

NBER WORKING PAPER SERIES

TOWARD AN UNDERSTANDING OF WHY PEOPLE DISCRIMINATE:  
EVIDENCE FROM A SERIES OF NATURAL FIELD EXPERIMENTS

Uri Gneezy  
John List  
Michael K. Price

Working Paper 17855  
<http://www.nber.org/papers/w17855>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
February 2012

We thank Gary Becker, James Heckman, and Steven Levitt for their discussions throughout the exploration process. Seminar participants at the 2012 ASSA meetings in Chicago also provided useful insights. Moshe Hoffman provided able research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Uri Gneezy, John List, and Michael K. Price. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Toward an Understanding of Why People Discriminate: Evidence from a Series of Natural Field Experiments

Uri Gneezy, John List, and Michael K. Price

NBER Working Paper No. 17855

February 2012

JEL No. C93,J71

**ABSTRACT**

Social scientists have presented evidence that suggests discrimination is ubiquitous: women, nonwhites, and the elderly have been found to be the target of discriminatory behavior across several labor and product markets. Scholars have been less successful at pinpointing the underlying motives for such discriminatory patterns. We employ a series of field experiments across several market and agent types to examine the nature and extent of discrimination. Our exploration includes examining discrimination based on gender, age, sexual orientation, race, and disability. Using data from more than 3000 individual transactions, we find evidence of discrimination in each market. Interestingly, we find that when the discriminator believes the object of discrimination is controllable, any observed discrimination is motivated by animus. When the object of discrimination is not due to choice, the evidence suggests that statistical discrimination is the underlying reason for the disparate behavior.

Uri Gneezy  
Rady School of Management  
University of California - San Diego  
Otterson Hall, Room 4S136  
9500 Gilman Drive #0553  
La Jolla, CA 92093-0553  
ugneezy@ucsd.edu

Michael K. Price  
Department of Economics  
University of Tennessee  
515 Stokely Management Center  
Knoxville, TN 27996  
and NBER  
mprice21@utk.edu

John List  
Department of Economics  
University of Chicago  
1126 East 59th  
Chicago, IL 60637  
and NBER  
jlist@uchicago.edu

## I. Introduction

Philosophers as far removed as Arcesilaus, Heraclitus, and Plato have scribed of injustice and extolled upon the virtues of removing it for the betterment of society. Perhaps taking a lead from these scholars, social scientists have extensively studied gender, race, and age discrimination in the marketplace. The empirical findings have provided suggestive evidence that systematic discrimination against women, nonwhites, and the elderly is prevalent in modern societies.<sup>1</sup>

While empirical studies have served to provide an empirical foundation measuring the extent of discrimination in the marketplace, they have been less helpful in distinguishing the nature of discrimination. This is significant because before one can begin to discuss social policies to address differential treatment of individuals based on ascriptive characteristics, it is critical to identify *both* the existence and the underlying source of discrimination.<sup>2</sup>

The contribution of this study is twofold. First, it uses a series of field experiments to measure discrimination based on gender, age, sexual orientation, race, and disability across several distinct product markets as well as information markets. For example, in one setting we explore whether the disabled face discrimination in the market for repairing their automobile. In another, we examine whether sexual orientation and race affect the price quotes an individual receives in the new car market. Our general approach is broad enough to have validity outside of typical product markets, and we highlight this fact by analyzing whether discrimination can be found within information markets.

---

<sup>1</sup> Riach and Rich (2002) provide a recent survey of the audit study literature; List and Rasul (2011) provide a recent summary of the regression based studies.

<sup>2</sup> Some headway has been made in these regards: Altonji and Pierret (2001) use data on wages and employment tenure to cleverly investigate racial discrimination—they report little evidence of statistical discrimination. Fershtman and Gneezy (2001) use experiments to study discrimination in the Israeli society, concluding that its source is statistical discrimination based on erroneous stereotypes. List (2004) utilizes lab and field experiments to examine discrimination across gender, age, and race in the sportscard market.

A second contribution of this study is to provide simple empirical frameworks for parsing the nature of discrimination observed. Beyond measuring the levels of discrimination, a series of field treatments permits us to determine the underlying motivations for the observed discrimination. For example, in the automobile repair market, we find that the disabled receive offers that are 30 percent higher than the offers received by the abled. One possible reason for this disparate treatment is search cost differences—one would expect search to be more costly for the disabled. Under this scenario, such agents might search less and repairmen might capitalize on that fact. To test this conjecture, and identify the underlying nature of discrimination, we employ a complementary field experiment where during the offer process *all* agents note that “I am getting a few price quotes.” Upon having our agents make this simple statement, we find that offer distributions become isomorphic: the abled and disabled receive statistically indistinguishable offers. Combining this result with data from the first experiment provides clean evidence that the mechanics are engaging in statistical discrimination based on perceived search differences.<sup>3</sup>

We make use of a similar approach to explore the nature and extent of discrimination in other markets. In sum, data from more than 3,290 transactions across several markets suggests that the object of discrimination is critically important. When the object of discrimination is considered to be controllable, or a choice of the individual, the evidence suggests that the nature of discrimination is animus. Alternatively, when the discriminator bases his actions on an object that is perceived to be out of the control of the individual—e.g., race and sex—the underlying form of discrimination tends to be statistical discrimination.

---

<sup>3</sup> Or “third-degree” price discrimination noted as per Pigou, who argued that such measured discrimination occurs when sellers divide customers “into two or more independent groups, each of which has its own continuous demand function reflecting quantities sold to that group at alternative prices.”

While this behavioral insight is outside of the standard economics model, it is consistent with what psychologists have denoted as attribution theory (Heider, 1958; Weiner, 1995). According to attribution theory, individuals make attributions about the cause and controllability of an action or stigma which lead to emotional reactions that affect the likelihood of helping or punishing behaviors. Heider (1958) showed that causal inference in which the decision maker takes into account the other person's motives and situational constraints, is an important cognitive process for perceiving social contexts. Greenberg and Frisch (1972) find that help that is deliberately given leads to more reciprocity than does accidental help. Most related to our work, individuals are found to be more prejudiced towards those demonstrating mental-behavioral conditions such as obesity, homosexuality, or drug addiction that are perceived as controllable (see, e.g., Teachman et al., 2003; Haider-Markel and Joslyn, 2008; Corrigan et al., 2009; Rukavina and Li, 2011). In contrast, those with physical disabilities that were perceived outside the control of the individual, such as blindness or paraplegia, were more likely to elicit pity and help from others (Weiner et al., 1988).<sup>4</sup>

Beyond academicians, our study also speaks to policymakers. This is because we cannot begin to craft serious legislation to address discrimination until we understand the underlying motivations for why disparate treatment occurs. Given its historical significance, the U.S. government for decades has been codifying and erecting rules that forbid animus-based discrimination. Our study suggests that a much different kind of discrimination—statistical discrimination—permeates many markets and is importantly shaping surplus allocations. If one

---

<sup>4</sup> Research on attitudes towards racial minorities shows that prejudice depends importantly on the perceived source of economic and social disadvantage. Those who attribute disadvantage to internal factors such as a lack of effort or motivation hold more prejudiced views than those who attribute disadvantage to external factors such as a lack of opportunity (see, e.g., Kluegel, 1990; Schuman et al., 1997). Relatedly, Reizenstein (1986) shows that willingness to help a person who was either collapsed on a subway or in need of class notes depends crucially on the perceived controllability of the situation – subjects were significantly less likely to help if the target person was described as drunk or having skipped class to go to the beach.

desires to curb such behaviors, a different kind of solution is apropos because animus-based solutions are not well equipped to attenuate statistical discrimination.

Methodologically, our study highlights that the general approach of combining several data sources to disentangle competing hypotheses can be fruitful. To facilitate causal inference, developing a framework that utilizes a battery of field experiments to isolate relevant comparative statics and measure the extent to which different factors influence observed outcomes can be quite powerful. In this regard, our study provides a blueprint for conducting large-scale field experiments to disentangle competing theories in the social sciences more broadly.

The remainder of the study proceeds as follows. Section II examines the nature and extent of discrimination against the disabled in the market for repairing automobiles. Sections III and IV examine whether race and sexual orientation affect price quotes in the new car market. Section V explores the extent to which discrimination can be detected within information markets. Section VI concludes.

## **II. Study I – Disability and Discrimination in Service Markets**

Our first study is designed to examine the nature and extent of discrimination against the disabled—those confined to wheel chairs. Interestingly, measuring whether, and to what extent, discrimination against the disabled exists has severely lagged the broader literature measuring discrimination along racial or gender lines. Given that the 2000 U.S. Census identified approximately 15.6 percent of the U.S. population (39,582,691 citizens) as considered disabled and that the number of such individuals worldwide is increasing rapidly, this dearth of research is noteworthy. Yet given the difficulties associated with measuring discrimination of any kind, this void is understandable.

Our approach departs from previous studies—all of which, to our best knowledge, are regression-based studies (see, e.g., Baldwin and Johnson, 1994, 1995)—by examining actual behavior in a well-functioning marketplace, the automobile repair market. The assignment given to testers is clear and a typical exercise: approach body shop  $j$  to receive a price quote to fix automobile  $i$ . We include six testers from each of two distinct groups—disabled white males age 29-45 and non-disabled white males age 29-45—who each visited 6 body shops.<sup>5</sup> All testers with disabilities were in wheelchairs and drove a specialized vehicle. And, each of the automobiles, which were personally owned by our disabled testers, had visible body problems. Therefore, the cost of fixing the car damage was independent of the fact that they were specialized cars.

The experiment took place in Chicago. Our original plan was to choose body shops randomly from a given zip code. In a pilot test, however, we discovered that this was not possible because many body shops are not easily approachable by testers confined to a wheelchair.<sup>6</sup> We therefore approached only body shops in which disabled agents could easily access services.

After each disabled tester was paired with one non-disabled tester, we randomly chose 12 accessible body shops; we then split these 12 body shops randomly between the disabled and the non-disabled tester in each pair, such that each tester visited 6 body shops. Importantly, both testers in any given pair approached body shops with the identical car. To control for unobserved body shop effects, we send two testers to each body shop, using a different car to avoid confounding factors. In this case, the non-disabled (disabled) tester in pair  $i$  visited body shops that were previously visited by the disabled (non-disabled) tester of pair  $j$ , where  $i \neq j$ .

