+3)	**9** (+2)	**10** (+1)
Message 12:			**7** (+5)	**8** (+4)	**9** (+3)	**10** (+2)
Message 13:				**8** (+5)	**9** (+4)	**10** (+3)
Message 14:					**9** (+5)	**10** (+4)
Message 15:						**10** (+5)

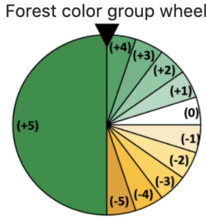
secret number is equally likely to be one of the following:
5, 6, 7, 8, 9, 10

(color number) is determined by spinning the color group wheel of the Observer:

Sky color group wheel



Forest color group wheel



Notes: The figure shows shows an example receiver decision screen faced with nondisclosure by the Disadvantaged sender in Treatment V.

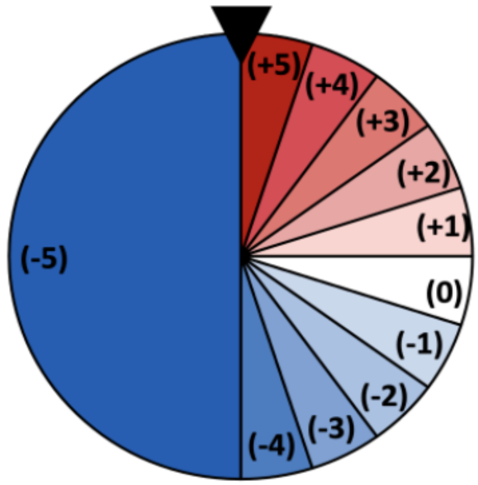
Figure A.3: Message (Signal) Table

	secret number (color number)					
Message 0:	**5** (-5)					
Message 1:	**5** (-4)	**6** (-5)				
Message 2:	**5** (-3)	**6** (-4)	**7** (-5)			
Message 3:	**5** (-2)	**6** (-3)	**7** (-4)	**8** (-5)		
Message 4:	**5** (-1)	**6** (-2)	**7** (-3)	**8** (-4)	**9** (-5)	
Message 5:	**5** (0)	**6** (-1)	**7** (-2)	**8** (-3)	**9** (-4)	**10** (-5)
Message 6:	**5** (+1)	**6** (0)	**7** (-1)	**8** (-2)	**9** (-3)	**10** (-4)
Message 7:	**5** (+2)	**6** (+1)	**7** (0)	**8** (-1)	**9** (-2)	**10** (-3)
Message 8:	**5** (+3)	**6** (+2)	**7** (+1)	**8** (0)	**9** (-1)	**10** (-2)
Message 9:	**5** (+4)	**6** (+3)	**7** (+2)	**8** (+1)	**9** (0)	**10** (-1)
Message 10:	**5** (+5)	**6** (+4)	**7** (+3)	**8** (+2)	**9** (+1)	**10** (0)
Message 11:		**6** (+5)	**7** (+4)	**8** (+3)	**9** (+2)	**10** (+1)
Message 12:			**7** (+5)	**8** (+4)	**9** (+3)	**10** (+2)
Message 13:				**8** (+5)	**9** (+4)	**10** (+3)
Message 14:					**9** (+5)	**10** (+4)
Message 15:						**10** (+5)

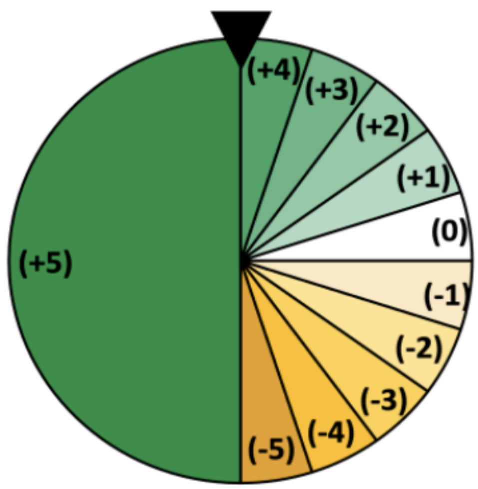
Notes: The figure depicts the table of possible messages (signals) in the game. The table is marked with examples that were used to describe how to read it during the instructions. This table was also available on decision screens throughout the sender-receiver game.

