THE NATURE AND EXTENT OF DISCRIMINATION IN THE MARKETPLACE:

EVIDENCE FROM THE FIELD[*]

John A. List

17 September 2003

Abstract

Empirical studies have provided evidence that discrimination exists in various markets, but they rarely allow the analyst to draw conclusions concerning the nature of discrimination. By combining data from bilateral negotiations in the sportscard market with complementary field experiments, this study provides a framework that amends this shortcoming. The experimental design, which includes data gathered from more than 1100 market participants, provides sharp findings: (i) there is a strong tendency for minorities to receive initial and final offers that are inferior to those received by majorities, and (ii) overall, the data indicate that the observed discrimination is not due to animus, but represents statistical discrimination.

Correspondence:

John A. List, Professor, The University of Maryland, 2200 Symons Hall, College Park, MD 20742-5535, email: Jlist@arec.umd.edu; website: http://www.arec.umd.edu/jlist/

I. Introduction

One would be hard-pressed to find an issue as divisive for a nation as race and civil rights. Yet our understanding of the sources of discrimination within the marketplace remains speculative. Within economics, empirically testing for marketplace discrimination has taken two quite distinct paths: regression-based methods and field experiments. The former technique typically tests for a statistical relationship between an outcome measure, such as wage or price, and a group membership indicator. By and large, regression studies find evidence of discrimination against minorities in the marketplace.[1] Field studies, which have arisen over the past 35 years, typically use matched pairs of transactors to test for discrimination. Due to the control that field studies offer the experimenter, they have become quite popular and have by now been carried out in at least ten countries [Riach and Rich 2002]. Across several heterogeneous labor markets, as well as product markets as diverse as home insurance and new car sales, field studies have made a strong case that systematic discrimination against minorities is prevalent in modern societies.[2]

While empirical studies have served to provide an empirical foundation that indicates discrimination is prevalent in the marketplace, they have been less helpful in distinguishing the nature of discrimination. As Riach and Rich [2002] note, findings from field studies *appear* to be more consistent with the majority white populations having a general "distaste" for minorities in the sense of Becker [1975] or a general "social custom" of discrimination in line with Akerlof [1980]; but statistical discrimination [Arrow 1972; Phelps 1972], or marketers using observable

---

[1] A comprehensive summary of the regression-based literature on discrimination is contained in Altonji and Blank [1999]; Yinger [1998] also provides a recent survey. For a recent study, the interested reader should see Goldberg [1996] who uses a regression-based approach to model discrimination in the new car market.

[2] Field studies measuring discrimination can naturally be split into two categories: audit studies and correspondence tests. The former method uses actual testers and personal contact of the testers with those transacting on the other side of the market. A good example of this type of study is Neumark et al. [1996]. The latter approach does not use personal contact, thereby avoiding potential bias from tester behavior differences across groups. A recent example of this approach is Bertrand and Mullainthan [2003].

characteristics to make statistical inference about productivity or reservation values of market agents, for example, cannot be ruled out, *ex ante* or *ex post*. This fact is highlighted in Heckman and Siegelman [1993, p. 224], who note that labor market "audit studies are crucially dependent on an unstated hypothesis: that the distributions of unobserved (by the testers) productivity characteristics of majority and minority workers are identical."

This study provides a framework for parsing the two forms of discrimination by using a series of field experiments in an actual marketplace—the sportscard market. I begin by documenting an interesting pattern of discrimination: there is a strong tendency for dealers to give nondealer minorities (women, nonwhite, and older agents) initial and final offers that are inferior to those received by their majority counterparts. These results hold when nondealers are acting as buyers and sellers, though the degree of discrimination is greater when agents are selling their wares, providing initial evidence that "consumer-side" discrimination is more pronounced than "seller-side" discrimination.

Similar to the extant discrimination literature, while the pattern of discrimination observed in this marketplace is interesting, one cannot draw firm conclusions about the nature of discrimination. More specifically, the differences in the empirical distribution of prices could be due to at least three sources: i) animus-based or taste-based discrimination, whereby dealers are willing to incur a cost to avoid transacting with a minority, ii) differences in bargaining ability— i.e., majorities are superior bargainers and obtain better outcomes in the marketplace, and iii) statistical discrimination, where minorities have a different distribution of reservation values and in their pursuit of profits dealers use observable variables to make inferences about a relevant but unobservable variable (reservation values).

The main contribution of this study is to provide a framework to rectify this shortcoming by performing a battery of complementary experimental exercises to determine the underlying

nature of discrimination. These additional experiments include i) dictator games, which isolate "tastes" for discrimination, ii) decentralized market experiments, in the spirit of Chamberlain [1948], in which I observe outcomes of face-to-face continuous bilateral bargaining in a multi-lateral market, iii) real auctions run in the marketplace, in which I estimate the underlying distribution of reservation values among the various consumer groups, and iv) dealer perception experiments, in which I observe dealer beliefs about the reservation value distributions of the various groups. In light of the pattern of discrimination observed in the marketplace, each of the three theories of discrimination provides clear predictions of behavior in the complementary experiments.

The complementary experiments provide surprisingly consistent insights. First, it would be incorrect to conclude that the observed differences in treatment are due to noneconomic tastes for discrimination among dealers or differences in bargaining ability across minorities and majorities. Rather, empirical results from all four experiments suggest that the observed differences in treatment are due to statistical discrimination—dealers use minority membership as a proxy for the distribution of reservation values. While this result is important in its own right, as renowned scholars such as Becker [1993] have argued that unfavorable treatment should be considered discrimination only if it is motivated by prejudice, the general contribution of the study is that within a proper framework field experiments can provide a deeper understanding of the nature of marketplace discrimination.

The remainder of this study proceeds as follows. The next section presents the experimental design and discusses the empirical results pertaining to the measurement of discrimination in the marketplace. Section III provides a description of the complementary experimental designs and their empirical results. Section IV concludes.

II. Experimental Design and Results

An important lesson learned from the vast literature on discrimination is that data availability places severe constraints on efforts to understand the nature of discrimination, forcing researchers to speculate about the source of the observed discrimination.[3] For example, Ayres and Siegelman [1995] present evidence from paired audits at new car dealerships in Chicago that suggest dealers quote significantly lower prices to white males than to black or female test buyers. As Ayres and Siegelman [1995, p. 317] note, however, "In car negotiations, dealers might use a customer's race or gender to make inferences about a buyer's knowledge, search and bargaining costs, or, more generally, her reservation price at the specific dealership." This statement highlights that any experimental design that aims to shed light on the nature of discriminatory behavior must take several sources into account, either by controlling them directly in the experimental design or by isolating them and examining the appropriate comparative static changes.

The goal of the experimental design herein is to begin with an examination of whether minority groups face discrimination in a particular bilateral bargaining market. If the evidence is in the affirmative, a systematic experimental analysis will be carried out to determine if the observed discrimination is consistent with predictions from the three theories outlined above: i) animus-based or taste-based discrimination, ii) differences in bargaining ability, and iii) statistical discrimination.[4]

---

[3] 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. And, List [2003b] uses field experimental methods to examine discrimination directed toward the disabled, but does not isolate the three sources of discrimination. The interested reader should also see Goldin and Rouse [2000], who use a "natural experiment" to lend important insights into labor market discrimination, and recent laboratory experiments that examine discrimination [e.g., Fershtman and Gneezy 2001; Mobius and Rosenblatt 2003].

[4] I use "experiment" liberally throughout the study—of course, the variables of interest, gender, race, and age are not assigned randomly (see List and Lucking-Reiley [2002] and Katz et al. [2001] for recent randomized field experiments in much different settings).

My experimental approach departs from previous audit studies by examining actual individual behavior in a well-functioning marketplace—the sportscard market. In this sense, the current experimental design matches real-world settings that economic theory attempts to explain: traders endogenously select into the market and they are likely to have previous experience buying, selling, and trading. Thus, the field experimental strategy employed herein may lead to different results compared with an audit study that exogenously induces certain behavioral patterns and roles on all classes of subjects, but a rigorous examination of natural behavior in an actual environment that our theory intends to explain is an important step in testing the validity of economic theories.

This first experiment (hereafter denoted Experiment I), and all of the complementary experiments described below, were carried out in the same geographic area from June 2002 to July 2003, which ensures comparability across the various experiments. Experiment I, which includes data gathered from four different sportscard shows, includes subjects from four distinct groups: i) white males, age 20-30, ii) white females, age 20-30, iii) nonwhite males, age 20-30, and iv) white males, age 60 or older. Each participant's experience typically followed four steps: (1) consideration of the invitation to participate in an experiment, (2) learning the market rules, (3) actual market participation, and (4) conclusion of the experiment and exit interview.

In Step 1, subjects approached the monitor's dealer table at the sportscard show and inquired about the 1989 *Upper Deck* Ken Griffey Jr. PSA graded "9" baseball card displayed on the table. The monitor then asked if the agent was interested in participating in an experiment that would last about 15 minutes. If the agent agreed to participate, and fit into one of the four groups defined above, in Step 2 a monitor thoroughly explained the experimental rules.[5]

---

[5] After agreeing to take part in the experiment I had the agent complete the first five questions in the survey (see Appendix I). If the agent was "eligible" (i.e. met the criteria of one of the four groups) he proceeded to Step 2. If the agent was not eligible, he took part in a pilot experiment for an unrelated study.

The monitor began by explaining how earnings were determined: the difference between the price paid for the commodity and the maximum reservation price determined market earnings. And, to ensure non-zero earnings, all individuals were informed that they would receive a $10 participation fee upon completion of the experiment.[6] The commodity that was used in the experiment was the identical card displayed on the monitor's table in which the consumer initially expressed an interest in purchasing: a 1989 *Upper Deck* Ken Griffey Jr. PSA graded "9" baseball card.[7] This particular card, and associate grade level, was chosen because it represented a card that many dealers (roughly 25%) at any given show would have in their possession; hence there was an opportunity to obtain multiple observations per subject (described more fully below).

Thus, the assignment given to buyers was clear and a common experience: enter the marketplace and purchase the Griffey card for as little as possible, but at a price that is lower than the reservation value. I was careful to choose two reservation values that could well be considered "low" and "high." The low reservation value ($50) was chosen to place the agent "out of the market" whereas the high reservation value ($80) was chosen to place the agent in the range of prices received in recent internet auctions ($75-$85). Each subject completed a treatment with a low and high reservation value (every subject was induced with the low reservation value first). The order in which I had an individual from each group approach dealers was random (i.e., the first subject approaching a dealer was a female 25% of the time), and the subject could visit only one dealer (the designated dealer) per reservation value. And, I

---

[6] Subjects were also informed that if they purchased the good above their reservation value that the difference would be taken out of their $10 participation fee. No subject purchased the good at a price above their reservation value.

[7] Note that the subject having an interest in the card provides attractive realism in that dealers naturally face this subject type in the marketplace, but it comes at a potential cost—dubious consumers may use this bargaining session to arrange for purchase of the card from the dealer later rather than purchase it in the experiment. To avoid this potential nuance, I ensured the subject that if he would like to purchase the Griffey card from me after the experiment I would guarantee to sell the card to him at the same price at which he purchased the card from the dealer, upon proof of a receipt.

should note that every dealer was approached up to eight times, and no dealer was approached more than twice by a subject in a particular group.[8] Given that payments are typically cash at sportscard shows, I verified that each subject had enough liquidity to purchase the card.