---

<sup>5</sup> We were referred to the disabled testers by an organization in Chicago responsible for helping disabled people. Our non-disabled testers were people that the disabled testers knew and trusted to drive their car.

<sup>6</sup> The most common problems were stairs and lack of special parking places in which the disabled agents could easily access services. As discussed more fully below, this limitation in and of itself might play an important role in explaining observed treatment differences across groups.

Before discussing the results, we should mention a few noteworthy design issues. First, by constructing the experiment so that we vary both the automobile and body shop approached across the disabled and non-disabled, we are able to reduce the noise associated with any uncontrolled differences between body shops and automobiles by use of panel data regression techniques. Second, note that the tester drove to the body shop alone and, once at the body shop, proceeded to the desk to ask for a price quote for fixing the car. Whenever a choice of parts was needed (e.g., replace a bumper), the tester was instructed to ask for a new one and not a used one. This is important because if we allowed “steering” of our testers to different repairs important confounds arise. After obtaining the price quote, the tester thanked the mechanic and stated that he would further reflect on whether to have the car fixed. In this sense, it is important to note that in this industry a “list price” for fixing problems typically does not exist and there is no price negotiation. Rather, the mechanic gives a take-it-or-leave-it offer. Thus, bargaining differences across groups are ruled out *ex ante* as a potential explanation of disparate treatment.<sup>7</sup>

Third, each tester was compensated \$150 for participation. We provided an extra \$400 to the tester with disability towards fixing each car at one of the body shops that participated in the study. Finally, given that each tester approached 6 different body shops and each body shop was approached twice, we have a total of 72 observations across 36 tester-body shops.

### *Experimental Results*

Table 1 summarizes our experimental results. Consider the non-disabled tester in pair 1. Entries in the first row reveal that when this tester approached body shop #1 (body shop indicator is in parenthesis) he received a price quote of \$450. The disabled tester he was paired with

---

<sup>7</sup> An interesting extension would be to relax this aspect of the design by examining behavior in a much different market, where haggling is common. To our knowledge, this has not been examined previously. We have gathered such data and the empirical results are consistent with the results presented herein. These findings are available upon request.



received a price quote from body shop #7 of \$575 to repair the exact same automobile. For convenience, we have placed the price quotes in ascending order. Summary statistics in Table 1 provide an indication that the disabled received inferior offers compared to the non-disabled. Overall, the disabled received considerably higher average price quotes, \$1,425, than the non-disabled, \$1,212.

While the data summary in Table 1 provides suggestive evidence that majority agents receive preferential treatment in the form of systemic lower offer prices, they should be viewed as only suggestive since the panel nature of the data are neglected. To account for the panel nature of the data we estimate variants of the following equation:

$$P_{ib} = f(X_{ib}) + \varepsilon_{ib}, \quad (1)$$

where  $P_{ib}$  is the price quote the  $i$ th tester received from body shop  $b$ .  $X_{ib}$  includes an individual group membership dummy variable, which equals one if the agent is disabled, zero otherwise, and five dichotomous variables controlling for automobile type. We specify the error structure as  $\varepsilon_{ib} = \alpha_i + u_{ib}$ ; where,  $\alpha_i$  are tester-specific random effects.

Summary empirical estimates of equation (1) are contained in columns 1 and 2 of Table 2. We include two distinct empirical specifications to provide insights into the robustness of the parameter estimates. The first empirical specification is an ordinary least squares (OLS) model, which does not control for data dependencies. The second model augments the OLS model by adding tester random effects.<sup>8</sup> Regression results in Table 2 provide evidence of our first result:

*Result 1: The disabled receive higher price quotes than the non-disabled receive.*

---

<sup>8</sup> These empirical results, as well as all results presented below, are robust to estimating a model that clusters standard errors.

This result is found in both specifications in Table 2. Coefficient estimates suggest that disabled agents receive roughly a \$213 (or 30 percent) higher quote price than the non-disabled, and this difference is statistically significant at the  $p < .05$  level in both models.

Inference as to why this disparate treatment exists is an important issue. Several factors are potentially at work: (i) access—as described above, many body shops are not easily accessed via wheelchairs; this considerably restricts the set of price offers the disabled can receive; and (ii) time—while the non-disabled can easily park and proceed to the front desk, the process is much more complex for the disabled. First, the disabled agent must find a suitable parking place: it is very uncommon to have designated places for the disabled in body shops. After finding an appropriate parking space, the disabled must commit much more effort and time to approach the service desk. And finally, in some cases it is necessary to leave the car for the day in order to obtain a price quote. Using a taxi is much more complex for the disabled than for the non-disabled.

#### *Complementary Evidence*

To investigate the source of discrimination further, we obtained data on search effort of consumers and perceived search effort from mechanics to examine if realizations of these variables are consistent with the pattern of discrimination observed. For the consumer survey, we had our testers complete the short questionnaire contained in Panel A of Appendix I. Most importantly, from this survey we find that even though we have small sample sizes, the disabled typically consult far fewer body shops: on average, the non-disabled visit 3.5 different mechanics whereas the disabled visit only 1.67 mechanics, a difference that is statistically

significant at the  $p < .01$  level using a small sample test of means.<sup>9</sup> Moreover, 33 (16.5) percent of the disabled (non-disabled) never seek a second opinion.

Concerning the supply side, we asked forty body shops in the Chicago area the two questions contained in Panel B of Appendix I.<sup>10</sup> These questions revolved around body shop perceptions of the degree of search among the disabled and non-disabled. In practice, twenty body shops answered question 1a while twenty different body shops answered question 1b. Upon answering this question, all forty body shops answered question 2.

The results are consonant with the consumer-side statements observed above: mechanics believe the disabled approach 1.85 body shops for price quotes while the non-disabled approach 2.85; the difference is significant at the  $p < .05$  level using a small sample t-test of distribution differences. Likewise, while nearly 60 percent of body shops believed the two groups obtained a similar number of price quotes, over 40 percent believed that the disabled obtained fewer quotes. Using a test of proportions, we find that this percentage is significantly different from zero at the  $p < .05$  level.

This evidence is consistent with economic discrimination (Arrow, 1972; Phelps, 1972) based on mechanics' beliefs about relative search costs and how they map into reservation value distributions. Yet, the survey evidence alone is only suggestive and further investigation is necessary to pinpoint the underlying mechanism at work. For these purposes, we provide a sharper focus on the underlying reason for discrimination by running a complementary field experiment that closely follows the approach of our main field experiment.

---

<sup>9</sup> This search difference was after they received a \$1000 price quote. When testers receive a smaller price quote (\$200) the search difference is directionally similar, but only significant at marginal levels.

<sup>10</sup> Due to IRB purposes, the body shops who received the questionnaire were different than those approached by the testers in our experiment.

In this complementary treatment, we visited 24 of the 36 body shops previously visited, but in this case we had new testers and new vehicles in need of repair. To begin, we had our tester pairs follow the original protocol to replicate the prior study. We then had these exact same testers approach body shops but had both the disabled and abled agents explicitly note that “I’m getting a few price quotes.” If search costs represent the main reason why discrimination is occurring, then we should observe that the offer discrepancies disappear in this treatment.

This is exactly the result that we observe, as shown in the summary data contained in rows 1-4 of Table 1A. Even though in the replication treatments the offer discrepancies are similar to the differences observed in the original study, simply by having our disabled agents note that “I’m getting a few price quotes,” they were able to secure much lower offers – a data pattern highlighted in Figure 1. Importantly, the lower offers received by disabled testers after signaling a willingness to search are not statistically different from those received by the abled. In fact, the disabled now receive slightly lower price quotes.

Empirical support of this result is contained in columns 3 and 4 of Table 2, where we estimate equation (1) using these new data. While the small samples render our estimates quite imprecise, both specifications highlight that disabled agents receive offers that are considerably higher than non-disabled counterparts. Yet when both agents clearly signal that this particular mechanic visit is just one part of their entire search process, these disparities are attenuated and indeed change signs. As a whole, the survey and complementary field experimental data lead to our next result:

*Result 2: The disparate treatment of the disabled is consistent with a model of statistical discrimination driven by search cost differences.*

Importantly, since physical disabilities such as being confined to a wheelchair is perceived as beyond the control of the consumer, attribution theory would predict that the observed disparate

treatment is due to statistical discrimination (Heider, 1958; Weiner, 1995). Consonant with this theory, we find little evidence that the disparate treatment is due to a general “distaste” for the disabled minorities in the Becker (1957) sense. Rather, the market results are consonant with statistical discrimination (see, e.g., Arrow, 1972; Phelps, 1972). Moreover, traits of the mechanics did not affect the level of discrimination observed, suggesting that the effect might be market wide.

### **III. Study II – Race Discrimination in Product Markets**

Our second experiment moves the exploration of discrimination to the market for purchasing new cars. In this second experiment, we focus on a different individual characteristic that is beyond the control of the consumer—race. Accordingly, if we observe discrimination against non-whites in this experiment, attribution theory would suggest that it is driven by statistical discrimination rather than animus.

A total of twenty male testers (10 African American and 10 Caucasian) bargained to purchase a new automobile at five different dealerships in the Chicago-area. All testers were employees at the University of Chicago and were randomly assigned the task of bargaining for either a high- or a low-priced car at each of the five dealerships. For each of the four tester type/price level pairings, we thus observe five audits per dealership.

The order in which tester types (Caucasian or African American) entered a particular dealership and bargained for a particular car (high- or low-priced) was randomized with a lag between visits to the same dealership. At each dealership, testers followed a uniform bargaining strategy that called for them to focus quickly on a pre-determined make and model and start negotiations. After deciding which car they were going to bargain over, testers were instructed

to wait for an initial offer from the salesman. Upon receiving the initial offer, the testers were instructed to confer in private for a few minutes and respond with a pre-determined counter offer. After making the initial counter offer, testers were instructed to wait for a counter offer by the dealer. If the salesperson responded with his/her own counter, the testers were instructed to confer in private and employ a “split-the-difference” bargaining strategy. This process continued until the dealer either (i) accepted a tester’s offer or (ii) refused to bargain further.

Table 3 summarizes our experimental results. As noted in the table, the average initial price quote for a minority tester seeking a low-end car was \$14,035 – a figure that is statistically indistinguishable from the \$14,113 average price offered Caucasian counterparts. However, when bargaining for a high-end car, minority testers face initial offers that are approximately 1.5 percent higher (\$37,129 versus \$36,627) than their Caucasian counterparts. We observe similar differences in final dealer offers. For the low-end cars, there is no discernable difference in the average final price quote across minority and Caucasian testers. However, for the high-end cars, minority testers receive final offers that are approximately 3 percent higher than their Caucasian counterparts.