Figure A.4: Color Group Wheels

A: Sky color group wheel



B: Forest color group wheel



Notes: The figure illustrates the likelihoods of color numbers (error terms) for each sender (color) group. We used a wheel visual and a gradient color scheme for ease of exposition. Further, we have picked different and contrasting colors that fit the color group names. Panel A depicts the color number likelihoods for Sky color group and panel B depicts it for the Forest color group.

Figure A.5: Payoff Tables

OBSERVER POINT	GUESS:										
PAYOFFS	**5**	**5.5**	**6**	**6.5**	**7**	**7.5**	**8**	**8.5**	**9**	**9.5**	**10**
secret number 5	-33	-13	6	23	40	56	70	84	95	104	110
secret number 6	-33	-13	6	23	40	56	70	84	95	104	110
secret number 7	-33	-13	6	23	40	56	70	84	95	104	110
secret number 8	-33	-13	6	23	40	56	70	84	95	104	110
secret number 9	-33	-13	6	23	40	56	70	84	95	104	110
secret number 10	-33	-13	6	23	40	56	70	84	95	104	110

GUESSER POINT	GUESS:										
PAYOFFS	**5**	**5.5**	**6**	**6.5**	**7**	**7.5**	**8**	**8.5**	**9**	**9.5**	**10**
secret number 5	110	104	95	84	70	56	40	23	6	-13	-33
secret number 6	95	104	110	104	95	84	70	56	40	23	6
secret number 7	70	84	95	104	110	104	95	84	70	56	40
secret number 8	40	56	70	84	95	104	110	104	95	84	70
secret number 9	6	23	40	56	70	84	95	104	110	104	95
secret number 10	-33	-13	6	23	40	56	70	84	95	104	110

Notes: Payoff tables were displayed at the bottom of each decision screen. Payoff table of the sender on the left. Payoff table of the receiver on the right.

Table A.1: Threshold signals for given χ values.

Threshold Signal	Disadvantaged Senders (χ)	Advantaged Senders (χ)
0	0	0
1	0.18	0.08
2	0.35	0.18
3	0.53	0.33
4	0.74	0.57
5	1	1

Table A.2: Receiver naivete by signal regions of of Treatment M, for the Disadvantaged sender group. Wilcoxon signed-rank test results compare the benchmarks to the observed values.

Signals	Bayesian Belief (Benchmark)	Naive Belief (Benchmark)	Mean Belief (Observed)	Sign-rank Test (p-value)	Observations
0	5	5	5		
1-5	7.62	6.58	7.21	0.0000, 0.0000	275
6-14	7.84	7.84	7.78	0.419	168
15	10	10	9.95		

Table A.3: Receiver naivete by signal regions of of Treatment M, for the Advantaged sender group. Wilcoxon signed-rank test results compare the benchmarks to the observed values.

Signals	Bayesian Belief (Benchmark)	Naive Belief (Benchmark)	Mean Belief (Observed)	Sign-rank Test (p-value)	Observations
0	5	5	5		
1-9	7.16	7.16	7.13	0.2394	159
10-14	7.39	8.42	8.02	0.0000, 0.0000	281
15	10	10	9.95		

Table A.4: Receiver naivete by signal regions of Treatment V, for the Disadvantaged sender group. Wilcoxon signed-rank test results compare the benchmarks to the observed values.

Signals	Bayesian Belief (Benchmark)	Naive Belief (Benchmark)	Mean Belief (Observed)	Sign-rank Test (p-value)	Observations
0-2	5.90	7.5	6.62	0.003, 0.014	136
3-5	7.62	6.58	7.33	0.0000, 0.0000	131
6-14	7.8	7.84	8.06	0.0023	143
15	10	10	9.95		

Table A.5: Receiver naivete by signal regions of Treatment V, for the Advantaged sender group. Wilcoxon signed-rank test results compare the benchmarks to the observed values.