In Step 3, the subject participated in the market. There was no time limit imposed on buyers' bargaining activity, but most interactions were completed within 15 minutes. It should be noted that throughout the experiment the sellers were not aware that an experiment was occurring—this was achieved by instructing subjects to not disclose to anyone that they were taking part in an experiment. This ensured that the process was as natural as possible for sellers, whose behavior is of primary interest in this first field experiment. And, I should stress that, unlike in audit studies, my nondealer experimental subjects did not know that this was a study on discrimination; rather they were informed that it was "an economics experiment." This approach avoids the potential nuance often criticized by commentators—e.g., Heckman [1998] notes that "auditors are sometimes instructed on the 'problem of discrimination in American Society' prior to sampling firms, so they may have been coached to find what audit agencies wanted them to find" (p. 104). Step 4 concluded the experiment—after subjects completed a confidential survey they had begun earlier (see footnote 5), they were paid their earnings in private.

In addition to this buying treatment, which I denote Treatment B, I also conducted a treatment where nondealers were *sellers* of the identical Griffey card (Treatment S). Thus, Treatment S measures "consumer-side" discrimination whereas Treatment B measures "seller-side" discrimination. Given that consumers arrive in the marketplace as buyers, sellers, and/or traders, this treatment again represents a natural experience for the subjects and provides a test of

---

[8] A participant in my experiment never executed a transaction before at least one subject from each subgroup visited the dealer. This was due, in part, by the fact that I carried out the low reservation values in the first round of dealer visitations. Yet transactions did take place in the high reservation treatment that precluded every subgroup from visiting each dealer an equal amount of times. In those cases I substituted another dealer to fill the sample. This is controlled in the statistical tests discussed below.

whether discrimination is present when agents are selling their goods. In every possible sense Treatment S was identical to Treatment B—I again took great care to ensure that these were interactions that would occur naturally in the market. In Treatment S, the monitor approached subjects entering the sportscard show and inquired into whether the agent was planning to sell goods at the show. If the agent answered in the affirmative, the monitor examined the agent's wares to determine if he was eligible to be in the experiment. The monitor ruled the agent eligible if he was entering the market to sell graded baseball cards from the late 1980s or 1990s (to parallel selling the 1989 Griffey card). While Treatment S data were gathered at the same sportscard shows as Treatment B (indeed, the treatments were intermixed), no dealers or nondealers participated in more than one treatment.

Similar to Treatment B, a transaction is expected to occur if the bargainer can find a suitable price. One difference between Treatment S and Treatment B is that I used a constant reservation value of $80 in Treatment S—thus market earnings equaled selling price minus $80. To gather multiple observations per subject, I had subjects in Treatment S approach 5 dealers in the marketplace. Like Treatment B, the order in which subjects approached dealers was random, and every dealer subject was approached by all four subgroups; in cases where the dealer purchased the good, I continued to have agents visit that dealer.

Column 1 in Table I contains a summary of the experimental design. In sum, Experiment I includes data gathered from 180 nondealers: 120 buying and 60 selling agents who varied across race, gender, and age. I therefore observed 240 outcomes in Treatment B since each buyer completes the exercise with a low and high reservation value. And, since each nondealer in Treatment S approaches 5 dealers, I observed 300 outcomes in Treatment S.[9] Of

---

[9] In practice, however, 14 of the 75 dealers in Treatment S showed no interest in the Griffey card—typically stating (to nondealer sellers across all four groups) that they "were not buying at the show" or "were not buying this card." I exclude these data from the statistical analyses below. Including these data and coding them as $0 does not change

the 120 buyers, 30 were white males, age 20-30; 30 were white females, age 20-30; 30 were nonwhite males age 20-30, and 30 were white males age 60 or older. The 60 nondealers in Treatment S were also equally split across the four subgroups.

On the dealer side of the market, I observe the behavior of 41 (75) dealers in Treatment B (Treatment S). It is important to note that 90 percent of my dealer subjects were white males—the other 10 percent were split approximately equally between white women and nonwhite men. These proportions are similar to what one observes in the dealer population in the geographic area of my experiment. On average, the 41 dealers in Treatment B were approached by approximately 6 agents (about half the dealers were approached by 4 (8) agents). In every case I was able to obtain important subject-specific information from the dealers, either via a survey they completed during an experiment in which they were a participant later at the show (complementary experiments described below) or through filling out a survey in exchange for a payment of $2 if they were unable or declined to participate in a complementary experiment. In both cases, the dealer filled out the survey near the end of the sportscard show and only after he had (unknowingly) participated in Experiment I.

A few design issues should be mentioned before proceeding to a discussion of the results. First, in pilot studies I did not observe systematic evidence that agents in any of the four membership groups favored a certain dealer type when attempting to sell or purchase wares, so I randomly chose dealers to have my subjects approach. Second, these data were gathered at large "regional" sportscard shows to avoid nondealer subjects having had previous interactions with the dealer subjects and to attenuate profit-maximizing dealers making inferences about the likelihood of future interactions, which could also cause statistical discrimination. As a check, I asked the nondealer subjects whether they had engaged in transactions with the dealer

---

the qualitative results reported below.

previously—only 3 (5) subjects in Treatment B (S) indicated they had.[10] Third, in Treatment B I observed that 2 dealers had multiple copies of the Griffey card on hand. Fourth, following the audit study literature, when making inquiries I spaced the visits by at least 30 minutes. Fifth, to avoid compromising the experiment, I had a monitor record bargaining time for each interaction from a distance.

Average participant summary statistics in the two treatments are summarized in Panel A of Table II. Participant characteristics within the same membership group are generally similar across Treatments B and S. And, participants across the four membership groups within a specific treatment are alike in many of the sampled characteristics. For example, average income (education) tends to be in the $30,000-39,999 (2-year college) range. Yet when I split subjects into income (below $20,000 and above $50,000) and education (high school or below and 4 year college or above) categories, I do observe differences in composition across the groups. And, there are differences in other variables: transaction intensity ranged from 5.27 transactions per month (white males 60+) to 7.53 transactions per month (white males age 20-30) and years of market experience was the lowest among women (6.73) and highest among older white men (12.67). Accordingly, in the conditional data analysis below I control for these factors.

II.A. Experiment I Results

Panel B of Table II contains summary statistics for Treatments B and S. Entries in Panel B are at the group level and include average initial and final offer and their standard errors, average bargaining length (in minutes), and number of executed transactions. Overall, the Treatment B data summary suggests that minorities received higher initial and final offer prices compared to white males age 20-30 (hereafter I will refer to white males age 20-30 as the

---

[10] Immediately after each dealer was approached the subject privately filled in responses for questions that pertained to that dealer (questions 6-8 of the survey).

"baseline" group). For example, females and older agents received initial offers that were more than 10 percent higher than those received by the baseline group. Data from Treatment S suggest similar insights, but a stark finding is that the level of discrimination is much more pronounced when agents are selling their wares: while the baseline group received an average initial (final) offer of $38.06 ($42.05), minorities received initial (final) offers in the $26-$29 ($33-$35) range. Average "trimmed" initial and final offers (excluding the 3 highest and 3 lowest offers) presented in Table II provide evidence that this pattern of discrimination is not caused by a few outliers.

Insights gained from Table II can be supplemented in a meaningful manner by exploring the prevalence of discrimination in the marketplace. In Table III, I match the data by dealer and provide the actual number of dealers that gave the minority groups an inferior, equivalent, and superior offer in each of the treatments. According to a series of tests of proportions, the proportion of minorities that receives inferior initial and final offers is significantly greater than the proportion that receives superior offers at the $p < .05$ level, providing evidence that the higher offer prices observed in Table II are not simply the result of a few dealers discriminating.

Empirical results in Tables II and III are suggestive that discrimination is prevalent in the marketplace, but they do not control for potentially important subject-specific characteristics that may be correlated with group membership, such as market experience, income, and education. In addition, they fail to make full use of the data since they do not exploit its panel nature. Note that multiple subjects visited the same dealer; thus it is important to determine the nature of the observed within-dealer variation in discriminatory behavior. To account for data dependencies, I estimate the following linear regression model:

(1) $$P_{ij} = X_{ij}\beta + \alpha_j + \varepsilon_{ij},$$

where P_{ij} is the initial (or final) offer price the ith nondealer received from dealer j; X_{ij} includes nondealer characteristics—years of market experience, transaction intensity, education and income dummy variables—as well as group membership dummy variables, and controls for i) whether the nondealer had previous interaction with the dealer and ii) whether a dealer had purchased the card already in the experiment (Treatment S).[11] α_j are dealer fixed effects, which transform each observation into a difference from the dealer-specific mean; ε_{ij} is the random error component, which includes non-dealer random effects to control for data dependencies.

Summary empirical estimates of equation (1) are contained in Table IV.[12] I include two distinct empirical specifications to provide insights into the robustness of the results. Specification 1 includes only the group membership indicators, dealer fixed effects, and buyer-specific random effects. Specification 2 augments this baseline model by adding years of market experience, transaction intensity, education and income category dummy variables, a dummy variable indicator for whether the nondealer had previous interaction with the dealer, and induced reservation price (and bargaining time for the final offer models). Perusal of the estimates in Table IV provides evidence that the pattern of results observed in Tables II and III is quite robust, and suggests:

Result 1: Initial offers to minorities are inferior.

In Treatment B, this result can be seen in columns 1-2 of Table IV. For example, results in column 2 suggest that women received an average offer that was approximately $10.90 more than their white male counterparts received. This difference in initial offer price is considerable—10 percent—and is statistically significant at conventional levels. Other results in column 2 suggest that older men are also targets of discrimination, as they received initial offers

---

[11] Since income and education are categorical variables I use a set of income and education dummy variables— seven income dummies and five education dummies.

[12] Similar regression results are obtained if I use the "trimmed" data set (excluding the 3 highest and 3 lowest

that were approximately 13 percent greater than offers received by the baseline majority group. While the evidence points to discrimination against nonwhite men as well, the point estimate in column 2 is not statistically significant at conventional levels.

In the case of Treatment S, evidence for Result 1 can be found in columns 5 and 6 of Table IV. When nondealers were sellers of the Griffey card, discrimination remains evident, but the level of discrimination is much more pronounced. Across the two empirical specifications, minorities received offers that were $8-$12 lower than initial offers received by the baseline group. In column 6, the observed differences map into roughly a 30% disparity in initial offer prices across minority and majority agents. These differences are not directly comparable to the literature on product market discrimination because they are, to my best knowledge, the first estimates concerning discrimination of sellers who are negotiating over prices of their wares.

Empirical estimates of equation (1) using final offers as the dependent variable suggest the following insight:

Result 2: Final offers to minorities are inferior, but via the bargaining process the differences are not as great as the initial offer discrepancies.

Much like Result 1, the first part of this finding is robust across the two specifications. In the case of Treatment B, Table IV shows that women and older men received final offers that were approximately 6-8 percent above the final offer received by the baseline group. Again, the difference in offer price is not significant at conventional levels for nonwhite men. Empirical estimates using data from Treatment S again show much stronger trends, and while the observed discrepancy is smaller than what is observed in the initial quoted offer prices, the difference between final offers is drastic: minorities received offers that were approximately 18%-20% less than what majority agents received. The observed differences are statistically significant at the $p < .05$ level for each membership group.

offers).

To provide evidence on the second part of Result 2, I estimated Specification 2, but as the regressand I used the difference of the initial and final offer prices. For both treatments each of the group membership dummy variables is statistically significant at the $p < .05$ level. The sole exception is the nonwhite dummy variable in the Treatment B data, which is the correct sign but not statistically significant at conventional levels.

While group membership is clearly important, the evidence in Table IV points to other aspects that also have an influence on the final quoted price. For example, years of market experience is a significant factor in both treatments: empirical estimates suggest that more experienced agents successfully negotiate lower buying prices. These comparative static findings concerning the effect of market experience provide a basis for further exploration into the distribution of final offers. Upon closer inspection of the final offer prices received across the various classes of bargainers, a further result follows:

> Result 3: Among experienced nondealer buyers, final offers are similar across minorities and majorities, but minorities commit greater temporal resources to arrive at the final offer.