We observe analogous data patterns when examining other aspects of the bargaining process. For example, when bargaining for low-end cars, minority and Caucasian testers are equally likely to have their initial counter offer rejected by the dealer. However, minority testers are almost twice as likely (44 percent versus 24 percent) to have the initial counter offer rejected when bargaining for a high-end car – a difference that is statistically significant at the  $p < 0.10$  level using a two sample test of proportions. Likewise, there is no detectable difference in the length of time spent bargaining over low-end cars across tester types. For the high-end cars, in

contrast, the bargaining process is terminated approximately 5.6 minutes sooner if the tester is a minority – a difference that is significant at the  $p < 0.10$  level.

While the data summary in Table 3 provides some evidence that Caucasian agents receive preferential treatment when bargaining over high-end cars, they should be viewed as only suggestive since the panel nature of the data are neglected. To account for the panel nature of the data we estimate variants of the following linear regression model:

$$P_{ij} = X_{ij}\beta + \alpha_j + u_{ij}$$

Where  $P_{ij}$  is the initial (final) offer received by the  $i$ th tester from the  $j$ th dealership;  $X_{ij}$  includes a dichotomous indicators for a Caucasian tester and a Caucasian tester bargaining for a high-end car; and  $\alpha_j$  are dealership by car-type fixed effects. To control for potential observed heterogeneities across testers, we include tester specific random effects.

Empirical estimates are contained in Table 4 and highlight patterns consonant with the raw data. For example, as noted in the first row of the table, there is no significant difference in either initial or final offers across minority and Caucasian testers when bargaining for low-end cars. However, when bargaining over high-end cars, initial (final) price offers for Caucasians are approximately \$630 (\$1010) less than those received by minority counterparts – differences that are significant at the  $p < 0.05$  level. Interestingly, individual specific traits of the car dealers do not significantly influence the nature or extent of discrimination.

Taken jointly, these data are *suggestive* of statistical discrimination against minorities in the new car market.<sup>11</sup> If the standard animus theory organized behavior, then one would expect minorities to receive higher price quotes for both high- and low-end cars. Yet, we only observe differential behavior when testers are bargaining for higher-end cars. In addition, if animus was at work, one might expect that the race of dealers would be correlated with how consumers are

---

<sup>11</sup> We stress suggestive because one can construct an animus based story based on heterogeneity.

treated, which we find little evidence to support. Combined with the data from the disabled experiment, we argue that there is empirical support of attribution theory for understanding the underlying nature of discrimination in markets.

#### **IV. Study III – Sexual Orientation and Discrimination in Product Markets**

Our first two field experiments examine the nature and extent of discrimination by varying a factor—race and being confined to a wheel chair—that sellers view as uncontrollable. A more demanding field test of attribution theory is an exploration of an object of discrimination that some sellers perceive as controllable whereas others do not.

After a broad-based search, we concluded that sexual orientation was a feature that matched the research needs. There is a growing body of literature highlighting that minorities are more likely than Caucasians to view sexual orientation as a choice rather than a genetically predetermined trait (see, e.g., Herek and Capitano, 1995; Jayaratnee et al., 2009).<sup>12</sup> This empirical body of work induced us to examine the effect of sexual orientation on price quotes within the market for new automobiles.

Beyond its import in testing theory, very few discussions around the globe are as divisive as those concerning issues related to sexual orientation. Earning penalties for homosexual men has been documented in surveys across Europe (United Kingdom: Arabsheibani et al., 2004; The Netherlands: Plug and Berkhout, 2004) and the question of gay marriage in the US was central in the 2004 Presidential Campaign.<sup>13</sup> Within the EU sexual orientation discrimination in new member states has been a hot button issue, and there are wide variations in statutes across the US

---

<sup>12</sup> Scientists are divided on the source of sexual orientation, with many arguing that it arises from a nature and nurture mix. We do not take a stand on this issue, rather we base our identification off perceptions across race of the sellers who we observe.

<sup>13</sup> The debate regarding same-sex marriage remains a divisive issue today. As of June 2010 only five states and the District of Columbia officially perform same-sex marriage. Three other states recognize, but do not perform, such marriages and eight additional states grant same sex couples benefits similar to those in a civil union. Of the remaining thirty-four states, all but two have statutes defining marriage to persons of the opposite-sex and do not award benefits to same sex partners.



prohibiting hate crime based on the sexual orientation of the victim. While thirty-one states recognize as hate crime those acts motivated by bias or prejudice against an individual's sexual orientation, there are nineteen states that do not recognize acts motivated by such predisposition as a violation of hate crime laws.

Yet, discussions regarding sexual orientation are not limited to the issue of same-sex marriage and hate crime. In May 2005 the conservative American Family Association (AFA) boycotted the Ford Motor Co. because of advertisements in publications the AFA deemed to serve gay and lesbian consumers. The automaker responded by promising to cease advertising campaigns explicitly targeting gay consumers for most of its brands.<sup>14</sup> In the face of criticism of its reaction to the AFA boycott, Ford released an announcement that the change in advertising policy did not reflect employment practices for the firm, in which nondiscriminatory policies cover employees regardless of race, religion, gender, or sexual orientation. Thus, while Ford was unwilling to allow sexual orientation to affect behavior in the labor market, it was willing to allow such factors to influence advertising strategies on the product market. Such actions are especially interesting given that differential treatment of individuals based on group affiliation is, in many cases, both illegal and a social taboo.<sup>15</sup>

To operationalize our test, we randomly matched twenty-four white male testers (undergraduate students at the University of Chicago) into twelve distinct "pairings," each of which bargained to purchase a new automobile at six different randomly selected Chicago-area dealerships. Each "pairing" was instructed to portray themselves as friends at three pre-determined dealerships and as partners at the remaining three dealerships. Every dealership was

---

<sup>14</sup> For more information on this boycott and Ford's subsequent reaction, we refer the interested to [http://money.cnn.com/2005/12/06/news/fortune500/ford\\_ads/index.htm?cnn=yes](http://money.cnn.com/2005/12/06/news/fortune500/ford_ads/index.htm?cnn=yes).

<sup>15</sup> It should be noted, however, that there are thirty states in the US that do not explicitly include differential treatment based on sexual orientation in anti-discrimination statutes.

visited by one pair of testers portraying themselves as friends and by a second pair of testers portraying themselves as partners. Our testers, thus, bargained for a total of seventy cars at thirty-five dealerships.<sup>16</sup>

The order in which tester types (partners or friends) entered a particular dealership was randomized with a lag between visits to the same dealership. At each dealership, testers followed a uniform bargaining strategy that called for them to focus quickly on a pre-determined make and model and start negotiations. All testers were instructed that the same individual should bargain with the salesperson at every dealership visited. This allows us to employ panel regression techniques that explicitly control for unobserved heterogeneities that may affect observed outcomes. Since not all dealerships offered the same make and models, testers bargained for nine different automobiles.

After deciding which car they were going to bargain over, testers were instructed to wait for an initial offer from the salesman. Upon receiving the initial offer, the testers were instructed to confer in private for a few minutes and respond with a pre-determined counter offer.<sup>17</sup> After making the initial counter offer, testers were instructed to wait for a counter offer by the dealer. If the salesperson responded with his/her own counter, the testers were instructed to confer in private and employ a “split-the-difference” bargaining strategy. This process continued until the dealer either (i) accepted a tester’s offer or (ii) refused to bargain further.<sup>18</sup>

---

<sup>16</sup> The model that testers were instructed to bargain for was unavailable at one of the randomly selected dealerships. We thus observe only five outcomes for one matched set of tester pairs.

<sup>17</sup> The initial counter offers were determined via a pilot study in which same sex couples bargained for the set of cars used in our experiment at randomly selected dealerships in the Chicago area. For each make, the initial counter offer was set at approximately 10-15% less than the best offer received during the pilot study.

<sup>18</sup> The testers did not purchase any cars. If a dealer accepted a given offer, the testers were instructed to thank the dealer and say that they needed to think things over before they could make a final decision. However, it is important to note that we purchased one car at the end of the study to compensate the dealerships for the time and effort of the salesperson and that we based this purchase decision on offers extended to our testers.

Testers were instructed to write down every offer made by the salesperson before conferring in private and to record the approximate time spent bargaining. Upon leaving the dealership, each tester completed a data sheet that included columns to record the race and gender of the salesperson with whom they negotiated, the approximate time spent negotiating, the various prices offered by the salesperson, etc.

Before visiting the dealerships, all testers attended a training session conducted by the authors. The training sessions provided testers with a copy of the bargaining script and information about the particular cars for which they would be negotiating. Next, the testers practiced mock negotiations with the trainer to gain comfort with the bargaining process. Immediate feedback on the negotiation strategy was provided to ensure comparability.

### *Experimental Results*

Table 5 provides summary statistics for the audit data, which includes information on various pecuniary and non-pecuniary outcomes along with observable characteristics that may affect bargaining. Perusal of Table 5 highlights a number of asymmetries in bargaining outcomes across our various tester types. First, conditioned on car specific averages, testers portraying partners receive initial offers that are approximately \$432 lower than counterparts portraying friends. Interestingly, this difference is correlated with the race of the salesperson. Partners receive initial offers that are \$728 higher (\$1181 lower) than friends when approaching a minority (Caucasian) salesperson. Similarly, the difference in initial offers across friends and partners is correlated with the testers' true sexual orientation - homosexual testers receive initial offers that are approximately \$1,190 higher when portraying partners as opposed to friends.<sup>19</sup>

---

<sup>19</sup> This difference holds for both minority and Caucasian dealers, but is significantly more pronounced for minority salespersons.

Second, we observe similar data patterns when examining price concessions that arise over the bargaining process. Final dealer offers for testers portraying friends are approximately 11.4 percent lower than the initial offer. For counterparts portraying partners, the concession in prices from the initial to final offer is only 9.15 percent. As with the initial offers, these differences depend crucially on dealer race.

Third, testers portraying friends have a higher percentage of offers accepted than do partners. Whereas 8.6 percent of the initial counter offers from testers portraying friends are accepted by a dealer, not a single initial offer is accepted from testers portraying partners. Similar differences arise when examining the likelihood that a tester's highest offer is accepted – friends are approximately 75 percent more likely than partners to have their highest offer accepted.