Signals	Bayesian Belief (Benchmark)	Naive Belief (Benchmark)	Mean Belief (Observed)	Sign-rank Test (p-value)	Observations
0-3	6	7.5	6.7	0.000, 0.0000	45
4-9	7.42	7.42	?	0.331	153
10-14	7.39	8.42	8.02	0.0000, 0.0000	231
15	10	10	9.95		

APPENDIX B

Chapter II Supporting Material

B.1 Appendix

B.1.1 Appendix: Figures & Tables

Table B.1: Survey Results

Arab/Muslim		non-Arab/Muslim	
Fatima Haddad	0.94	Allison Sullivan	0.98
Rabia Ahmed	0.96	Meredith O'Brien	0.99
Salma Ahmed	0.96	Laurie Ryan	0.98
Aamina Haddad	0.96	Anne Murphy	0.98
Surayya Mohammed	0.97	Kristen Sullivan	0.98
Abdul Aziz Haddad	0.93	Brad Walsh	0.99
Ramadan Ibrahim	0.94	Todd McCarthy	0.99
Hussein Bilal	0.97	Brent Baker	0.98
Hasan Ahmed	0.98	Greg O'Brien	0.99
Ahmed Ramadan	0.96	Brett Walsh	0.99

Notes: This table tabulates the different names used in the experiment and their identifiability. The numbers report the probability that the name was picked as Arab/Muslim (or White with European origin) in an independent survey among 102 students.
Sample size = 48

Table B.2: Effect of Guest Identity by State (IL, NJ, NM, OR)

	Dependent Variable: 1(Host Accepts)			
	(IL)	(NJ)	(NM)	(OR)
Guest is Arab/Muslim	0.0152	0.0238	-0.323***	-0.0311
	(0.174)	(0.138)	(0.101)	(0.0886)
Constant	0.682***	0.810***	1***	0.895***
	(0.149)	(0.104)	(0)	(0.0489)
R-squared	0.000	0.001	0.172	0.002
Observations	55	39	55	41

Table B.3: Effect of Guest Identity by State (FL, AZ, MO, SC)

	Dependent Variable: 1(Host Accepts)			
	(FL)	(AZ)	(MO)	(SC)
Guest is Arab/Muslim	-0.260*	0.131	-0.323***	-0.0548
	(0.132)	(0.103)	(0.102)	(0.159)
Constant	0.950***	0.792***	0.793***	0.792***
	(0.0513)	(0.0941)	(0.0618)	(0.0798)
R-squared	0.101	0.036	0.110	0.004
Observations	49	50	46	43

Table B.4: Effect of Guest Identity by State (UT, MT, MS, OK)

	Dependent Variable: 1(Host Accepts)			
	(UT)	(MT)	(MS)	(OK)
Guest is Arab/Muslim	-0.123	-0.174***	0.0524	0.0128
	(0.112)	(0.0509)	(0.153)	(0.104)
Constant	0.923***	1***	0.625***	0.870***
	(0.0784)	(0)	(0.147)	(0.0666)
R-squared	0.028	0.080	0.003	0.000
Observations	33	39	47	40

Table B.5: Instant Book Option, Superhost Badge and Verified ID)

	Dependent Variable: 1(Host Accepts)		
Guest is Arab/Muslim	-0.0796	-0.0312	-0.000860
	(0.0542)	(0.0483)	(0.0768)
Listing is instant bookable	0.179***		
	(0.0449)		
Listing is instant bookable * Guest is Arab/Muslim	-0.0205		
	(0.0680)		
Superhost		0.124***	
		(0.0422)	
Superhost * Guest is Arab/Muslim		-0.131**	
		(0.0655)	
Host has verified ID			0.0882
			(0.0665)
Host has verified ID * Guest is Arab/Muslim			-0.117
			(0.0842)
Constant	0.759***	0.786***	0.774***
	(0.0389)	(0.0368)	(0.0605)
R-squared	0.055	0.023	0.017
Observations	537	537	537

Notes: Standard errors are clustered by (guest name)*(state) in parenthesis; *** p<0.01, ** p<0.05, * p<0.1

BIBLIOGRAPHY

Agan, Amanda, and Sonja Starr. 2018. "Ban the box, criminal records, and racial discrimination: A field experiment." *The Quarterly Journal of Economics*, 133(1): 191–235.