Table V provides one cut of the data to illustrate Result 3. In Table V, I split agents by their level of market experience—if an agent has either i) a greater number of transactions per month *or* ii) a greater number of years of market experience than the average subject in his or her treatment type (i.e., more than 6.6 (7.8) transactions per month or 9.4 (9.1) years of market experience in Treatment B (S)), I label the agent "highly" experienced.

Panel A in Table V contains a summary of Treatment B data. The stark pattern in these data is the similarity in final quoted prices across the highly experienced agents in each membership group. While initial offers to the highly experienced agents continue to exhibit the discrimination that was observed in the overall data, after negotiation, minorities obtained prices that were very similar to the prices obtained by the majority group: $98.8 (baseline) versus

$99.9 (females), $103.6 (nonwhite), and $103.4 (older men). Estimating variants of equation (1) using these data yields insights consistent with no final offer price discrimination. Yet these gains by minorities are not free—in each case, they must commit greater temporal resources to achieve lower prices. For example, in comparison to the 2 minutes of bargaining time by the average experienced baseline group member, the average female, nonwhite, and age 60+ experienced agent committed 6.5, 3.6, and 3.6 minutes bargaining with dealers. These differences in bargaining time are significant at the $p < .05$ level using a Wilcoxon test for matched pairs (matched across dealers).

Considering data from lesser-experienced agents in Panel A paints a much different picture, as in these data the final offer discrepancies remain. Whereas the baseline group received a final offer of $102, women, nonwhite men, and older men received final average offers of roughly $112, $105, and $115. Estimating equation (1) using the final offer data suggests that these differences are statistically significant at the $p < .05$ level for women and older men, but not significant at conventional levels for nonwhite men. Combining these results with insights gained from the experienced agent data discussed above suggests that the bulk of dealer rents accrue when they strike a deal with naïve minority buying agents. This finding may make an important contribution to our understanding of why dealers, in general, rely on bilateral negotiations rather than posted prices.[13]

Insights gained from Result 3 naturally lead one to inquire into whether dealers display similar behavioral heterogeneity. To explore this issue, I estimated the linear regression model:

(2) $$Disc_j = V_j\beta + \varepsilon_j,$$

---

[13] Treatment S data are not as sharp as Treatment B data, but directionally they do suggest that a similar phenomenon occurs when nondealers are selling their wares.

where $Disc_j$ is the average level of dealer discrimination as measured by dealer j's average offer to baseline members minus dealer j's average offer to minority group members. V_j includes measures of dealer-specific experience. In particular, since many of the dealers were uncomfortable discussing their total number of transactions in a typical month, I included the typical number of monthly trades and years of market experience in V_j. I also included chronological age in V_j to ensure that the experience variables, especially years of market experience, were not merely tracking age effects.

Table VI provides summary estimation results that lead to the following insight:

Result 4: Experienced dealers exhibit a greater level of discrimination than their inexperienced counterparts.

Estimates in Table VI suggest that regardless of whether dealers are buying or selling, or whether they are making an initial or final offer, discrimination is more prevalent among experienced dealers, as measured by years of market experience. For example, every specification in Table VI suggests that years of market experience is an important determinant of the degree of discrimination: whether selling or buying goods, differences in offers are largest among experienced dealers, as suggested by the negative (positive) coefficient estimates in the Treatment B (Treatment S) models.

III. Complementary Experimental Designs and Results

Experiment I provides evidence that discrimination is present in the sportscard marketplace, but similar to the vast empirical literature these data in isolation cannot pinpoint the nature of discrimination. Under certain plausible scenarios, the results are consonant with each of the three theories discussed earlier: i) animus-based or taste-based discrimination, ii) differences in bargaining ability, and iii) statistical discrimination.

This section begins by summarizing an experiment that isolates animus-based, or taste-based, discrimination among dealers. I then turn to a summary of an empirical examination of

behavior in an induced-value experiment, in the spirit of Chamberlain [1948], in which I examine aspects of bargaining abilities across groups by observing outcomes of face-to-face continuous bilateral negotiations in a multilateral context. I proceed to an examination of data drawn from actual auctions run in the marketplace for a 1989 *Upper Deck* Ken Griffey Jr. PSA graded "9" baseball card (the identical card from Experiment I); this experiment allows me to estimate the underlying distribution of reservation values among the various consumer groups. This auction experiment is supplemented by a fourth field experiment that gathers data on dealers' perceptions of the reservation value distributions, which permits a more detailed examination of whether the discrimination observed in the marketplace is indeed based on dealer profit motives.

III.A. The Dictator Game

To more closely examine whether dealers have a "taste" for discrimination in the form of animus or bigotry, I employ a $5 dictator game. Unlike the popular two-player ultimatum bargaining game, in the dictator game strategic concerns are absent, as the proposer simply states what the split of the $5 will be and the responder has no veto power. The dealer decision in this experiment is, therefore, straightforward: decide how much, if any, of the $5 to transfer to their nondealer partner. The novelty in this experiment is that I inform dealers about an important characteristic of their partner: in the baseline Treatment WM, I inform the dealer that his partner is a white male nondealer, age 20-30. In Treatment WF, I inform the dealer that his partner is a white female nondealer who is 20-30 years old. Likewise, in Treatment NWM (WMM) I inform the dealer that his partner is a nonwhite male nondealer, age 20-30 (white male nondealer age 60+). Accordingly, any observed difference in rates of giving across treatments provides evidence of tastes for discrimination [Fershtman and Gneezy 2001].

I was careful to ensure that individual behavior was anonymous to avoid attenuating influences—election analysts remind us that exit polls considerably underestimate the number of votes a candidate such as David Duke receives in the actual election. To do so, I followed as closely as possible the experimental instructions of the "Double Blind 1" treatment in Hoffman et al. [1996]. In using this protocol, no one, including the experimenter or any subsequent observer of the data, could possibly know any particular subject's decision. This was achieved by using opaque envelopes, dummy envelopes, blank slips of paper, etc. More specifically, if the session included 10 dealers, I included 10 opaque envelopes—eight of the envelopes contained five $1 bills and five blank slips of paper; the other two envelopes contained ten blank slips of paper. Thus, if each of the eight dealers who received the envelopes containing money took all five dollars, there was no way to discern which of the eight dealers received the five dollars and which two dealers received the ten slips of paper. I was also careful to strictly follow this procedure in each treatment to ensure that social isolation would not change across treatments.

Before discussing the results, a few noteworthy items should be mentioned. First, in the actual implementation of the experiment, one-by-one dealers randomly chose their envelope from a box. After choosing an envelope, the dealer sat behind a large cardboard box and decided how many one dollar bills to keep and how many to place in the envelope. All bills taken were replaced by a blank slip of paper to ensure similar thickness of envelopes. Second, dealers never physically saw their partners; rather they were merely informed of their partner's race, gender, and age. Third, upon session completion, I randomly selected each dealer's partner who fit the specific criteria. This was done later at the same sportscard show and carried out in experiments unrelated to this study. The selected nondealer partner then received the envelope's contents (nondealers were never proposers).

Fourth, of the 116 dealers who participated in Experiment I, I was able to include 55 in

the dictator game experiment.[14] These 55 dealers comprised 5 of the dictator game sessions (i.e., they were the only participants included in these five sessions). Other aspects of the experimental design are summarized in column 2 of Table I. In each treatment, I include 30 dealers and 30 nondealers; hence each treatment has 60 subjects, yielding a sample size of 240 subjects in this experiment. Finally, I point the reader to column 1 of Table VII which summarizes predictions from the three theories. In sum, if the animus theory explains the data patterns observed in Experiment I, then in the dictator game dealers should be less generous towards minorities. The other two theories predict no differences in generosity.

Table VIII summarizes the data obtained from the four dictator treatments and Figure I uses the data to graphically depict the frequency distribution for each treatment. The data display similarities with the published literature [e.g., Hoffman et al. 1996], as they indicate that the average offer is in the range of 20-40 percent ($1.07-$1.90 of the $5), and the rate of positive offers is in the range of 50 percent. More importantly for the purposes herein, the results are not consonant with the animus theory—using a Wilcoxon rank-sum test, which tests the null hypothesis that the distributions are identical across treatment conditions, the only statistically significant difference I find is that white women receive greater levels of transfers than the three other cohorts (WF vs. WM: $z = 2.07$; WF vs. NWM: $z = 2.07$; WF vs. WMM: $z = 1.72$).

The result that dealers transfer higher amounts to women than to other men is consonant with findings in a laboratory experiment carried out in Sweden: Dufwenberg and Muren [2002] find that in dictator games men transfer more to women than to other men. Similarly, this result is in line with Eckel and Grossman's [1998] "chivalry" effect, where in ultimatum games men

---

[14] But note that my experimental approach precludes directly matching dealer behavior across Experiment I and the dictator game. The trade-off is clear: match data across experiments by compromising the dictator game (by allowing the experimenter to observe sensitive information) or do not match data across games by carrying out a dictator game with complete anonymity. In light of the sensitivity associated with discrimination, I opted to avoid confounding factors and chose to collect anonymous dictator game data, thereby ensuring that the test of animus is clear-cut. Yet I should note that I am able to match groups of dealers.

are more generous to women than to other men. More importantly, for our purposes, the key pieces of evidence from the dictator games suggest that:[15]

Result 5: Dealers do not exhibit noneconomic tastes for discrimination that systematically favor the majority group.

III.B. Chamberlain Markets

Evidence from the dictator game data is not consistent with the theory that dealer "tastes" for discrimination are driving the observed results in Experiment I. To provide insights into bargaining differences across the four groups, I make use of the experimental design of Chamberlain [1948], who over half a century ago executed what is believed to be the first experiment to test neoclassical competitive market theory. Within this market setting I examine two distinct treatments—Treatment Random and Treatment Unclear. Treatment Random is similar to the field experimental design in Experiment I, with a few major differences: i) dealers *know* that they are part of an experiment and ii) dealers know that buyers' reservation values are drawn *randomly.* The design of Treatment Unclear is identical to Treatment Random with one exception: instead of making it clear to dealers that buyer reservation values are drawn randomly, I left that aspect of the experimental design ambiguous (i.e., I did not inform dealers of the determination of buyer values). And, to provide dealers with a sense that consumers may actually depart the experiment with the cards that were bargained over, in Treatment Unclear I used a different set of sportscards each market period that had market value outside of the experiment (explained more fully below).

Given that these treatments are similar in spirit to Experiment I, I only briefly highlight the major aspects of the four step experimental design. In Step 1, a monitor approached dealers

---

[15] These insights are reinforced if I consider data from only the 55 dealers that participated in Experiment I— focusing only on these data, empirical results mirror those from the overall sample. Note that given Result 4 the animus theory makes a further prediction: more experienced dealers are relatively less generous towards minorities compared to lesser experienced dealers. While this would have been interesting to test, it is not feasible given the

at a sportscard show and inquired about their interest in participating in an experiment that would last about 60 minutes. To gather the nondealer subject pool, a monitor approached potential subjects entering the show and inquired about their level of interest in participating in an experiment that would last about 60 minutes. As soon as 12 dealers and 12 nondealers agreed to participate, Step 2 began. Note that the 12 nondealer subjects in any given session were drawn from any of the four groups—white males, age 20-30; nonwhite males, age 20-30; white females, age 20-30; and white males, age 60+.