Finally, dealers are more likely to terminate bargaining following the initial counter offer from testers portraying partners. Dealers are approximately 80.9 percent more likely to terminate the bargaining process with such testers than they are with counterparts portraying friends. In aggregate, these raw data are suggestive of discrimination against individuals perceived to be homosexual – pecuniary outcomes in the new car market depend critically on observable characteristics such as the sexual orientation of the buyer and race of the salesperson.

#### *Differential Treatment at the Same Dealership*

We observe eleven audits dispersed over ten unique dealerships where a tester's best (highest) offer was accepted. Seven of these dealerships reject offers from testers perceived as homosexuals. In five of these instances, dealers reject offers that equal or exceed the agreed upon price at which they are willing to sell the same car to testers perceived as friends. On average, dealers are willing to forego nearly \$890 in potential profits to avoid transacting with

perceived same-sex couples with two dealerships indicating a willingness to forego at least \$1500 in potential profits.<sup>20</sup> Interestingly, in both instances where the rejected offer comes from heterosexual testers portraying friends, the testers were offered the opportunity to purchase a less expensive car. Yet, only 42.9% of the dealerships (3 out of 7) extend such offers to testers portraying homosexuals.

Similarly, we observe audit pairs at nine unique dealerships where the high offers from both sets of testers are rejected, but only one audit pair is offered the opportunity to purchase a less expensive car. In approximately 77.8% of these cases (7 out of 9), the audit pair offered the opportunity to purchase a less expensive car were portraying friends. Using a test of proportions, we find that the percentage of testers portraying partners that receive the option of purchasing a less expensive car is less than the proportion of testers portraying friends that receive such offers ( $p < 0.05$ ).

Combined with insights garnered from Table 5, these data suggest a first result:

***Result 1: Dealers appear to discriminate against perceived same-sex couples.***

While the raw data are suggestive that discrimination against homosexuals is prevalent in the car market, they do not control for potentially important unobserved tester-specific characteristics that may be correlated with observed bargaining skill and associated bargaining outcomes.

#### *Asymmetries in Discrimination: The Role of a Dealer's Race*

Recall that each set of testers visited multiple dealers portraying themselves as both partners and friends. To account for data dependencies, we estimate the following linear regression model:

$$(2) \quad P_{ij} = X_{ij}\beta + u_{ij}$$

---

<sup>20</sup> It should be noted that there are also potential profits left on the table in both instances where a dealer rejects the best offer of testers portraying friends and accepts the best offer from partners. However, the average measure of foregone potential profits in these instances is approximately 50% lower.

Where  $P_{ij}$  is the initial offer received by the  $i$ th tester from the  $j$ th dealership;  $X_{ij}$  includes a dichotomous indicator for a Caucasian salesman bargaining with an audit pair portraying friends, indicators for a minority salesman bargaining with an audit pair portraying partners and friends, and dichotomous indicators for the various car types used in the study. To control for potential observed heterogeneities in bargaining skill across testers, we include tester specific random effects.

Summary empirical estimates of equation (2) are contained in Table 6 and provide evidence that price differentials across tester types are correlated with the race of a salesman. For example, parameter estimates in Table 6 suggest that initial offers made by Caucasian salesmen to partners are lower than the initial offers made to friends. On average, testers portraying friends receive initial offers from Caucasian salesmen that are about 12.6% (or \$1471) greater than the offers received from these salesmen when portraying partners, with this difference is significant at the  $p < 0.05$  level.

This result is consistent with a model of statistical discrimination. The literature examining the relationship between sexual orientation and labor market outcomes notes an approximate 15-30% wage *penalty* for homosexual men (see, e.g., Badgett, 1995; Berg and Lien, 2002; Black et al., 2003, Blandford, 2003; Carpenter, 2007). Given these wage differentials, it is plausible to assume that dealers perceive that the distribution of reservation prices for homosexual men lies to the left of that for heterosexual counterparts. If so, they would maximize expected profits by offering the former group *lower* prices.

Importantly, however, this pattern of behavior is reversed when examining price offers from minority salesmen. As highlighted in Model A, minority salesmen make initial offers that are approximately \$1100 (or about 8.9%) greater when approached by an audit pair portraying

partners. This generates an approximate \$1703 (or about 14.6%) difference in the offers received by testers portraying partners when approaching a minority salesman as opposed to a Caucasian counterpart. Both of these differences are statistically significant at the  $p < 0.05$  level. Moreover, as highlighted in Model B, these data patterns continue to hold if we control for the true sexual orientation of testers when portraying partners.<sup>21</sup>

Combined, the data suggest a second result:

***Result 2: Discrimination against same sex couples depends upon the race of a dealer. Minority dealers provide higher initial offers to same sex couples whereas Caucasian salesmen show preferential treatment to such consumers.***

Figures 2 and 3 provide the cumulative distribution functions for the initial offers received by a given tester type (partners or friends) from minority and Caucasian salesmen and highlight data patterns consistent with Result 2. Whereas the distribution of initial offers for partners first-order stochastically dominates the distribution for friends when the salesman is a minority, we see the reverse pattern when the salesman is Caucasian - the distribution of initial offers for partners is first-order stochastically dominated by the distribution for friends.

These data patterns are in line with attribution theory. Given that controllability of sexual orientation is perceived differently across races, we would expect animus from those sellers who view the object of discrimination as a choice. And, if any disparate treatment occurs, statistical discrimination should be practiced by those sellers who view sexual orientation as uncontrollable. This is exactly what we find from our natural field experiment. Such observed data patterns are in the spirit of studies that argue that minorities are more likely to exhibit hostile attitudes towards homosexual men (see, e.g., Harper, 1991; Herek and Glunt, 1991;

---

<sup>21</sup> Additional results from Model B suggest that there is a fundamental asymmetry in the initial offers received by an audit pair portraying partners that is correlated with true sexual preferences: homosexual testers portraying partners receive initial offers that are approximately \$1,350 (or 11.7%) greater than heterosexual testers portraying partners with this difference statistically significant at the  $p < 0.10$  level.

Black, Oles, and Moore, 1998). Viewed through the lens of Becker's (1957) model, these data provide an example of a group of sellers who are foregoing a profit to cater their own prejudice. In fact, they appear to be foregoing a chance to price discriminate to cater their own tastes of bigotry.

## **V. Study IV – Race and Gender Discrimination in Information Markets**

Thus far, our analysis has focused on identifying the nature of discrimination within the context of formal market behavior. Our final experiment tests for discrimination outside the traditional product market. To do so, we focus on one aspect of social interaction: helping strangers. The psychology literature on helping behavior is rich and highlights that people are willing to help others even when help is costly and no material reward to the helper exists. Examples range from minor help with low cost (telling people the time), to helping behavior in which people risk their own life to save others (jumping into a turbulent sea to save a drowning person).

Do strangers discriminate in their willingness to help? In this section of the study, we show that the willingness of people to help strangers is affected by the receiver of the help's race (African American versus white), gender, and age (20 years old versus 50 years old). In a series of field studies we find that people use these differences to discriminate in assisting others. This discrimination depends also on the help giver's group affiliation (race, gender, and age).

### *Related literature*

One night in 1964 in New York, Kitty Genovese was stabbed to death when returning home from work at 3 a.m. What made this case famous is that the murderer took about 30 minutes to kill his victim. He repeatedly stabbed her, ran away and then returned to stab her again. Her screams attracted the attention of dozens of people; who watched the attack from the



safety of their apartments. Not one person intervened, and no one telephoned the police. This shocking murder started an important literature in social psychology studying what influences the tendency of people to help, which is now textbook material (e.g., Darley and Latane, 1968, 1982).

In their pioneering work on helping behavior, Latane and Darley (1970) proposed a model consisting of five steps a person must undertake to assist the fellow: noticing the need for help, being able to identify it as such, taking responsibility for the action, deciding on a way to help, and weighing perceived costs and risks.<sup>22</sup> Although there is a vast empirical literature examining helping behavior, relatively few of these studies have explored whether and how discrimination affects such acts.

Gaertner and Bickman (1971) and Shaffer and Graziano (1980) explore whether individuals are more likely to assist a caller in need when both the caller and helper are of the same race. Results from these studies provide limited evidence of discrimination. For example, Shaffer and Graziano (1980) find no cross-racial differences in the rate with which subjects were willing to assist callers. In contrast, Gaertner and Bickman (1971) show that while black subjects were equally likely to help both black and white callers, white subjects were significantly less likely to help black callers.<sup>23</sup>

Gaertner et al. (1982) also find some evidence of racial discrimination. In their study white female college students, either alone or in the presence of passive bystanders, witnessed an

---

<sup>22</sup> Their model was based on prior empirical work such as Latane and Rodin's (1969) "fall of a chair" study in which participants heard a crash and loud screams of a woman in pain. When alone, 70% of the subjects came to the room or offered to call for help within two minutes, while only 20% of the participants responded when they were in pairs. Latane and Rodin termed this effect the "diffusion of responsibility" – the presence of other bystanders allows people to diffuse both responsibility for helping and guilt for not helping and, thus, decreases the helping response.

<sup>23</sup> It is important to note that these studies differ in the gender of the helper. Whereas callers in Gaertner and Bickman (1971) were males, callers in Shaffer and Graziano (1980) were females. Hence, the seemingly different results across these studies are consonant with our findings reported below.

emergency involving a black or white victim. In the presence of bystanders, blacks were helped more slowly than whites. In contrast, subjects helped black and white victims equally quickly when they were alone.

Eagly and Crowley (1986) provide a meta-analysis of 172 studies exploring the effect of gender on helping behavior. Results from their review show that while men are more helpful than women, women are more likely than men to receive help. For example, Mallozzi et al. (1990) have both male and female testers, standing by the side of a road in obvious need of assistance, count the number of cars that passed before one stopped to offer help. Empirical results indicate that while women were more likely than men to receive help, men were significantly more likely to stop and offer help.

West et al. (1975) studied the effect of the sex and race of the victim on helping a person whose car has ostensibly broken down. Four different tester – a black male, a black female, a white male, and a white female – stopped a car on the shoulder of a road with the hood open and waited for someone to stop and offer assistance. As in Mallozzi et al. (1990), female victims were more likely to receive help and men more likely to offer help. Moreover, black (white) testers received help faster when stopped in neighborhoods that were predominantly black (white).