Agan, Amanda, Bo Cowgill, and Laura Katherine Gee. 2020. "Do workers comply with salary history bans? a survey on voluntary disclosure, adverse selection, and unraveling." Vol. 110, 215–19.

Åslund, Olof, and Oskar Nordströum Skans. 2012. "Do anonymous job application procedures level the playing field?" *ILR Review*, 65(1): 82–107.

Ballance, Joshua, Robert Clifford, and Daniel Shoag. 2020. ""No more credit score": Employer credit check bans and signal substitution." *Labour Economics*, 63: 101769.

Barach, Moshe A, and John J Horton. 2021. "How do employers use compensation history? Evidence from a field experiment." *Journal of Labor Economics*, 39(1): 193–218.

Bartik, Alexander Wickman, and Scott Thomas Nelson. 2016. *Credit reports as resumes: The incidence of pre-employment credit screening.* Massachusetts Institute of Technology, Department of Economics.

Behaghel, Luc, Bruno Crépon, and Thomas Le Barbanchon. 2015. "Unintended effects of anonymous resumes." *American Economic Journal: Applied Economics*, 7(3): 1–27.

Bergman, Oscar, Tore Ellingsen, Magnus Johannesson, and Cicek Svensson. 2010. "Anchoring and cognitive ability." *Economics Letters*, 107(1): 66–68.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." *American economic review*, 94(4): 991–1013.

Bessen, James E, Chen Meng, and Erich Denk. 2020. "Perpetuating inequality: What salary history bans reveal about wages." *Available at SSRN 3628729*.

Blackaby, David, Alison L Booth, and Jeff Frank. 2005. "Outside offers and the gender pay gap: Empirical evidence from the UK academic labour market." *The Economic Journal*, 115(501): F81–F107.

Blau, Francine D, and Lawrence M Kahn. 2020. "The gender pay gap: Have women gone as far as they can?" In *Inequality in the United States: A Reader*. 345–362. Routledge.

Chen, Daniel L, Martin Schonger, and Chris Wickens. 2016. "oTree—An open-source platform for laboratory, online, and field experiments." *Journal of Behavioral and Experimental Finance*, 9: 88–97.

Cullen, Zoë B, and Bobak Pakzad-Hurson. 2019. "Equilibrium effects of pay transparency in a simple labor market." 193–193.

Deversi, Marvin, Alessandro Ispano, and Peter Schwardmann. 2018. "Spin doctors: A model and an experimental investigation of vague disclosure." *Available at SSRN 3275418*.

Dohmen, Thomas, and Armin Falk. 2011. "Performance pay and multidimensional sorting: Productivity, preferences, and gender." *American economic review*, 101(2): 556–90.

Doleac, Jennifer L, and Benjamin Hansen. 2020. "The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden." *Journal of Labor Economics*, 38(2): 321–374.

Dranove, David, and Ginger Zhe Jin. 2010. "Quality disclosure and certification: Theory and practice." *Journal of economic literature*, 48(4): 935–63.

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

Edelman, Benjamin, Michael Luca, and Dan Svirsky. 2017. "Racial discrimination in the sharing economy: Evidence from a field experiment." *American economic journal: applied economics*, 9(2): 1–22.

Eyster, Erik, and Matthew Rabin. 2005. "Cursed equilibrium." *Econometrica*, 73(5): 1623–1672.

Frederick, Shane. 2005. "Cognitive reflection and decision making." *Journal of Economic perspectives*, 19(4): 25–42.