In Step 2, monitors thoroughly explained the experimental rules. The experimental instructions were standard and taken from Davis and Holt [1993, pp. 47-55] with the necessary adjustments: first, all individuals were informed that they would receive a $10 participation fee upon completion of the experiment. And, to ensure that marketers would engage in a transaction at their reservation prices, I provided a $0.05 commission for each executed trade for both buyers and sellers. Second, nondealers (dealers) were informed that the experiment consisted of 5 periods and that they would be buyers (sellers) in the experiment. In each of 5 periods, I used induced values by providing each buyer (seller) with a "buyer's card" ("seller's card"), which contained a number, known only by that buyer (seller), representing the maximum price that he or she would be willing to pay for (sell) *one* unit of the commodity. Third, the monitor explained how earnings (beyond the participation and commission fees) were determined: for buyers, the difference between the contract price and the maximum reservation price determined market earnings. Likewise, sellers' earnings were determined by the difference between the actual contract price and the minimum reservation price. Several examples illustrated the irrationality associated with buying (selling) the commodity above (below) the induced value.

---

anonymity of dealers. I should note, however, that when I split the data by average experience in a session I do not find evidence in favor of this prediction.

Fourth, the homogeneous commodities used in Treatment Random were 1982 *Topps* Ben Oglivie baseball cards, on which I had artfully drawn a moustache, making the cards valueless outside of the experiment. In Treatment Unclear, to provide dealers with a sense that consumers may actually depart the experiment with the cards that were bargained over, i) I used a different set of sportscards each market period that had market value outside of the experiment ("non-mustachioed" cards) and ii) after executing a transaction, nondealer buyers actually kept the card in hand until the experiment was completed. Fifth, after completion of each contract, the exchange price was announced so that all buyers and sellers were made aware of the most recent transaction. Sixth, buyers and sellers engaged in two 5-minute practice periods to gain experience with the market.

In Step 3, subjects participated in the market. Each market session consisted of 5 market periods that lasted 10 minutes; after each 10-minute period, a monitor privately gathered with buyers and gave them a new buyer's card while a different monitor privately gave sellers a new seller's card. Step 4 concluded the experiment—after subjects completed a survey, they were paid their earnings in private (I used the survey contained in Appendix I with the necessary adjustments).[16]

Given that I executed 6 distinct five-period market sessions in each treatment, this experiment includes data drawn from 60 unique market periods. And, since buyers and sellers competed in only one treatment, this experiment included 288 subjects: 144 nondealers and 144 dealers. As summarized in column 3 of Table I, of the 72 consumers in Treatment Random, 24 were white males, age 20-30; 18 were white females, age 20-30; 15 were nonwhite males, age 20-30; and 15 were white males, age 60 and over. Thus, I observe 120 periods of data for white

---

[16] Buyer- and seller-induced values are taken from Davis and Holt [1993, pp. 14-15]. The efficient perfectly competitive prediction using this design yields $37 in rents per period, which occurs where competitive price theory

males, age 20-30; 90 periods of data for white females, age 20-30; 75 periods of data for nonwhite males, age 20-30; and 75 periods of data for white males, age 60 and over. Of the 72 consumers in Treatment Unclear, 25 were white males, age 20-30; 18 were white females, age 20-30; 14 were nonwhite males, age 20-30; and 15 were white males, age 60 and over. The vast majority of dealers (over 90%) were white males. And, I should note that 22 of the dealers were participants in Experiment I.[17]

Prior to discussing the empirical results, I point the reader to predictions from the three theories in column 2 of Table VII. Each of the three theories has unequivocal predictions across several aspects of the data, and as I progress through the results summary I will highlight whether evidence is consistent with each of the theories.

The raw data summary is presented in Table IX. To eliminate much of the randomness due to the variation in induced values, in Table IX I present summary statistics of the average "lost surplus" in each cell. To derive the average lost surplus (lost surplus = predicted surplus – actual surplus), I first computed the average for each market participant over the five rounds within each session. Predicted surplus is what would have happened had competitive predictions prevailed throughout the market session—I use a price of $13.50 when computing predicted surplus.[18] I then computed the average and standard error of the individual averages.

Given that years of market experience was an important predictor of whether, and to what extent, dealers engaged in discrimination (see Table VI), I split the dealer data along the first moment of the years of market experience variable: if a dealer had a greater number of years of market experience than the average dealer in this experiment (˜ 15 years), then he was labeled

predicts a tendency for the static price/quantity equilibrium of Price = $13.00-$14.00 and Quantity = 7 to be reached, which is the extreme point of the intersection of the buyer and supplier rent areas.

[17] Since Treatment Unclear was run after Experiment I was completed, all of these 22 dealers were participants in Treatment Random.

[18] All buyers (sellers) given an induced value below (above) $13.50 received a predicted surplus of $0.

"experienced"; otherwise I labeled the dealer "inexperienced." I closely follow this approach for nondealers as well: a non-dealer is labeled experienced if he has 9 or more years of market experience (9 is the approximate mean number of years of market experience).[19] The raw data point to a first finding:

> Result 6: Majority buyers outperform minority buyers when dealers do not know that reservation values are determined randomly.

This result can be seen in rows 2-5 of columns 4-6, where the data suggest that the baseline majority group outperforms the minority groups in every case. To provide statistical evidence of Result 6, I estimated an ordinary least squares regression model using average individual lost surplus as the regressand and the three membership dichotomous variables as well as five dummy variable indicators for experimental session (note there were six sessions) as regressors. I find that I can reject the homogeneity null for each of the three minority categories in the pooled data at the $p < .05$ level; directionally the finer categories of experienced and inexperienced agents point to similar inferences, but the small sample sizes make it difficult to detect the differences statistically. These results reveal that even when dealers are aware that they are taking part in an experiment, and that their behavior is recorded, discriminatory behavior is not attenuated.

One can provide further evidence into the predictive power of the theories by examining data from Treatment Random, where a second insight is obtained:

> Result 7: Minority and majority buyers perform similarly when dealers know that reservation values are determined randomly.

Evidence for this result can be found from a comparison of data in rows 2-5 of columns 1-3, which suggest that lost surplus is similar across nondealers in the four membership groups. Using an identical regression model as described above, I now find that none of the membership

---

[19] Results are similar if I follow the approach outlined above (see Table V) to split nondealers into groups.

dummy variables ever gain statistical significance at even the $p < .25$ level, suggesting that there is no difference in lost surplus across the groups. Overall, given that dealer information on reservation values is the only design aspect that changes between Treatment Unclear and Treatment Random, Results 6 and 7 are not in line with the animus, bargaining ability, or dealers *unknowingly* performing statistical discrimination theories, but they are consonant with the notion of dealers *knowingly* statistically discriminating.

Turning to the dealer data summary, I find that:

Result 8: Experienced dealers lose less surplus than inexperienced dealers.

The raw data summary in columns 2-3 and 5-6 in row 1 of Table IX point to this finding, and via estimation of an ordinary least squares regression model of average individual lost surplus regressed on a dummy variable for whether the dealer is experienced and the five session indicators, I find that in the random treatment experienced dealers lost less surplus than inexperienced dealers at the $p < .05$ level. While this finding is significant in the random treatment data, using similar empirical methods I cannot reject this same null hypothesis in the Treatment Unclear data.

The random treatment data are therefore at odds with the animus and dealers unknowingly performing statistical discrimination theories. Additionally, the empirical "null" result for the unclear treatment data does not match any of the theoretical predictions, but directionally it is in line with the bargaining theory and at odds with the other three theories. A comparison of the performance of experienced dealers across the two treatments, however, yields an interesting insight that is consistent with dealers knowingly practicing statistical discrimination: experienced dealers in Treatment Unclear lost more surplus than experienced dealers in Treatment Random (0.32 vs. 0.17). Yet I should stress that this difference is not statistically significant at conventional levels using a t-test of means.

III.C. Reservation Values

Insights from the dictator games and the bilateral negotiating markets provide key pieces of evidence hinting at the nature of discrimination observed in Experiment I, but what remains necessary is to understand differences in reservation value distributions across agents. In this sense, it is important to pinpoint an underlying theoretical framework of statistical discrimination to place the analysis in proper perspective. The structure of such a mechanism has been the subject of intense research for decades and has resulted in a very rich literature. For the purposes herein, work in the price discrimination and bargaining literatures [Salop and Stiglitz 1977; Cramton 1991] predicts that conditional on equivalent mean reservation values, dealers will quote a higher (lower) price to buying (selling) agents who are from groups that have reservation values that are more widely distributed. Intuitively, this insight follows from dealers attempting to secure deals with high (low) reservation value agents when selling (buying) their wares.[20]

This section aims to provide direct evidence of such reservation value distributions by gathering willingness to pay (WTP) and willingness to accept (WTA) measures of value for a 1989 *Upper Deck* Ken Griffey Jr. PSA graded "9" baseball card—the identical cards that were used in Experiment I. In an attempt to complete the simplest possible exercise to gather reservation values, I chose an allocation mechanism—Vickrey's 2nd price auction—that has proven straightforward in other field experiments [List, 2001]. Each subject's auction experience typically followed three steps: (1) inspection of the good/learning the auction rules, (2) actual bid (offer), and (3) survey completion and debriefing.

---

[20] One can construct an intuitive bargaining model to motivate the role of reservation value distributions by assuming a buyer has a private reservation value, w, distributed according to pdf $f(w)$ with a CDF $F(w)$. The seller has marginal cost, c, and announces a final take-it-or-leave-it price p. The seller's problem is therefore to maximize $(p - c)$ Prob $(w = p) = (p - c)(1 - F(p))$, which has a first order condition of $p = c + (1 - F(p))/f(p)$. The first order condition shows that as $f(p)$ decreases (more widespread distribution), p is higher. For an analytical survey of theoretical models of discrimination, see Fryer [2001].

In Step 1 of the WTP treatment, much like Treatment B in Experiment I, a potential subject approached the monitor's table and inquired about the sale of the sportscard displayed on the table. The monitor then invited the potential subject to participate in an auction for the sportscard that would take about five minutes. The participant could pick up and visually examine the card. The card was sealed with the appropriate grade clearly marked on the cardholder. The monitor worked one-on-one with the participant and no time limit was imposed on his or her inspection of the card or learning the auction rules (experimental instructions are available upon request). In Step 1 of the WTA treatment, similar to Treatment S of Experiment I, the monitor approached subjects entering the sportscard show and inquired into whether the agent was planning to sell goods at the sportscard show. If the agent answered in the affirmative, the monitor examined the agent's wares to determine if he or she was "eligible" to be in the auction. The monitor then invited eligible subjects to participate in an auction for the sportscard that would take about five minutes.

In Step 2 of the WTP treatment each participant privately wrote a bid on the bidding sheet and placed it in an opaque box. The monitor informed the participant that the bid would not be opened until after the show and that all bids would be destroyed when the research project was completed. In Step 2 of the WTA treatment, the monitor informed the subject that one of the participants, determined randomly, would be given the Griffey card *free of charge*. If this lucky participant was the lowest offerer in the auction, the participant would be paid the second lowest offer in exchange for the Griffey card. If the lucky participant was not the lowest offerer, he would receive the Griffey card in the mail, postage paid. After the rules were fully understood, the subject completed the recording sheet by stating his minimum WTA to sell the Griffey card.

Finally, to conclude, in Step 3 the subject completed a survey similar to the one found in Appendix I. While completing the survey, the monitor explained that if the participant won the auction, she would be contacted by telephone within three days after the show. In the WTP treatment, upon receipt of payment for the card, the card would be sent to her via courier, postage free. In the WTA treatment, upon receipt of the card, the monies would be sent to her via courier, postage free.