### *Experimental Design and Results*

We use four different tasks to measure discrimination in helping behavior towards strangers. In the first two tasks, testers dropped either a pen or their keys on the ground in front of a passing stranger. For each task, our measure of helping behavior was whether the stranger stopped to pick up the pen/keys and return them to the tester. In the third task, the tester asked a stranger if they had change for a dollar bill so the tester could pay for a parking meter. In the

final task, testers approached a passing stranger and asked for directions to a well-known, local landmark. For this task, we observed two different measures of helping behavior; whether the helper provided directions and the associated length of the conversation.

We employed a 2\*2\*2 design with eight groups of testers that differ by race (African American or white), gender, and age (20 year old versus 50 year old). In each of the eight groups, we employed five testers per task – each of whom interacted with seventy-five potential helpers. Our data are therefore drawn from interactions between forty unique testers and 3,000 potential helpers. Testers were recruited from the students and staff of a large university, and each tester was paid \$50 for participating. All data was collected between 10am and 4pm on sunny weekdays. A research assistant observed every interaction from a distance and recorded the relevant outcomes.

Table 7 summarizes the experimental results, which can be interpreted as showing that young black men are less likely to receive help than other testers. Across each of the four tasks, the percentage of individuals who help young black men is lower than the corresponding percentage of individuals helping *all* other tester types. For example, as noted in Column 1, only 30 percent of potential helpers return a dropped pen to such testers – a rate that is 27 to 51 percent lower than that observed for other tester types. Similarly, only 36 percent of potential helpers give such testers directions – a rate that is 31 to 52 percent lower than that observed for other tester types. Importantly, for each of these fourteen pair-wise comparisons, the observed difference is statistically significant at the  $p < 0.05$  level.

Young white females, in contrast, receive preferential treatment in information markets. Across all metrics, such testers are more likely to receive help than other tester types. For example, 75 percent of potential helpers give such testers directions – a rate that is approximately

19 percent to more than twice that observed for other tester types. Further, the average duration of such conversations (17.8 seconds) is at least 42.5 percent greater than that for all other counterparts. Similarly, young white females are significantly more likely to receive change when asking a man than all male and middle aged counterparts.

Despite the pronounced differences across race and gender lines for young testers, there are no discernable differences along these same dimensions when we restrict attention to the subsample of middle aged testers. For example, the percentage of potential helpers that return a dropped pen (keys) to our different middle aged tester types range from 46 to 51 (63 to 68) percent – differences that are not significant at any meaningful level. Similarly, there is no discernable difference in the frequency with which individuals give directions to such testers or the associated duration of such conversations.

While the data summary in Table 7 provides anecdotal evidence that young black males (young white females) are discriminated against (receive preferential treatment) in the market for information, they should be viewed as only suggestive since the panel nature of the data are neglected. To account for the panel nature of the data, we make use of Butler and Moffitt's (1982) random-effects probit model to estimate whether or not tester  $i$  successfully received directions from helper  $j$ . Specifically we estimate

$$T_{ij} = \beta X_{ij} + e_{ij}$$

where  $T_{ij}$  equals unity if tester  $i$  received directions from helper  $j$ , and equals zero otherwise;  $X_{ij}$  includes dichotomous indicators for our different tester types (where young black males is the baseline), dichotomous indicators for tester/helper pairings that are of the same race and same gender, and the interaction of these two indicators.

Empirical estimates are contained in the first two columns of Table 8 and provide insights in line with the raw data summary. For example, as noted in column 1, young black male testers are significantly less likely to receive directions than *all* other tester types. Conversely, young white females are significantly more likely to receive directions than *all* other tester types. Secondary results in column 2 highlight the importance of like social grouping on helping behavior – individuals are more likely to help testers who are of the same race as themselves.

To examine the resultant duration of such interactions, we estimate a random-effects regression of the form

$$D_{ij} = \beta X_{ij} + e_{ij}$$

where  $D_{ij}$  is the length (in seconds) of the conversation between tester  $i$  and helper  $j$  and all covariates are as defined above. To account for the fact that not all helpers provide directions and the resultant censoring of the data at zero, we estimate the model using a random effects tobit specification.

Empirical estimates for the model are contained in the final two columns of Table 8 and provide insights consonant with the raw data summary; (i) there are no discernable difference in conversation lengths across middle aged tester types, (ii) the average conversation length for young white female testers is longer than those recorded by other tester types, and (iii) conversations with young black men are significantly shorter than those recorded by all other tester types. Moreover, as noted in the final column of the table, average conversation lengths between testers and helpers of the same race (gender) are significantly longer (shorter) than those observed amongst pairings of different races (genders).

Taken jointly, these data are suggestive of statistical discrimination against young minority males in the market for information. We present two pieces of evidence that suggests

animus might not be at work. First, if animus were the driver of behavior, one would expect both younger and older minority males to receive less assistance than their Caucasian counterparts. Yet, we only observe differential behavior when restricting the analysis to younger testers. Second, there needs to be an underlying model of statistical discrimination consistent with such helping facts.

To better understand these data patterns, we examine whether differences in criminal activity across our tester types predicts differences in the likelihood of receiving aid. One cost of stopping and engaging with a stranger is the risk of being robbed and/or physically assaulted. To explore whether such costs affect helping behaviors, as would be expected under a model of statistical discrimination, we first consider the proportion of all arrests recorded in Chicago during 2008 that are attributed to each of our tester types. We then plot these proportions against the frequency with which the particular tester type receives aid from a stranger.

These data are contained in Figure 4 and are suggestive of statistical discrimination. Young, African-American, males account for a higher proportion of arrests than other tester types.<sup>24</sup> Figure 4 also shows the fitted line from a linear regression of arrest on the decision to aid and shows that arrests explain much of the variance in the decision to aid (the regression model has an  $R^2$  of 0.69). In this regard, our data, again, indicate that attribution theory can lend insights into the underlying nature of discrimination in markets. Within personal information markets, the statistical discrimination *might* operate through perceptions of fear. Yet, we view this final result as speculative, and further research is warranted.

---

<sup>24</sup> These data come from the Chicago Police Department's 2008 Annual Report, available here: <https://portal.chicagopolice.org/portal/page/portal/ClearPath/News/Statistical%20Reports/Annual%20Reports/08AR%5B1%5D.pdf>. It is important to note two details about our use of the above data. First, the above data has 312 arrests attributed to either an unknown race, age or gender; we remove these from the data set, reducing the total number of observations from 196,613 to 196,302. Second, the data set provides a breakdown of arrests by age and gender and separately by race and gender, but not by all three. As a simplification, we assume no covariance between race and age to create imputed arrest counts by age, gender, and race. We use Census demographic information on Chicago to generate the proxy for likely criminal.

## VI. Conclusion

Social scientists have shown that a certain amount of discrimination is directed toward females, nonwhites, and the elderly. This study departs from a traditional analysis of discrimination by investigating the nature *and* extent of discrimination against several distinct groups across several different markets. The paper begins by presenting field evidence that supports the notion that discrimination is present in each of the markets.

Perhaps the novelty of the study, however, is the combination of insights drawn from complementary field treatments, which taken together display an interesting pattern of behaviors. In certain markets, the observed discrimination is not bigotry or animus-based, but consistent with the notion of profit-maximization, or Pigou's (1920) "third-degree price discrimination": in their pursuit of the most profitable transactions, marketers use observable characteristics to make statistical inference about reservation values of market agents. In others, the discrimination is more in line with Becker's (1957) taste based theory of discrimination, or animus.

Interestingly, the nature of discrimination is less driven by particulars of the market or institutions, rather the nature of the disparate behavior is driven by whether the object of discrimination is a choice of the individual or is uncontrollable. This result is important on several fronts. For policymakers, understanding the underlying source of discrimination is a necessary condition for the design and implementation of policies to address differential treatment of individuals in the marketplace. Without such understanding, policies crafted to address discriminatory practices may prove ineffective or promote unintended consequences that exacerbate the ills of such behaviors. Importantly, our results should help policymakers craft legislation that targets the *cause*, rather than the *effect* of discriminatory behaviors.

For theorists, it provides a test of attribution theory and highlights a fundamental, but heretofore overlooked, factor driving the underlying nature of disparate behavior – the extent to which the object of discrimination is perceived to a choice and hence controllable. For empiricists, our work potentially opens up a new area of empirical study that permits one to determine the nature of disparate treatment. Future empirical inquiries should test the robustness of our insights in other markets and extend the toolkit by pairing experimental data with naturally-occurring data.



## References:

- Altonji, J.G. and C.R. Pierret, "Employer learning and statistical discrimination," *Quarterly Journal of Economics*, February, (2001), pp. 1-37.
- Arabsheibani, G Reza, Alan Marin, Jonathon Wadsworth. (2004). "In the Pink: Homosexual-Heterosexual Wage Differences in the UK," *International Journal of Manpower*, 25(3,4): 343-354
- Arrow, Kenneth. (1972). "The Theory of Discrimination," in *Discrimination in Labor Markets*, O. Ashenfelter and A. Rees, eds., Princeton, NJ: Princeton University Press.
- Badgett, M.V. Lee. (1995). "The Wage Effects of of Sexual Orientation Discrimination," *Industrial and Labor Relations Review*, 48(4): 726-739
- Baldwin, Marjorie L. and William G. Johnson. (1994). "Labor Market Discrimination Against Men with Disabilities," *Journal of Human Resources*, 29: 1-19.
- Baldwin, Marjorie L. and William G. Johnson. (1995). "Labor Market Discrimination Against Women with Disabilities," *Industrial Relations: A Journal of Economics and Society*, 34(4): 555-557
- Becker, Gary. (1957). *The Economics of Discrimination*, University of Chicago Press
- Berg, Nathan and Donald D. Lien. (2002). "Measuring the Effects of Sexual Orientation on Income: Evidence of Discrimination?" *Contemporary Economic Policy*, Vol. 20: 294-414
- Black, Beverly, Thomas P. Oles, Linda Moore. (1998). "The Relationship Between Attitudes: Homophobia and Sexism Among Social Work Students," *Affilia*, 13(2): 166-189
- Black, Dan, Hoda R. Makar, Seth G. Sanders, Lowell J. Taylor. (2003). "The Earnings Effects of Sexual Orientation," *Industrial and Labor Relations Review*, 56(3): 449-469
- Blandford, John M. (2003). "The Nexus of Sexual Orientation and Gender in the Determination of Earnings," *Industrial and Labor Relations Review*, 56(4): 622-642
- Carpenter, Christopher. (2007). "Sexual Orientation, Income, and Non-Pecuniary Economic Outcomes: New Evidence from Young Lesbians in Australia," *Review of Economics of the Household*, 6(4): 391-408
- Corrigan, Patrick W., Sachiko A. Kuwabara, John O'Shaughnessy. (Apr. 2009). "The Public Stigma of Mental Illness and Drug Addiction: Findings from a Random Stratified Sample," *Journal of Social Work*, 9(2): 139-147