Goldin, Claudia. 1990. *Understanding the gender gap: An economic history of American women.* National Bureau of Economic Research.

Goldin, Claudia. 2015. "Hours flexibility and the gender gap in pay." *Center for American Progress.*

Grossman, Sanford J. 1981. "The informational role of warranties and private disclosure about product quality." *The Journal of Law and Economics*, 24(3): 461–483.

Grossman, Sanford J, and Oliver D Hart. 1980. "Disclosure laws and takeover bids." *The Journal of Finance*, 35(2): 323–334.

Hagenbach, Jeanne, and Frédéric Koessler. 2017. "Simple versus rich language in disclosure games." *Review of Economic Design*, 21(3): 163–175.

Hagenbach, Jeanne, Frédéric Koessler, and Eduardo Perez-Richet. 2014. "Certifiable Pre-Play Communication: Full Disclosure." *Econometrica*, 82(3): 1093–1131.

Hall, Robert E, and Alan B Krueger. 2012. "Evidence on the incidence of wage posting, wage bargaining, and on-the-job search." *American Economic Journal: Macroeconomics*, 4(4): 56–67.

Hansen, Benjamin, and Drew McNichols. 2020. "Information and the Persistence of the Gender Wage Gap: Early Evidence from California's Salary History Ban." National Bureau of Economic Research.

Holt, Charles A, and Susan K Laury. 2002. "Risk aversion and incentive effects." *American economic review*, 92(5): 1644–1655.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin. 2015. "Is no news (perceived as) bad news? An experimental investigation of information disclosure." National Bureau of Economic Research.

Juhn, Chinhui, and Kristin McCue. 2017. "Specialization then and now: Marriage, children, and the gender earnings gap across cohorts." *Journal of Economic Perspectives*, 31(1): 183–204.

Khanna, Shantanu. 2020. "Salary history bans and wage bargaining: Experimental evidence." *Labour Economics*, 65: 101853.

Krupka, Erin L, and Roberto A Weber. 2013. "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association*, 11(3): 495–524.

Mathios, Alan D. 2000. "The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market." *The Journal of Law and Economics*, 43(2): 651–678.

Mazei, Jens, Joachim Hüffmeier, Philipp Alexander Freund, Alice F Stuhlmacher, Lena Bilke, and Guido Hertel. 2015. "A meta-analysis on gender differences in negotiation outcomes and their moderators." *Psychological bulletin*, 141(1): 85.

Meli, Jeffrey, and James C Spindler. 2019. "Salary history bans and gender discrimination." *U of Texas Law, Law and Econ Research Paper*, , (E587).

Milgrom, Paul, and John Roberts. 1986. "Price and advertising signals of product quality." *Journal of political economy*, 94(4): 796–821.

Milgrom, Paul R. 1981. "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics*, 380–391.

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

Oechssler, Jörg, Andreas Roider, and Patrick W Schmitz. 2009. "Cognitive abilities and behavioral biases." *Journal of Economic Behavior & Organization*, 72(1): 147–152.

Ponthieux, Sophie, and Dominique Meurs. 2015. "Gender inequality." In *Handbook of income distribution*. Vol. 2, 981–1146. Elsevier.

Rozada, Martín González, Eduardo Levy Yeyati, et al. 2018. "Do women ask for lower salaries? the supply side of the gender pay gap." Universidad Torcuato Di Tella.

Sinha, Sourav. 2019. "Salary history ban: Gender pay gap and spillover effects." *Available at SSRN 3458194*.

Solnick, Sara J, and Maurice E Schweitzer. 1999. "The influence of physical attractiveness and gender on ultimatum game decisions." *Organizational behavior and human decision processes*, 79(3): 199–215.

Viscusi, W Kip. 1978. "A note on" lemons" markets with quality certification." *The Bell Journal of Economics*, 277–279.

Wozniak, Abigail. 2015. "Discrimination and the effects of drug testing on black employment." *Review of Economics and Statistics*, 97(3): 548–566.