A few noteworthy aspects of the experimental design merit further consideration. First, no subjects participated in more than one treatment. Second, since I am not testing the incentive compatibility of the institution, and want to avoid any excess noise, I inform the subjects that it is in their best interest to bid (offer) their true reservation value in the auction. I reinforce this notion via several examples that illustrate the optimal strategy of truth telling in the two mechanisms. Third, the treatment was changed at the top of each hour, so subjects' treatment type was based on the time they attended the card show.

As summarized in column 4 of Table I, I included 30 consumers from each of the four groups in both the WTP and WTA auctions. Thus, these auction treatments include 120 bidders and 120 offerers—the average participant across the four groups share similar qualities with those who participated in the experiments above. Again, I point the reader to Table VII (column 3), which summarizes the theoretical predictions from the three theories.

Panel A in Table X and Figures II and III summarize the experimental auction data.[21] The pattern of results exhibited leads to an interesting insight:

---

[21] These results have intrinsic importance outside of testing for reservation value differences. For example, in this case prospect theory [e.g., Kahneman and Tversky 1979] predicts that WTA > WTP. Due to income (shifts in the indifference curve) and substitution (curvature of the indifference curve) effects, neoclassical theory also predicts value divergences. In practice, it would appear that the substitution effect would hold more prominence in shaping the value disparity than would the income effect. The intuition behind this thought is contained in the Engel aggregation condition, which requires that the income elasticities of demand sum to one. While this certainly limits the size of the income effect associated with the card, the substitution effect can range from zero to infinity. See List [2003a] for tests of these conjectures in the field.

Result 9: Minority reservation price distributions are more widely dispersed than the majority reservation price distribution.

As can be seen in Table X, in both the WTP and WTA auctions the first moments of the reservation values are roughly equivalent across membership groups, though the second moments are consistently larger among the minority groups compared to the baseline. In the WTP case, a robust f-test suggests that in each comparison these variance differences are statistically significant at the $p < .08$ level. While the variances of the WTA measures of value display similar patterns, only the older white males have a larger variance than the baseline group at the $p < .05$ level.

A finding that the reservation values of minorities are more disparate is important, but whether the observed differences in reservation value distributions map into minorities providing dealers with more lucrative bargaining opportunities is of utmost importance. To illustrate differences in the WTP distributions across the four groups, consider the fraction of buyers who bid more than $100: Figure II reveals that only 1 of the 30 baseline subjects bid more than $100 for the Griffey card, whereas 6 of 30 females bid over $100, with two of these females bidding more than $125. Similar insights are obtained if I compare the baseline group to nonwhite and older men. Using a test of proportions, I find that significantly more of the three minority groups fall in the upper tail of the value distribution (bid over $100) at the $p < .05$ level.

Considering the WTA distributions in Figure III, and focusing on nondealers who offered the Griffey card for $50 or less, similar insights are obtained: considerably fewer baseline subjects offer in the lower portion of the value distribution. Only 3 of 30 baseline members offer to sell the Griffey card for $50 or less whereas 10 of the 30 female offerers quoted an asking price of $50 or less. These two proportions are significantly different at the $p < .05$ level. Directionally, results are similar when one compares the baseline group to nonwhite and older

men: 7 and 4 of these subject types offered the Griffey card for less than $50, though these figures are not statistically different from the baseline at conventional levels.

Overall, data from the auction treatments provide evidence that important parts of the reservation value distributions are different across the four cohorts: in comparison with the baseline cohort, reservation value distributions among minorities tend to have "fatter" WTP (WTA) upper (lower) tails. This finding is consonant with the statistical discrimination theory.

III.D. Dealers' Perceptions

The actual distribution of values observed across the four cohorts certainly paints a compelling picture of the underlying causes of the observed discrimination in Experiment I, yet it is dealers' perceptions of these distributions that drive behavior. This section explores whether dealers accurately perceive the value distributions described above. To test whether dealers accurately predict reservation value distributions, I take data plots similar to Figures IV and V to the field and run an experiment to determine whether, and to what extent, dealers recognize the actual value distributions. The experimental design, which includes 60 dealers, as summarized in column 4 of Table I, is straightforward and proceeds in four steps: (1) consideration of the invitation to participate, (2) learning the rules, (3) participation, and (4) conclusion of the experiment and exit interview.

In Step 1, the monitor approached a dealer at a sportscard show and inquired about his interest in participating in an experiment that would take about 5 minutes. If the dealer agreed, the monitor explained the experiment: dealers would be shown various value distributions, or "consumer values," from a recent auction of a 1989 *Upper Deck* Ken Griffey Jr. PSA graded "9" baseball card—the identical cards that were used in the above experiments. The dealer would be shown two distributions at a time and asked to choose which subgroup each particular distribution was from: for example, in one treatment, the monitor showed Figure IV and the

dealer decided which bid values were from white males, age 20-30, and which bid values were from white males, age 60+. After answering this question, the monitor presented the same dealer with Figure V and asked the dealer to choose which "ask" values were from white males, age 20-30, and which "ask" values were from white females, age 20-30.

Given that there are four WTP and four WTA distributions, and that I always used the baseline group's WTP and WTA distributions in every treatment, there are nine possible comparisons, which the monitor rotated randomly in the experiment. To ensure that every dealer saw any given distribution only once, I had each dealer make one decision in the WTP treatment and one decision in the WTA treatment. Each dealer, therefore, had the opportunity to perform the matching exercise twice; and the monitor informed the dealer that he would receive $2 for each correctly matched pair. Thus, each dealer had an opportunity to earn $4 in the 5-minute experiment. The dealer was not informed of correct responses until he completed both choices.

In Step 3 the dealer made two choices, after which he filled out a survey (similar to Appendix I) and received his payment in private. In total, I observed the choices of 60 dealers; thus I observed 120 decisions. Of the 60 dealers, 23 of them were included in Experiment I; again, these 23 data points are similar to the others so I pool them.

Column 3 in Table VII provides the theoretical predictions for this final experiment: while the animus and bargaining theories make no predictions, the statistical discrimination theory conjectures that dealers are aware of the reservation price distributions, but that more experienced dealers have a keener sense of the differences.

Panel B of Table X summarizes the experimental dealer perception data. The panel provides summary statistics across the three comparisons of the WTP and WTA treatments and leads to the following insight:

Result 10: Dealers recognize differences in reservation price distributions, but experienced dealers are more aware of the differences.

The first part of this result can be seen vividly in the raw data. For example, consider when dealers are presented with the baseline and white female distributions: column 1, row 1 in Panel B indicates that 15 of 20 dealers (75 percent) correctly matched the reservation value distributions. Perusal of the other results in column 1 suggests the correct percentages are all greater than 50 percent, where 50 percent is the expected outcome if dealers were making choices randomly. Indeed, considering the WTP and WTA comparisons in isolation, a t-test of proportions suggests that in both cases the correct percentage is significantly greater than 50 percent at the $p < .05$ level.

In accord with the analysis of the dictator game data, in Panel B I also split the data along the first moment of the years of market experience variable. Under this classification scheme, results in column 2 (3), row 1, for example, indicate that 8 of 12 (7 of 8) inexperienced (experienced) dealers correctly matched the reservation value distributions when they were presented with the baseline and white female distributions. In aggregate, 21 of 28 experienced dealers correctly matched (75 percent), whereas only 56 percent of inexperienced dealers correctly matched (18 of 32). The match rate for experienced (inexperienced) dealers is (not) significantly different from 50 percent at the $p < .05$ level using a test of proportions. Furthermore, using a similar statistical test I find that the percentage of correct matches is significantly greater among experienced dealers than inexperienced dealers at the $p < .10$ level.

The numbers tend to increase when dealers are examining the offer distributions: in these cases experienced (inexperienced) dealers correctly match 82 (59) percent of the distributions, a difference that is significant at the $p < .05$ level.[22] Both dealer types are able to compare and contrast the baseline group with women and nonwhites, but in the case of the baseline and older

---

[22] Similar to the WTP data, the match rate for experienced (inexperienced) dealers is (not) significantly different from 50 percent at the $p < .05$ level using a test of proportions.

male agents neither dealer type performs well. This is probably due to the similarities in WTA distributions across these cohorts (see Figure III). Overall, data from the perception experiment tell a story that is consistent with the statistical discrimination theory. Combining insights from the perception experiment with results from the previous experiments provides a persuasive case that dealers knowingly practice statistical discrimination in this marketplace.

IV. Concluding Remarks

Before one can even begin to discuss social policies to address discrimination, it is critical to understand the causes of the underlying preferential treatment that certain groups receive. In this study, I depart from a traditional empirical investigation in an attempt to provide a framework to advance our understanding of the nature of discrimination in the marketplace. The study begins by presenting experimental evidence that is consistent with minorities being discriminated against: women, nonwhites, and older agents receive initial offers that are inferior to those received by their majority counterparts. While experienced minority buying agents are able to obtain final offers that are statistically equivalent to the final offers received by their majority counterparts, they must commit greater temporal resources to arrive at the final offer.

While these insights are interesting in their own right, one may consider the novelty of the study to be the combination of these insights with those drawn from four complementary field experiments, which each indicate that the observed discrimination is consistent with the notion of profit maximization: dealers in the marketplace are using their knowledge on reservation values to engage in the most profitable transactions. This finding holds whether the dealers are selling their wares or buying goods from nondealers, yet the level of discrimination is much larger when dealers are in the buying role. And, whether buying or selling, the degree of discrimination is found to be at its apex among the most experienced dealers.

As with any research endeavor, this study undoubtedly raises more issues than it settles. For example, to what extent does market structure influence social discrimination? In one sense, by varying the rules within the marketplace, one could examine discrimination in explicit collusive scenarios, which would provide insights into customer markets heretofore unexplored in the literature. From a much different research perspective, by varying the demand parameters in the Chamberlain experiment, one could gain considerable insights about the bargaining role and its influence on social discrimination. I leave discussion of these issues for another occasion.

References

Akerlof, George A., "The Theory of Social Custom, of Which Unemployment may be One Consequence," *Quarterly Journal of Economics*, IXIV (1980), 749-775.

Altonji, Joseph G., and Rebecca Blank, "Race and Gender in the Labor Market," in *Handbook of Labor Economics*, O. Ashenfelter and D. Card, eds. (Vol. 3C, Elsevier Science B.V., 3143-3259, 1999).

Altonji, Joseph G., and Charles Pierret, "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics*, CXVI (2001): pp. 293-312.

Arrow, Kenneth. "The Theory of Discrimination," in *Discrimination in Labor Markets*, O. Ashenfelter and A. Rees, eds. (Princeton, NJ: Princeton University Press, 1972).

Ayres, Ian, and Peter Siegelman, "Gender and Race Discrimination in Bargaining for a New Car," *American Economic Review*, LXXXV (1995), 304-321.

Becker, Gary, *The Economics of Discrimination*, 2nd ed., (Chicago: University of Chicago Press, 1975).

Becker, Gary, "The Evidence Against Banks does not Prove Bias," *Business Week*, April 19, 1993, 18.

Bertrand Marianne, and Sendhil Mullainthan, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," NBER working paper #9873, July 2003.

Camerer, Colin, *Behavioral Game Theory: Experiments on Strategic Interaction*, (Princeton University Press, 2002).

Chamberlain Edward H., "An Experimental Imperfect Market," *Journal of Political Economy*, LVI (1948), 95-108.

Cramton, Peter C. "Dynamic Bargaining with Transactions Costs," *Management Science*, XXXVII (1991), 1221-1233.

Davis, Douglas, and Charles Holt, *Experimental Economics,* (Princeton: Princeton University Press, 1993).

Dufwenberg, Martin, and A. Muren, "Discrimination by Gender and Social Distance," Stockholm University working paper, 2002.

Eckel, Catherine C., and Phillip J. Grossman, "Are Women Less Selfish than Men? Evidence from Dictator Experiments," *Economic Journal*, CVIII (1998), 726-735.