- Derlega, V.J. and J. Grzelak (1982). *Cooperation and helping behavior: Theories and research*, New York: Academic Press
- Eagly, Alice and Maureen Crowley. (1986). "Gender and Helping Behavior: A Meta-Analytic Review of the Social Psychological Literature," *Psychological Bulletin*, 100(3): 283-308
- Fershtman, C., and U. Gneezy, "Discrimination in a Segmented Society: An Experimental Approach," *Quarterly Journal of Economics*, (2001), pp. 351-377.
- Gaertner, Samuel and Leonard Bickman. (1971). "Effects of Race on the Elicitation of helping Behavior: The Wrong Number Technique," *Journal of Personality and Social Psychology*, 20(2): 218-222
- Gaertner, Samuel, John Dovidio, Gary Johnson. (1982). "Race of Victim, Nonresponsive Bystanders, and Helping Behavior," *The Journal of Social Psychology*, 117(1): 69-77
- Greenberg, Martin S. and David M. Frisch. (1972). "Effect of Intentionality on Willingness to Reciprocate a Favor," *Journal of Experimental Social Psychology*, 8(2): 99-111.
- Haider-Markel, Donald P. and Mark R. Joslyn. (2008). "Beliefs About the Origins of Homosexuality and Support for Gay Rights: An Empirical Test of Attribution Theory," *Public Opinions Quarterly*, 75(2): 291:310
- Harper, Phillip Brian. (1991). "Eloquence and Epitaph: Black Nationalism and the Homophobic Impulse in Responses to the Death of Max Robinson," *Social Text*, No. 28: 68-86
- Heider, Fritz (1958) *The Psychology of Interpersonal Relations*, New York: Wiley.
- Herek, Gregory M. and Eric K. Glunt. (1991). "Interpersonal Contact and Heterosexuals' Attitudes toward Gay Men: Results from a National Survey," *The Journal of Sex Research*, 30(3): 239-244
- Herek, Gregory M. and John P. Capitano. (1995). "Black Heterosexuals' Attitudes Toward Lesbians and Gay Men in the United States," *The Journal of Sex Research*, 32(2): 95-105
- Jayaratne, Toby Epstein, Susan A. Gelman, Merle Feldbaum, Jane P. Sheldon, Elizabeth M. Petty, Sharon L.R. Kardia. (2009). "The Perennial Debate: Nature, Nurture, or Choice? Black and White Americans' Explanations for Individual Differences," *Review of General Psychology*, 13(1): 24-33
- Kluegel, James R. (1990). "Trends in Whites' Explanations of the Black-White Gap in Socioeconomic Status, 1977-1989," *American Sociological Review*, 55(4): 512-525
- Latane, Bibb and John M. Darley. (1970). *The Unresponsive Bystander: Why Doesn't He Help?* Appleton-Century Crofts: New York

- Latane, Bibb and Judith Rodin. (1969). "A lady in distress: Inhibiting effects of friends and strangers on bystander intervention," *Journal of Experimental Social Psychology*, 5(2): 189-202
- List, John A. "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field," *Quarterly Journal of Economics*, 108 (2004), pp. 49-89
- List, John A. and Imran Rasul. (2011). "Chapter 2: Field Experiments in Labor Economics," *Handbook of Labor Economics*, 4(1):103-228
- Mallozi, John, Vincent McDermott, Wesley A. Kayson (1990). "Effect of Sex, Type of Dress, and Location on Altruistic Behavior," *Psychological Reports*, 67(3 pt. 2): 1103-1106
- Phelps, Edmund S. (1972). "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62(4): 659-661
- Pigou, Arthur. (1920). *The Economics of Welfare*. Macmillan and Co.: London
- Plug, Erik and Peter Berkhout. (2004). "Effects of Sexual Preferences on Earnings in the Netherlands," *Journal of Population Economics*, 17(1): 117-131
- Reisenzein, Rainer. (1986). "A Structural Equation Analysis of Weiner's Attribution-Affect Model of Helping Behavior," *Journal of Personality and Social Psychology*, 50(6): 1123-1133.
- Riach, P.A. and J. Rich. (19 Nov. 2002). "Field Experiments of Discrimination in the Marketplace," *The Economic Journal*, 112(483): F480-F518
- Rukavina, Paul B. and Weidong Li. (2011). "Adolescents' Perceptions of Controllability and its Relationship to Explicit Obesity Bias," *Journal of School Health*, 81(1): 8-14
- Schuman, Howard, Charlotte Steeh, Lawrence Bobo, Maria Krysan. (1997). *Racial Attitudes in America: Trends and Interpretations*, Harvard University Press
- Shaffer, David R. and William H. Graziano. (1980). " Effects of Victims' Race and Organizational Affiliation on Receiving Help from Blacks and Whites," *Personality and Social Psychology Bulletin*, 6(3): 366-372
- Teachman, Bethany A., Katherine D. Gapinski, Kelly D. Brownell, Melissa Rawlins, Subathra Jeyaram (Jan. 2003). "Demonstrations of Implicit Anti-Fat Bias: The impact of providing causal information and evoking empathy," *Health Psychology*, 22(1): 68:78
- Weiner, Bernard, Raymond P. Perry, Jamie Magnusson. (Nov. 1988). "An Attributional Analysis of Reactions to Stigmas," *Journal of Personality and Social Psychology*, 55(5): 738-748

Weiner, Bernard. (1995). "Attribution Theory in Organizational Behavior: A Relationship of Mutual Benefit," *Attributional Theory: An Organizational Perspective*, 3-6

West, Stephen G., Gayle Whitney, Robert Schnedler. (1975). "Helping a Motorist in Distress: The Effects of Sex, Race, and Neighborhood," *Journal of Personality and Social Psychology*, 31(4): 691-698

## **Appendix I: Market Surveys**

---

### **Panel A. Tester Survey**

These questions will be used for statistical purposes only. THIS INFORMATION WILL BE KEPT CONFIDENTIAL AND WILL BE DESTROYED UPON COMPLETION OF THE STUDY.

1. In a typical year, how often do you visit the car mechanic? \_\_\_\_
2. After you receive a price quote of the diagnosed problems from the mechanic, do you consult a second mechanic for a quote:  
  
1 (never) 2        3        4        5 (always)
3. How many mechanics will you typically consult after receiving a:  
  \$200 price quote? \_\_\_\_\_  
  \$1,000 quote?    \_\_\_\_\_
4. Age \_\_\_\_\_
5. What is the highest grade of education that you have completed. (Circle one)  
1) Eighth grade   3) 2-Year College                5) 4-Year College  
2) High School    4) Other Post-High School    6) Graduate School Education
6. What is your approximate yearly income from all sources, before taxes?  
1) Less than \$19,999  
2) \$20,000 to \$39,999  
3) \$40,000 or over

### **Panel B. Bodyshop Survey**

"Hi; my name is XXX and I am a Professor at the BBB. As part of my research I would like to ask you two very simple questions. If you'll answer them--it will be great help."

1a. "If a person in a wheelchair needs to fix his car and gets a quote of \$1,000 (not covered by the insurance). How many price quotes do you think he will try to get?"

1b "If a person needs to fix his car and gets a quote of \$1,000 (not covered by the insurance). How many price quotes do you think he will try to get?"

After they answered this question, we asked them

2. "Do you think that there is a difference between disabled people and non-disabled people in the number of price quotes they get?"

**Table 1: Experimental Design and Summary Results – Experiment I**

		<u>Price quote</u>						<u>Mean</u>
Pair 1	Non-Disabled	450(1)	500(2)	635(3)	700(4)	700(5)	900(6)	648
	Disabled	575(7)	750(8)	820(9)	1000(10)	1050(11)	1200(12)	899
Pair 2	Non-Disabled	800(7)	1000(12)	1000(9)	1050(11)	1100(10)	1200(8)	1025
	Disabled	800(3)	1100(4)	1300(1)	1500(2)	1500(5)	2000(6)	1367
Pair 3	Non-Disabled	900(13)	1000(14)	1000(15)	1250(16)	1300(17)	2000(18)	1242
	Disabled	1000(19)	1000(20)	1500(21)	1500(22)	1600(23)	1800(24)	1400
Pair 4	Non-Disabled	80(23)	150(21)	150(19)	200(20)	320(24)	427(22)	221
	Disabled	200(14)	250(15)	350(13)	500(17)	500(18)	700(16)	417
Pair 5	Non-Disabled	150(25)	160(26)	190(27)	330(28)	500(29)	500(30)	305
	Disabled	420(31)	500(32)	500(33)	535(34)	550(35)	550(36)	509
Pair 6	Non-Disabled	3000(36)	3500(35)	3500(32)	4000(31)	4000(33)	5000(34)	3883
	Disabled	3000(29)	3250(26)	3300(30)	4500(25)	4900(27)	4800(28)	3958
						<i>Non-Disabled</i>	<b>Overall</b>	\$1212 se=\$214
						<i>Disabled</i>	<b>Overall</b>	\$1425 se=\$212

Notes: Entries are at the individual level. For example, figures in the first row reveal that when this non-disabled agent approached body shop #1 (number in parenthesis), he received a price quote of \$450 to fix automobile #1. The disabled agent he was paired with received a price quote from body shop #7 of \$575 to repair the exact same automobile. For convenience, we have placed the price quotes in ascending order.