Fershtman, Chaim, and Uri Gneezy, "Discrimination in a Segmented Society: An Experimental Approach," *Quarterly Journal of Economics*, CXVI (2001), 351-377.

Fryer, Roland, "Economists' Models of Discrimination: An Analytical Survey," University of Chicago working paper, 2001.

Goldberg, Pinelopi, "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, CIV (1996), 622-654.

Goldin, Claudia, and Cecilia Rouse, "Orchestrating Impartiality: the Impact of 'Blind' Auditions on Female Musicians," *American Economic Review*, IC (2000), 715-741.

Heckman, James J., "Detecting Discrimination," *Journal of Economic Perspectives*, XII (1998), 101-116.

Heckman, James J., and Peter Siegelman, "The Urban Institute Audit Studies: Their Methods and Findings," in *Clear and Convincing Evidence: Measurement of Discrimination in America*, M. Fix and R. Struyk, eds. (Washington, D.C.: The Urban Institute Press, 1993).

Hoffman, Elizabeth; Kevin McCabe, and Vernon L. Smith, "Social Distance and Other-Regarding Behavior in Dictator Games." *American Economic Review*, LXXXVI (1996), 653-660.

Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, ILVII (1979), 263-291.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to Opportunity in Boston: Early Results of a Randomize Mobility Experiment," *Quarterly Journal of Economics*, CXVI (2001), 607-654.

List, John A., "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards," *American Economic Review,* ICI (2001), 1498-1507.

List, John A., "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, CXVIII (2003a), 41-71.

List, John A. "Are the Disabled Discriminated Against in Product Markets?" University of Maryland working paper, 2003b.

List, John A., and David Lucking-Reiley, "Seed Money Matters: Experimental Evidence from a University Capital Fundraising Campaign," *Journal of Political Economy*, CX (2002), 215-233.

Mobius, Markus, and Tanya Rosenblatt, "Why Beauty Matters," Harvard University working paper, 2003.

Neumark, David, R. Bank, and K. van Nort, "Sex Discrimination in Restaurant Hiring: An Audit Study," *Quarterly Journal of Economics,* CXI (1996), 915-941.

Phelps, Edmund. "The Statistical Theory of Racism and Sexism," *American Economic Review*, LXII (1972), 659-661.

Riach, Peter A., and Judy Rich, "Field Experiments of Discrimination in the Market Place," *Economic Journal*, CXII (2002), F480-F518.

Salop, Steven, and Joseph E. Stiglitz, "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion," *Review of Economic Studies*, ILIV (1977), 493-510.

Yinger, John, "Evidence on Discrimination in Consumer Markets," *Journal of Economic Perspectives*, XII (1998), 23-40.

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

1. How long have you been active in the sportscard and memorabilia market? _____yrs

2. Approximately how many transactions (trades, buys, and sells) do you make in a typical month?

Trades:_____ Buys:_____Sales:_____Total:_____.

3. Are you a sportscard or sports memorabilia professional dealer?_____

4. Gender: 1) Male   2) Female

5. Age _____   Date of Birth _____

1st dealer:

6. What was the initial offer made by the dealer?_____

7. What was the final offer made by the dealer?_____

8. Have you ever had interaction with that dealer?_____

2nd dealer:

6a. What was the initial offer made by the dealer?_____

7a. What was the final offer made by the dealer?_____

8a. Have you ever had interaction with that dealer?_____

9. 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

10. What is your approximate yearly income from all sources, before taxes?

  1) Less than $10,000   5) $40,000 to $49,999

  2) $10,000 to $19,999   6) $50,000 to $74,999

  3) $20,000 to $29,999   7) $75,000 to $99,999

  4) $30,000 to $39,999   8) $100,000 or over

11. Approximate height and weight: _____.

University of Maryland and NBER

TABLE I

EXPERIMENTAL DESIGNS

Experiment I	Dictator Game Experiment	Chamberlain Experiment	Reservation Value Experiment
Treatment B: Nondealers as buyers	Nondealers as recipients	Treatment Random: Nondealers as buyers	Treatment WTP: Nondealers as buyers
• 30 white males, age 20-30 • 30 white females, age 20-30 • 30 nonwhite males, age 20-30 • 30 white males, age 60 and older	• 30 white males, age 20-30 • 30 white females, age 20-30 • 30 nonwhite males, age 20-30 • 30 white males, age 60 and older	• 24 white males, age 20-30 • 18 white females, age 20-30 • 15 nonwhite males, age 20-30 • 15 white males, age 60 and older	• 30 white males, age 20-30 • 30 white females, age 20-30 • 30 nonwhite males, age 20-30 • 30 white males, age 60 and older
• 41 dealers	• 120 dealers	• 72 dealers	• 60 dealers (perceptions)
Treatment S: Nondealers as sellers:	---	Treatment Unclear: Nondealers as buyers	Treatment WTA: Nondealers as sellers:
• 15 white males, age 20-30 • 15 white females, age 20-30 • 15 nonwhite males, age 20-30 • 15 white males, age 60 and older		• 25 white males, age 20-30 • 18 white females, age 20-30 • 14 nonwhite males, age 20-30 • 15 white males, age 60 and older	• 30 white males, age 20-30 • 30 white females, age 20-30 • 30 nonwhite males, age 20-30 • 30 white males, age 60 and older
• 75 dealers		• 72 dealers	• 60 dealers (perceptions)

Each column represents a distinct type of experiment. For example, the top panel of column 1 denotes that I included 30 subjects from each subgroup as buyers in Experiment I. Nondealer subjects participated in only one experiment. As discussed in the text, some dealers were included in multiple experiments.

TABLE II

PARTICIPANT CHARACTERISTICS AND BARGAINING OUTCOMES—EXPERIMENT I

Treatment	White Males Age 20-30 Mean (Std. dev.)		White Females Age 20-30 Mean (Std. dev.)		Nonwhite Males Age 20-30 Mean (Std. dev.)		White Males Age 60+ Mean (Std. dev.)	
	B	**S**	**B**	**S**	**B**	**S**	**B**	**S**
Panel A. Characteristics								
Transaction	7.53	6.73	6.67	8.06	6.37	7.93	5.27	8.40
intensity	(10.0)	(4.3)	(6.5)	(4.9)	(5.3)	(3.8)	(3.6)	(5.9)
Yrs. of market	10.00	8.40	6.73	6.67	8.37	7.70	12.67	13.73
experience	(12.5)	(6.2)	(6.7)	(5.2)	(5.9)	(5.0)	(9.83)	(8.7)
Income	4.97	5.27	4.70	4.60	4.40	4.53	4.70	4.60
	(2.0)	(1.5)	(2.2)	(1.7)	(1.9)	(1.4)	(2.1)	(2.2)
< $20,000	13%	0%	17%	7%	20%	13%	17%	27%
= $50,000	40%	33%	43%	27%	30%	33%	30%	40%
Education	3.67	4.26	3.27	3.26	3.20	3.80	3.67	3.73
	(1.3)	(1.2)	(1.5)	(1.3)	(1.2)	(1.2)	(1.5)	(1.3)
= HS	17%	13%	50%	40%	33%	20%	33%	27%
= 4-yr. college	23%	53%	20%	20%	13%	33%	27%	33%
Panel B. Bargaining Summary								
	Mean (s.e.)		Mean (s.e.)		Mean (s.e.)		Mean (s.e.)	
Initial offer	$108.75	$38.06	$119.41	$26.64	$113.50	$26.47	$120.58	$29.34
	(2.5)	(2.8)	(3.3)	(2.2)	(3.6)	(2.3)	(4.7)	(2.5)
Final offer	$100.38	$42.05	$106.98	$33.99	$104.25	$33.52	$107.33	$35.16
	(2.1)	(3.2)	(2.7)	(2.8)	(2.6)	(2.6)	(3.3)	(2.6)
"Trimmed"	$107.31	$38.12	$117.31	$25.45	$110.65	$25.36	$116.20	$28.45
initial (final)	($100.05)	($41.73)	($105.62)	($33.15)	(102.96)	($32.91)	($104.44)	($34.73)
offer								
Bargaining	1.97	1.69	4.08	3.13	3.49	2.26	3.02	2.06
time (minutes)	(0.4)	(0.3)	(0.7)	(0.6)	(0.6)	(0.4)	(0.6)	(0.4)
Transactions	3	5	1	1	2	0	1	1
N	60	61	60	61	60	61	60	61

a. *Transaction intensity* represents the number of transactions made in a typical month.

b. *Yrs. of market experience* denotes number of years that the subject has been active in the market.

c. *Income* denotes categorical variable (1-8): 1) Less than $10,000, 2) $10,000 to $19,999, 3) $20,000 to $29,999, 4) $30,000 to $39,999, 5) $40,000 to $49,999, 6) $50,000 to $74,999, 7) $75,000 to $99,999 8) $100,000 or over.

d. *Education* denotes categorical variable 1) Eighth grade or less, 2) High School 3) 2-Year College, 4) Other Post-High School, 5) 4-Year College, 6) Graduate School Education.

e. *Initial* and *final offers* represent the offers actually received by the bargainers to begin and end the session. *"Trimmed" offers* are computed after excluding the three lowest and three highest offers.

f. *Bargaining time* is number of minutes of actual bargaining time that the agent negotiated. The monitor measured this during the experiment.

TABLE III

SUMMARY STATISTICS FOR EXPERIMENT I: NUMBER OF DISCRIMINATING DEALERS

	Initial Offer			Final Offer		
	Females	Nonwhite	Older Males	Females	Nonwhite	Older Males
Panel A: Treatment B						
Offer > Baseline Offer	25	14	21	19	15	14
Offer = Baseline Offer	14	22	18	18	20	18
Offer < Baseline Offer	2	5	2	4	6	4
Panel B: Treatment S						
Offer > Baseline Offer	8	9	7	8	7	4
Offer = Baseline Offer	16	16	24	25	24	33
Offer < Baseline Offer	37	36	30	28	30	24

a. Entries represent the actual number of dealers that fell in each behavioral category. For example, in row 1 column 1, an entry of "25" suggests that 25 dealers quoted women higher initial selling prices than they quoted the baseline, male age 20-30 agent.
b. Baseline offer represents offer to white males, age 20-30.
c. Treatment B (S) has dealers (nondealers) selling the Griffey card.

TABLE IV

REGRESSION RESULTS: EXPERIMENT I

| Variable | Treatment B | | | | Treatment S | | | |
| | Initial Offer | | Final Offer | | Initial Offer | | Final Offer | |
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Female *(age 20-30)*	10.8** (2.8)	10.9** (3.4)	6.6** (2.7)	6.2** (3.0)	-11.8** (2.1)	-12.3** (1.7)	-8.2** (1.8)	-8.5** (1.7)
Nonwhite *(age 20-30)*	4.9* (2.8)	4.9 (3.3)	3.7 (2.7)	3.3 (2.8)	-11.9** (2.1)	-10.8** (1.8)	-8.7** (1.8)	-6.8** (1.7)
Male *(age 60+)*	11.9** (2.8)	13.6** (3.4)	6.8** (2.7)	8.6** (2.9)	-8.8** (2.1)	-11.3** (1.9)	-7.0** (1.8)	-7.1** (1.8)
Transaction *intensity*	---	0.12 (0.18)	---	-0.13 (0.15)	---	-0.09 (0.15)	---	-0.04 (0.2)
Yrs. of market *experience*	---	-0.17 (0.14)	---	-0.28** (0.12)	---	0.38** (0.12)	---	0.25** (0.12)
Bargaining *Time*	---	---	---	-0.36 (0.24)	---	---	---	0.35* (0.20)
Dealer *Fixed Effects*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Buyer *Random Effects*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	240	240	240	240	244	244	244	244

a. Dependent variable is either the initial or final dealer quoted price to sell the Griffey card. Education and income category dummy variables as well as controls for i) reservation price (only Treatment B), ii) whether the nondealer had previous interaction with the dealer, and iii) whether a dealer had purchased the card already in the experiment (only Treatment S) are included in Specification 2. Coefficient estimates of these variables are available upon request.
b. Standard errors are in parentheses beneath coefficient estimates; ** denotes significant at the $p < .05$ level. * denotes significant at the $p < .10$ level.

TABLE V

EXPERIENCE LEVELS AND BARGAINING OUTCOMES : EXPERIMENT I

Exp. Level	White Males Age 20-30 Mean (Std. error)		White Females Age 20-30 Mean (Std. error)		Nonwhite Males Age 20-30 Mean (Std. error)		White Males Age 60+ Mean (Std. error)	
	High	Low	High	Low	High	Low	High	Low
Panel A. Treatment B								
Initial Offer	$109.2	$108.3	$116.7	$121.5	$113.3	$113.7	$119.1	$123.5
	(3.8)	(3.1)	(4.0)	(5.0)	(4.6)	(5.9)	(4.9)	(10.2)
Final Offer	$98.8	$102.0	$99.9	$112.4	$103.6	$105.2	$103.4	$115.3
	(2.8)	(3.1)	(4.0)	(3.9)	(3.1)	(4.5)	(2.8)	(8.2)
Bargaining time (minutes)	2.1 (0.5)	1.8 (0.5)	6.5 (1.2)	2.2 (0.6)	3.6 (0.8)	3.3 (1.1)	3.6 (0.8)	1.8 (0.7)
N	30	30	26	34	36	24	40	20
Panel B. Treatment S								
Initial Offer	$31.9	$29.5	$21.5	$21.8	$23.8	$18.2	$24.0	$23.1
	(3.9)	(4.1)	(2.8)	(3.5)	(3.0)	(3.3)	(2.7)	(5.6)
Final Offer	$35.1	$32.8	$28.5	$26.4	$28.8	$25.0	$28.4	$29.3
	(4.4)	(4.6)	(3.8)	(3.9)	(3.4)	(4.1)	(3.0)	(5.9)
Bargaining time (minutes)	1.7 (0.4)	1.7 (0.4)	3.6 (0.9)	2.4 (0.7)	2.4 (0.5)	2.0 (0.6)	1.8 (0.4)	3.7 (1.7)
N	45	30	45	30	45	30	60	15

An individual is assumed to have a "high" level of experience if she has either i) a greater number of transactions per month *or* ii) a greater number of years of market experience than the average subject in her treatment type (i.e., more than 6.6 (7.8) transactions per month or 9.4 (9.1) years of market experience in Treatment B (S)).
.

TABLE VI

REGRESSION ESTIMATES—EXPERIMENT I DEALER DATA

	Treatment B		Treatment S	
	Initial Offer	Final Offer	Initial Offer	Final Offer
Years of market experience	-0.47** (0.12)	-0.44** (0.11)	0.45** (0.12)	0.42** (0.12)
Trading intensity	-0.21 (0.17)	-0.12 (0.15)	0.17 (0.13)	0.16 (0.12)
Age	0.06 (0.16)	0.21 (0.15)	0.03 (0.13)	-0.09 (0.12)
Constant	-3.60 (6.6)	-6.17 (5.9)	-0.45 (5.3)	2.48 (5.2)
R^2	0.30	0.34	0.30	0.24
Adjusted R^2	0.24	0.28	0.27	0.20
n	41	41	61	61

a. Dependent variable is the average level of dealer discrimination as measured by dealer i's average offer to baseline members minus dealer i's average offer to minority group members.
b. Standard errors are in parentheses beneath coefficient estimates.
c. ** denotes coefficient estimate is significant at the $p < .05$ level.

TABLE VII

PREDICTIONS OF THE THREE THEORIES IN THE COMPLEMENTARY EXPERIMENTS

Dictator Game Experiment	*Chamberlain Experiment*	*Reservation Value-Elicitation Experiments*
Animus: Dealers are less generous towards minorities	**Animus**: in both treatments: (i) minority buyers do worse than baseline members (ii) experienced dealers do worse than inexperienced dealers	**Animus**: No prediction
Bargaining: No difference in generosity	**Bargaining:** in both treatments: (i) minority buyers do worse than baseline members (ii) experienced dealers outperform inexperienced dealers	**Bargaining:** No prediction
Statistical discrimination: No difference in generosity	**Statistical discrimination:** Case 1. dealers *knowingly* perform statistical discrimination: (i) minority buyers perform as well as baseline group members in the random treatment, but worse than baseline group members in the unclear treatment (ii) experienced dealers in the random treatment perform better than experienced dealers in the unclear treatment (because they mistakenly apply statistical discrimination in the latter) (iii) experienced dealers perform worse than inexperienced dealers in the unclear treatment, but no such difference emerges in the random treatment Case 2, dealers *unknowingly* perform statistical discrimination: (i) minority buyers perform worse than baseline group members in both treatments (ii) experienced dealers perform worse than inexperienced ones in both treatments	**Statistical discrimination:** (i) minorities have a different WTP/WTA distribution (ii) dealers are aware of this (iii) experienced dealers are more aware of the differences than inexperienced dealers.

TABLE VIII

DICTATOR EXPERIMENT SUMMARY RESULTS

Treatment	Average Offer Mean (Std. Dev.)	Rate of Positive Offers	Average Positive Offer Mean (Std. Dev.)
WM (partners are white males age 20-30) n = 60	1.07 (1.22)	0.50	2.13 (0.83)
WF (partners are white females age 20-30) n = 60	1.90 (1.60)	0.70	2.71 (1.19)
NWM (partners are nonwhite males age 20-30) n = 60	1.07 (1.39)	0.47	2.28 (1.14)
WMM (partners are white males age 60+) n = 60	1.20 (1.27)	0.53	2.25 (0.77)

Column 3 is reported in percentage of total amount available in the allocation decision (average positive offer ignores the zero-offers).

TABLE IX

CHAMBERLAIN EXPERIMENTAL DATA SUMMARY

	Treatment Random			Treatment Unclear		
	Pooled Lost Surplus (s.e.)	Inexp. Lost Surplus (s.e.)	Exp. Lost Surplus (s.e.)	Pooled Lost Surplus (s.e.)	Inexp. Lost Surplus (s.e.)	Exp. Lost Surplus (s.e.)
Sellers	0.24 (0.05) n=72	0.33 (0.08) n=34	0.17 (0.07) n=38	0.35 (0.07) n=72	0.38 (0.08) n=37	0.32 (0.09) n=35
White males, age 20-30	0.21 (0.11) n=24	0.29 (0.12) n=17	0.01 (0.24) n=7	0.07 (0.02) n=25	0.23 (0.12) n=15	-0.17 (0.08) n=10
White females, age 20-30	0.31 (0.11) n=18	0.39 (0.14) n=14	0.03 (0.26) n=4	0.45 (0.10) n=18	0.70 (0.11) n=10	0.13 (0.10) n=8
Non-white males, age 20-30	0.10 (0.08) n=15	0.21 (0.11) n=8	-0.03 (0.09) n=7	0.67 (0.07) n=14	0.74 (0.09) n=10	0.51 (0.12) n=4
White males, age 60+	0.23 (0.15) n=15	0.22 (0.20) n=5	0.23 (0.21) n=10	0.42 (0.10) n=15	0.43 (0.20) n=6	0.41 (0.11) n=9

a. Summary statistics in each cell are the average lost surplus, its standard error, and the number of observations. To derive the average lost surplus (lost surplus = predicted surplus – actual surplus), I first computed the average for each market participant over the five rounds within each session. I then computed the average and standard error of the individual averages. Predicted surplus is what would have happened had competitive predictions prevailed throughout the market session—I use a price of $13.50 when computing predicted surplus.

b. "Inexp." and "Exp." denote inexperienced and experienced sellers (dealers) and buyers (nondealers), respectively. A dealer (nondealer) is labeled "experienced" if he or she has 15 (9) or more years of market experience.

TABLE X

SUMMARY OUTCOMES—RESERVATION PRICE AND DEALER PERCEPTION EXPERIMENTS

	White Males Age 20-30 Mean (Std. deviation)		White Females Age 20-30 Mean (Std. deviation)		Nonwhite Males Age 20-30 Mean (Std. deviation)		White Males Age 60+ Mean (Std. deviation)	
Treatment	WTP	WTA	WTP	WTA	WTP	WTA	WTP	WTA

Panel A. Reservation Price Summary

Bid	69.59	---	81.61	---	67.10	---	77.40	---
	(26.4)		(36.8)		(38.5)		(45.7)	
Offer	---	130.05	---	127.46	---	145.06	---	140.76
		(96.8)		(110.0)		(118.3)		(132.6)
N	30	30	30	30	30	30	30	30

Panel B. Dealer Perception Summary   Percentage Correct

Comparison	Overall	Inexperienced Dealer	Experienced Dealer
WTP White males and:			
White females	15/20 (75%)	8/12 (67%)	7/8 (88%)
Nonwhite males	11/20 (55%)	5/11 (45%)	6/9 (67%)
Older males	13/20 (65%))	5/9 (55%)	8/11 (72%)
Aggregate:	39/60 (65%)	18/32 (56%)	21/28 (75%)
WTA White males and:			
White females	17/20 (85%)	7/10 (70%)	10/10 (100%)
Nonwhite males	15/20 (75%)	8/12 (67%)	7/8 (88%)
Older males	10/20 (50%)	4/10 (40%)	6/10 (60%)
Aggregate:	42/60 (70%)	19/32 (59%)	23/28 (82%)

a. Panel A summarizes data from the auctions for the Griffey Jr. sportscard. Panel B summarizes data gathered from the dealer perception experiment in which dealers chose which distribution was from the majority group and which was from the minority group.