**Table 1A: Complementary Experimental Design and Summary Results – Experiment I**

		<u>Price quote</u>						<u>Mean</u>
Pair 1	Non-Disabled	435(12)	500(1)	539(13)	670(24)	700(29)	1000(18)	640
	Disabled	507(8)	650(3)	710(6)	750(22)	800(5)	1250(4)	778
Pair 2	Non-Disabled	220(25)	250(2)	250(20)	380(32)	430(36)	630(33)	360
	Disabled	180(30)	278(26)	300(9)	499(28)	500(35)	690(27)	408
Pair 1 “More Quotes”	Non-Disabled	485(27)	530(9)	610(35)	645(30)	750(26)	775(28)	633
	Disabled	450(33)	500(32)	500(20)	615(25)	750(36)	799(2)	602
Pair 2 “More Quotes”	Non-Disabled	210(3)	250(22)	350(6)	350(8)	415(4)	430(5)	334
	Disabled	200(1)	200(13)	324(29)	324(24)	430(12)	450(18)	321

Notes: Entries are at the individual level. For example, figures in the first row reveal that when this non-disabled agent approached body shop #12 (number in parenthesis), he received a price quote of \$435 to fix automobile #1. The disabled agent he was paired with received a price quote from body shop #8 of \$507 to repair the exact same automobile. For convenience, we have placed the price quotes in ascending order.

**Table 2: Regression Results – Experiment I**

Variable	Main Experiment		Complementary Experiment	
	(OLS)	(Random Effects)	(OLS)	(Random Effects)
<i>Disabled</i>	212.7* (88.8)	212.7* (88.8)	87.5 (64.7)	87.5 (64.3)
<i>Auto2</i>	422.5* (153.8)	422.5* (151.8)	310* (45.8)	310* (45.7)
<i>Auto3</i>	547.5* (153.8)	547.5* (151.8)	---	---
<i>Auto4</i>	-454.4* (153.8)	-454.4* (151.8)	---	---
<i>Auto5</i>	-366.3* (153.8)	-366.3* (151.8)	---	---
<i>Auto6</i>	3122.5* (153.8)	3122.5* (151.8)	---	---
<i>Constant</i>	666.9* (117.5)	666.9* (117.5)	345.3* (51.2)	345.3* (51.1)
<i>“Many Quotes”</i>	---	---	-17 (64.7)	-17 (64.3)
<i>Disabled*</i> <i>“Many Quotes”</i>	---	---	-109 (91.5)	-109 (91.4)
<i>Tester</i> <i>Random Effects</i>	No	Yes	No	Yes
<i>N</i>	72	72	48	48

Notes:

1. Dependent variable is the log of the price quote received. *Disabled* = 1 if the agent is disabled, 0 otherwise. *Autoi* variables are dichotomous and equal 1 when the agent approached body shops with automobile *i*. *“Many Quotes”* = 1 if the agent is stated upfront that he was gathering many quotes, 0 otherwise.

2. Standard errors are in parentheses beneath coefficient estimates; \*denotes significant at the  $p < .05$  level.

3. In the complementary experiment random effects model tester random effects needed to be excluded because convergence was not achieved due to the fact that the model could not find a positive estimated component.



**Table 3: Summary Statistics – Experiment II**

	Caucasian Tester – Low Price Car	Minority Tester – Low Price Car	Caucasian Tester – High Price Car	Minority Tester – High Price Car
<b><i>Pecuniary Outcomes</i></b>				
Initial Dealer Offer	\$14,113.4 (2942.8)	\$14,035 (2913.5)	\$36,627 (6137)	\$37,129.2 (6367.6)
Initial Dealer Counter	\$13,323.8 (2809.8)	\$13,373.8 (2798.3)	\$35,557.5 (5668.7)	\$36,417.4 (6036.6)
Final Dealer Offer	\$13,073.6 (2741.1)	\$13,089.2 (2696.9)	\$35,167.3 (5680.1)	\$36,227.7 (6113.21)
% of Initial Counter Offers Rejected	12%	12%	24%	44%
Bargaining Time	18.6 (9.7)	19.8 (10.8)	19.1 (12.4)	13.52 (12.2)
<b><i>Non-Pecuniary Outcomes</i></b>				
% of Testers Offered a Used Car	32%	56%	20%	28%
% of Testers Offered a Business Card	92%	88%	72%	52%
% of Testers Offered a Seat	68%	76%	68%	52%

**Table 4: Initial and Final Dealer Offers**

	Initial Dealer Offer	Final Dealer Offer
Caucasian Tester Bargaining for a Low-End Car	102.4 (281.9)	-32.92 (314.5)
Caucasian Tester Bargaining for a High-End Car	-628.9** (320.6)	-1010.11** (368.4)
Dealership by Car Type Fixed Effects	Yes	Yes
# of Observations	100	100
Log-Likelihood	-815.4	-828.6

\*\* Denotes statistical significance at the  $p < 0.05$  level

**Note:** Cell entries are parameter estimates (and associated standard errors) for a series random effects regression models of initial and final dealer offers on indicators for our different tester types. All models include tester specific random effects and dealership by car type fixed effects.

**Table 5: Summary Statistics – Experiment III**

	All Testers portraying Friends	Straight Testers portraying Friends	Homosexual Testers portraying Friends	All Testers portraying Partners	Straight Testers portraying Partners	Homosexual Testers portraying Partners
Diff from Avg Initial Offer for Car	\$215.98 (1467.03)	\$587.99 (1242.5)	-\$858.72 (1604.1)	-\$215.98 (1241.87)	-\$378.24 (1301.2)	\$331.66 (307.9)
Diff from Avg Initial Offer for Car – Caucasian Dealer	\$514.65 (1175.99)	\$803.33 (1141.8)	-\$351.37 (869.4)	-\$666.42 (1129.33)	-\$788.05 (1006.3)	\$549.89 (522.2)
Diff from Avg Initial Offer for Car – Minority Dealer	-\$182.26 (1747.12)	\$294.34 (1367.4)	-\$1492.91 (2210.9)	\$546.31 (1065.97)	\$792.64 (1140.7)	\$258.92 (990.8)
% of Initial Offers Accepted	8.6%	3.8%	22.2%	0%	0%	0%
% of Initial Offers Rejected	14.2%	11.5%	22.2%	25.7%	29.6%	12.5%
% Reduction from Dealer’s Initial Offer	11.4%	12.5% (10.9)	8.2% (6.3)	9.15%	8.4% (8.9)	11.8% (10.1)
Total Percentage of Offers Accepted	20%	19.2%	22.2%	11.4%	14.8%	0%
Total Bargaining Time	18.00 (11.38)	16.46 (9.53)	22.44 (15.4)	16.67 (11.37)	14.13 (6.96)	25.25 (18.4)
<b><i>Dealer Characteristics</i></b>						
% of Male Dealers	91.4%	92.3%	88.9%	88.6%	85.2%	100%
% of Caucasian Dealers	51.4%	53.9%	44.4%	57.1%	66.7%	25%
% of Minority Dealers	42.8%	42.3%	44.4%	37.1%	25.9%	75%
<b><i>Treatment of Tester</i></b>						
% Offered a Test Drive	60%	65.4%	44.4%	45.7%	48.1%	37.5%
% Offered a Less Expensive Car	60%	61.5%	55.5%	45.7%	37.0%	62.5%
Total # of Obs	35	26	9	35	27	8

**Note:** Figures in the table represent summary statistics across the different treatments.

**Table 6: Random Effects Regressions of Initial Dealer Offers**

	Model A Initial Offer	Model B Initial Offer
Baseline – Caucasian Salesman bargaining with Partners	11700.82** (459.5)	11593.21** (470.4)
Caucasian Salesman Bargaining with Friends	1471.71** (390.7)	1691.1** (383.1)
Minority Salesman Bargaining with Partners	1703.2** (478.8)	1400.5** (478.9)
Minority Salesman Bargaining with Friends	602.9 (425.5)	817.6** (415.6)
Homosexual Testers Portraying Partners		1350.7* (689.0)
Tester Random Effects	Yes	Yes
Make and Model Indicator – Aveo Excluded	Yes	Yes
Total Number of Observations	70	70
Log Likelihood	-601.1	-599.3

\*\* Denotes statistical significance at  $p < 0.05$  level

\* Denotes statistical significance at  $p < 0.10$  level

**Note:** Cell entries provide parameter estimates for a random effects linear regression model of initial price offers. Cell entries can be read as follows – average initial offers to testers portraying partners are approximately \$1703.20 greater when approaching a minority salesman as opposed to a Caucasian counterpart.

**Table 7: Summary Statistics – Experiment IV**

	Task 1 – Dropping a Pen	Task 2 – Dropping Keys	Task 3 – Asking a Man for Change	Task 4 – Asking a Female for Change	Task 5 – Asking for Directions
Young White Male Tester	41%	67%	18.4%	4.2%	52% (7.91 Secs)
Young White Female Tester	59%	72%	35.6%	13.4%	75% (17.8 Secs)
Young Black Male Tester	30%	62%	1.4%	1.6%	36% (4.57 Secs)
Young Black Female Tester	56%	69%	31.4%	14%	60% (11.83 Secs)
Middle-Aged White Male Tester	50%	68%	14.8%	10.4%	59% (11.75 Secs)
Middle-Aged White Female Tester	50%	66%	16%	7%	63% (12.49 Secs)
Middle-Aged Black Male Tester	46%	63%	5.4%	6%	61% (12.49 Secs)
Middle-Aged Black Female Tester	51%	64%	11.6%	13.8%	63% (12.47 Secs)

**Note:** Cell entries give the percentage of individuals providing help to a given tester type. For Task 5, we also include the average length of time (in seconds) the helper spent talking with the tester. By design, we employed five testers in each cell – each of whom approached 20 individuals in Tasks 1 and 2, respectively; 20 individuals combined between Tasks 3 and 4; and, 15 individuals in Task 5. We thus observe a total of 3000 unique interactions.