b. Figures can be read as follows: in the WTP white males and white females comparison, 15 of 20 dealers correctly matched the reservation value distributions. More specifically, 8 of 12 (7 of 8) inexperienced (experienced) dealers correctly matched the reservation value distributions. A dealer is labeled "experienced" if he has 15 or more years of market experience.
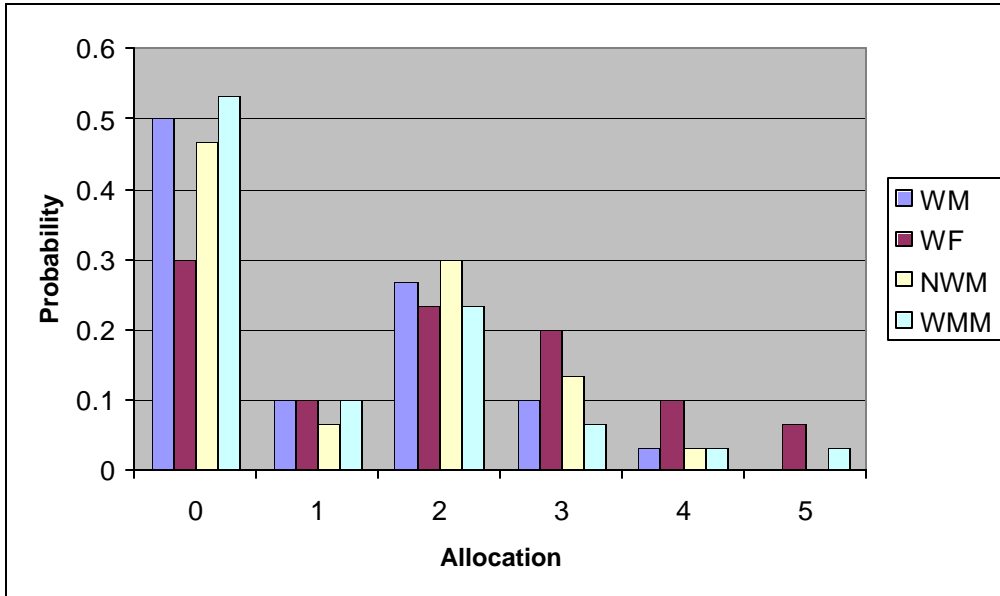
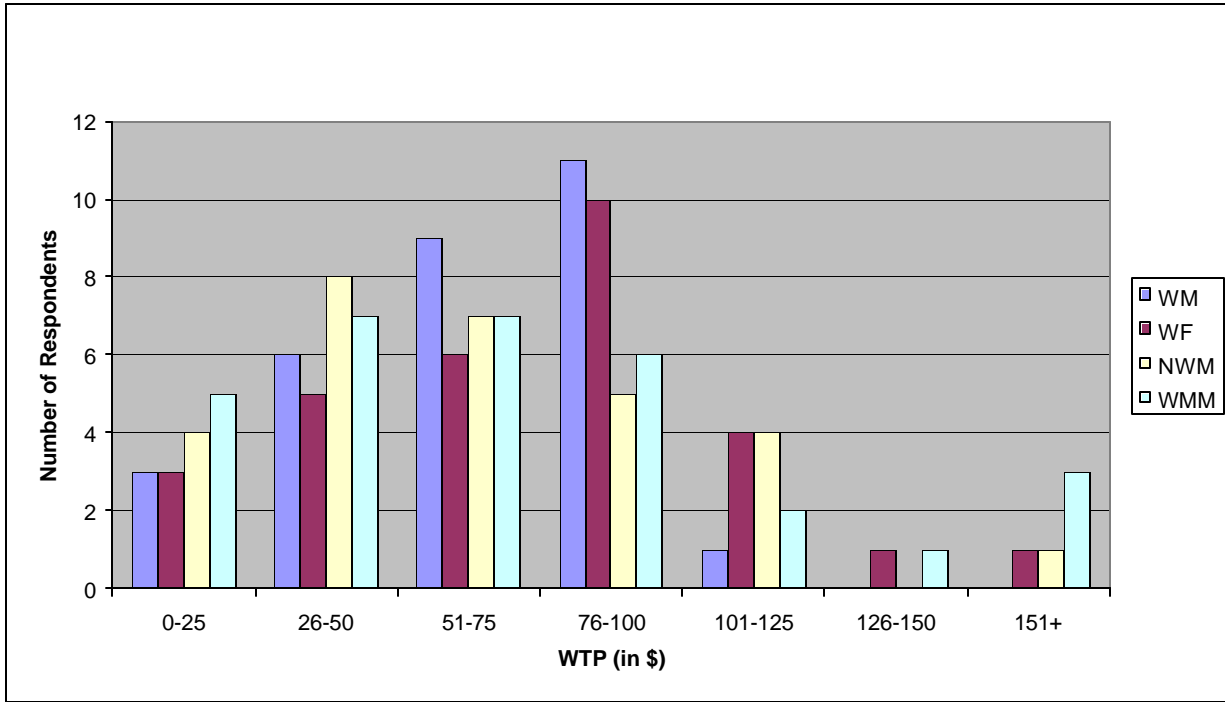**Figure 1. Dictator Experiment Results Summary**
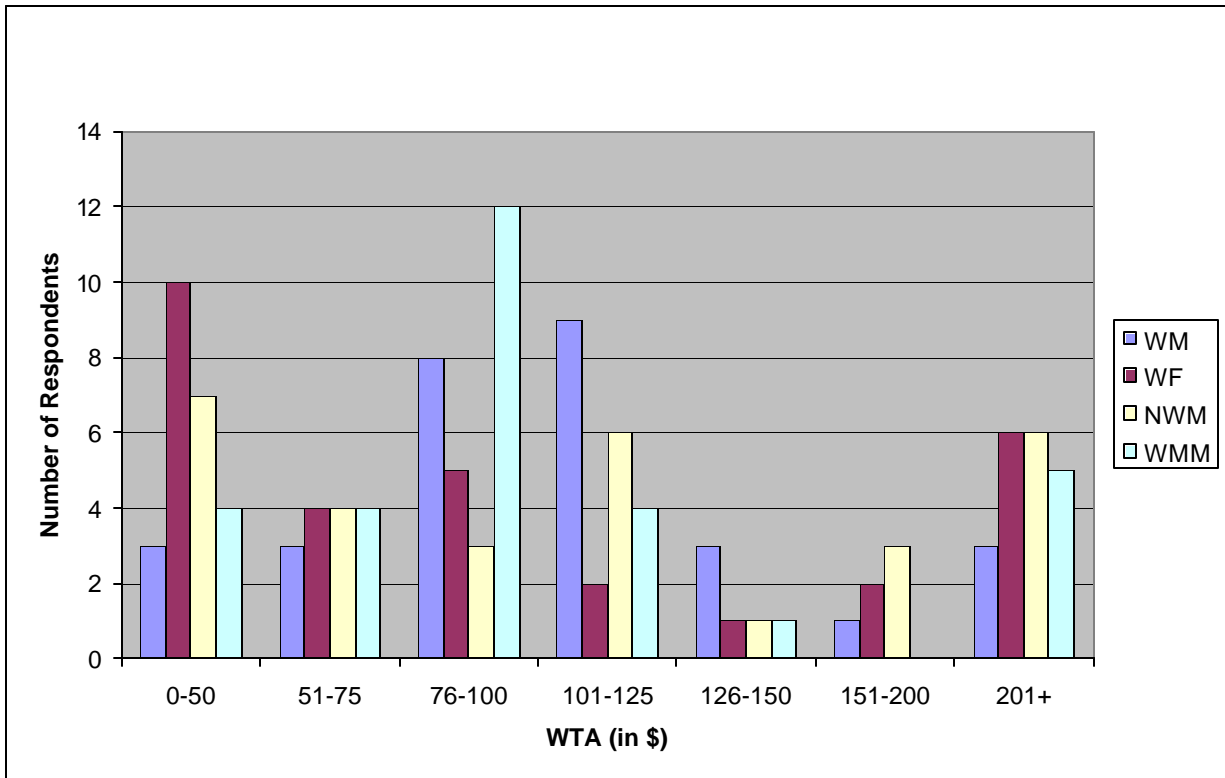
Figure 2. Reservation Value Distributions: WTP

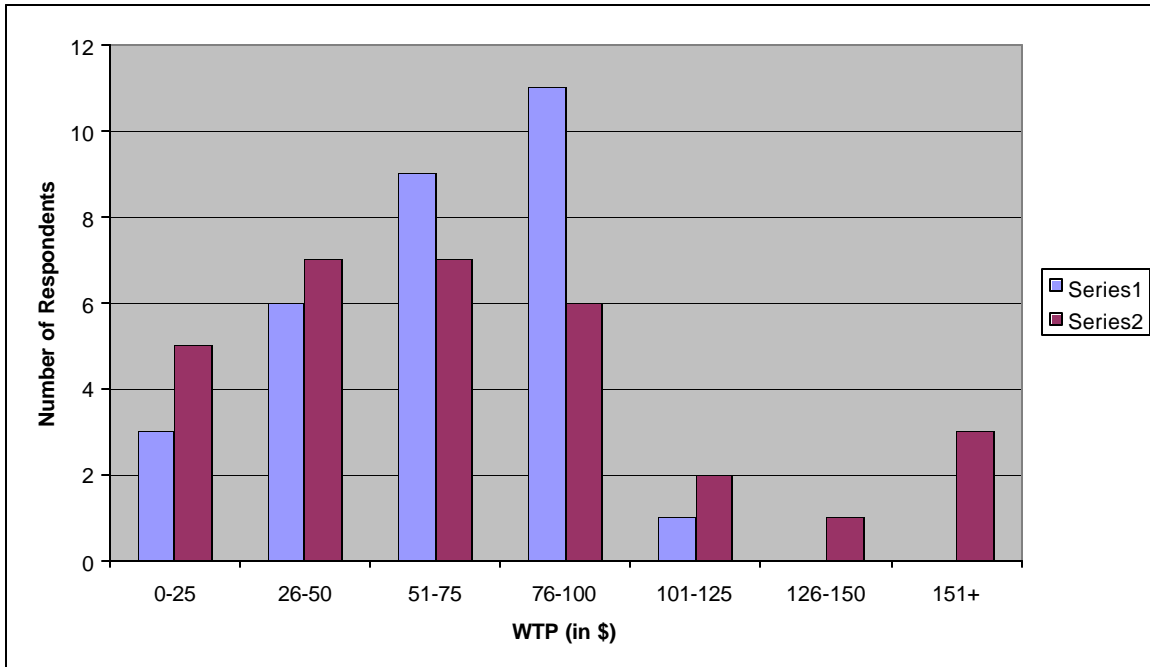**Figure 3. Reservation Value Distributions: WTA**
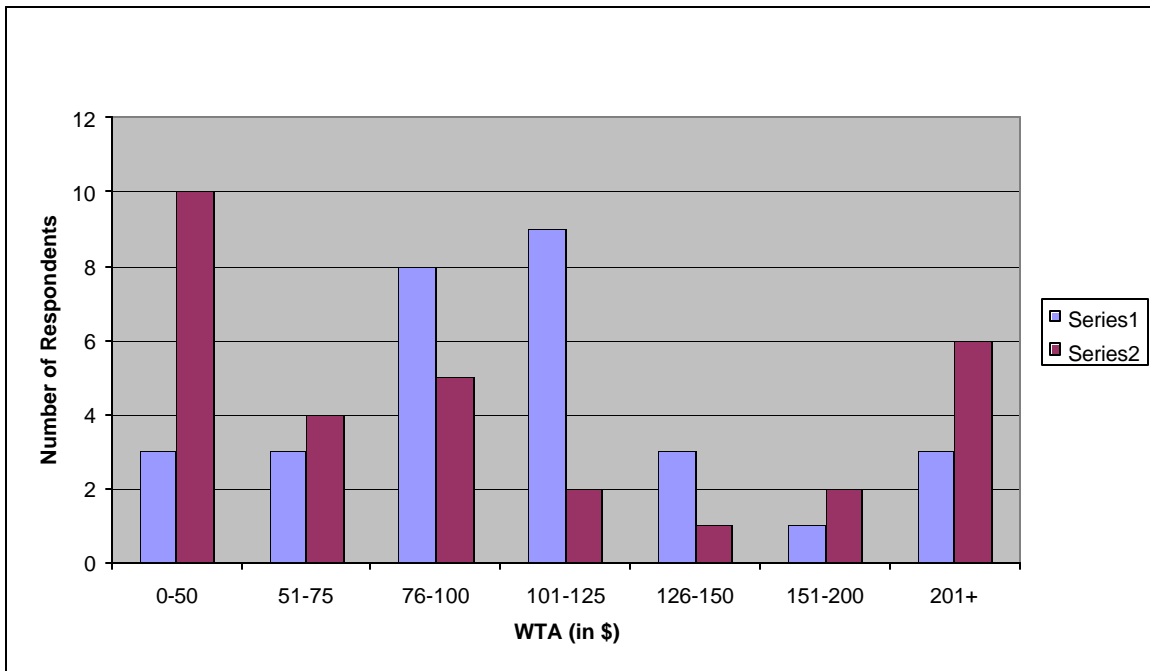
Figure 4. Dealer Perception Experimental Example: WTP

Figure 5. Dealer Perception Experimental Example: WTA