**Table 8: Likelihood of Getting Directions and Average Time of Interaction**

	Prob Getting Directions	Prob Getting Directions	Length of Interaction (Secs)	Length of Interaction (Secs)
Constant - 20 Year Old Black Male Tester	-0.36*** (0.15)	-0.34** (0.16)	-9.93*** (3.28)	-8.24*** (3.48)
20 Year Old White Male Tester	0.41** (0.21)	0.23 (0.22)	8.71** (4.37)	4.76 (4.54)
20 Year Old White Female Tester	1.02*** (0.22)	0.83*** (0.23)	24.12*** (4.27)	19.44*** (4.48)
20 Year Old Black Female Tester	0.61*** (0.21)	0.63** (0.21)	15.03*** (4.32)	15.67*** (4.27)
50 Year Old White Male Tester	0.58*** (0.21)	0.37* (0.22)	14.68*** (4.33)	10.02** (4.53)
50 Year Old Black Male Tester	0.65*** (0.21)	0.67** (0.21)	16.04*** (4.31)	17.00*** (4.26)
50 Year Old White Female Tester	0.68*** (0.21)	0.46** (0.23)	14.69*** (4.31)	9.91** (4.51)
50 Year Old Black Female Tester	0.68*** (0.21)	0.71*** (0.21)	16.26*** (4.31)	17.03*** (4.26)
Tester and Helper of Same Race		0.28* (0.17)		5.86* (3.20)
Tester and Helper of Same Gender		-0.20 (0.14)		-7.57*** (2.83)
Tester and Helper Same Race & Gender		0.17 (0.21)		5.57 (4.06)
# of Observations	600	600	600	600
# of Censored Obs	-	-	249	249
Log Likelihood	-393.86	-389.10	-1784.38	-1775.33

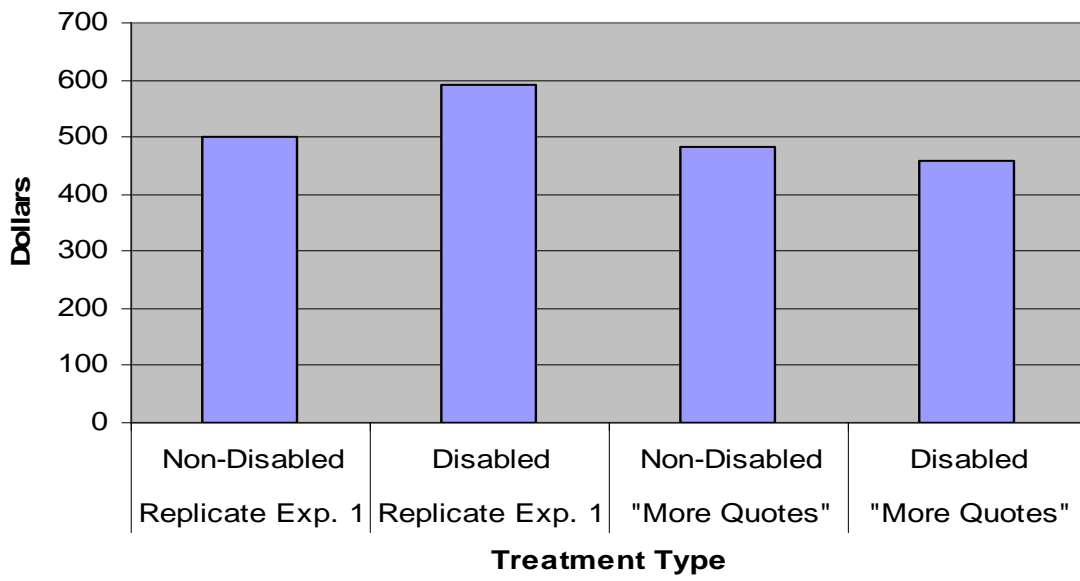
\*\*\* Denotes Statistical Significance at  $p < 0.01$  level

\*\* Denotes Statistical Significance at  $p < 0.05$  level

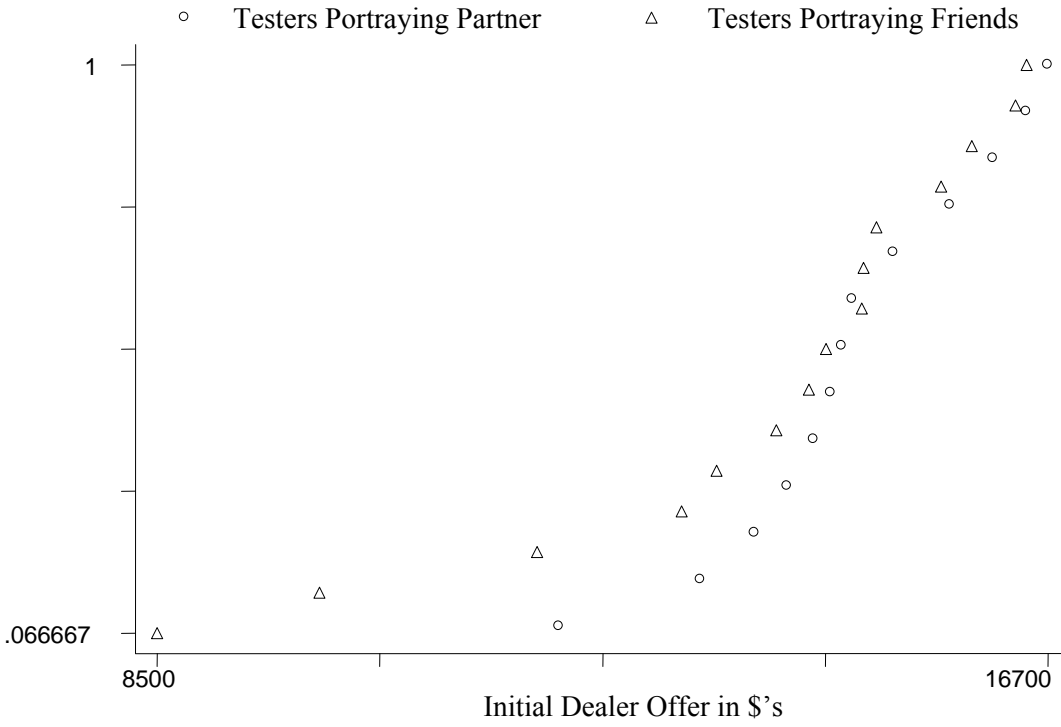
\* Denotes Statistical Significance at  $p < 0.10$  level

**Note:** Cell entries in columns 1 and 2 are parameter estimates from a series of random effects probit models with associated standard errors in parentheses. Cell entries in columns 3 and 4 are parameter estimates from a series of random effects tobit models examining the time (in seconds) that the tester and helper spoke. Associated standard deviations are in parentheses.

**Figure 1: Complementary Experiment I Summary**



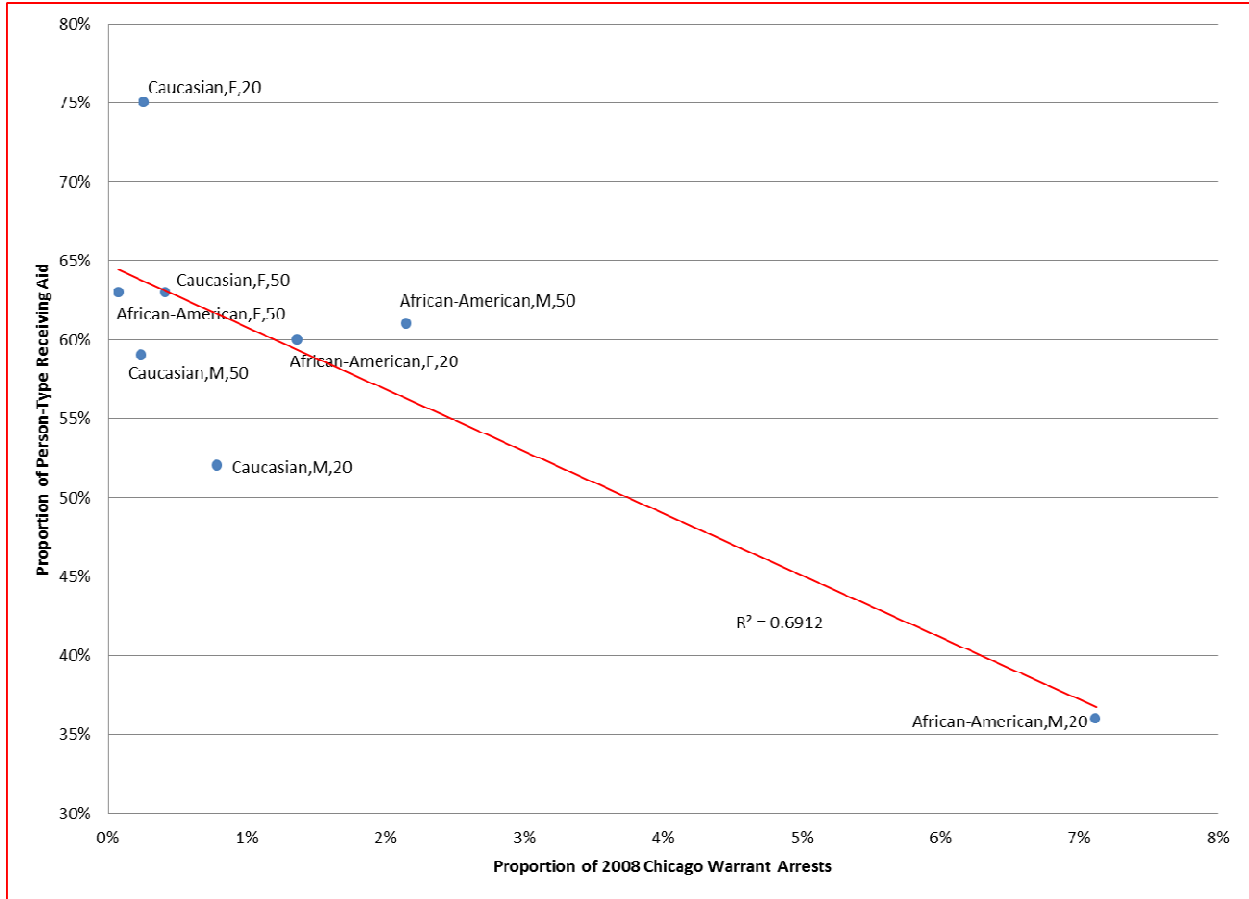
**Figure 2: Cumulative Distribution Function of Initial Offers – Minority Salesmen**



**Figure 3: Cumulative Distribution Function of Initial Offers – Caucasian Salesmen**



**Figure 4: Proportion Given Aid Versus Warrant Arrest Proportion**



**Note:** Arrest data is from the 2008 Chicago Annual Report. The Report has data on arrests in the year 2008. We only look at Warrant Arrests. The Annual Report only has data on race-gender-crime type and age-gender-crime type, not race-age-gender-crime type. To breakout crime type by age-race we assume no covariance. Furthermore, age is reported in bins. We impute the crime rate at a given age by the data's trend in age-crime incidence. Arrest totals are then compared with Census data on demographics of Chicagoans. Total Population is from 2006. Demographic percentages are from the 2000 Census. To get the percentage of Chicagoans in a given age-gender-race we assume uniform distribution of the population under the age of 65.