

## American Economic Association

---

Race and Gender Discrimination in Bargaining for a New Car

Author(s): Ian Ayres and Peter Siegelman

Source: *The American Economic Review*, Vol. 85, No. 3 (Jun., 1995), pp. 304-321

Published by: American Economic Association

Stable URL: <http://www.jstor.org/stable/2118176>

Accessed: 10/02/2009 09:31

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=aea>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *The American Economic Review*.

<http://www.jstor.org>

# Race and Gender Discrimination in Bargaining for a New Car

By IAN AYRES AND PETER SIEGELMAN\*

*More than 300 paired audits at new-car dealerships reveal that dealers quoted significantly lower prices to white males than to black or female test buyers using identical, scripted bargaining strategies. Ancillary evidence suggests that the dealerships' disparate treatment of women and blacks may be caused by dealers' statistical inferences about consumers' reservation prices, but the data do not strongly support any single theory of discrimination. (JEL J70, J15, J16)*

The purchase of a new car typically involves negotiations between buyer and seller. Such negotiations may leave room for sellers to treat buyers differently on the basis of race or gender, especially because any individual buyer has little or no means of learning the prices paid by others. The tests we report in this paper confirm this possibility; we find large and statistically significant differences in prices quoted to test buyers of different races and genders. This is true even though the testers were selected to resemble each other as closely as possible, were trained to bargain uniformly, and followed a prespecified bargaining script.

Race or gender discrimination by sellers might be motivated by two broad kinds of forces. The first is noneconomic tastes for discrimination (including traditional forms

of animus or bigotry) introduced into the market by a firm's owner, employees, or customers (Gary Becker, 1957). Even a market in which no participants are prejudiced might exhibit discrimination, however, if dealers use buyers' race or gender to make statistical inferences about the expected profitability of selling to them. Our study finds some evidence that is consistent with both broad theories of discrimination. Some discrimination may be attributable to seller animus. But our data also suggest that at least part of the observed disparate treatment of women and blacks is caused by dealers' inferences about consumer reservation prices.

Statistical inferences might disadvantage black or women consumers even though they are on average poorer than white males and should therefore have lower (opportunity) costs of search (George Stigler, 1968). Differences in information and (direct) search or negotiation costs might give white males lower reservation prices, despite their greater ability to pay and higher opportunity costs of search time. Moreover, profit-maximizing discrimination could well depend on more than a group's mean reservation price (Steven Salop and Joseph Stiglitz, 1977). It may be profitable for dealers to offer higher prices to a group of consumers who have a lower *average* reservation price, if the variance of reservation prices within the group is sufficiently large. Thus for example, suppose that a larger proportion of black (than white) consumers are willing to pay a high markup, even

\*Ayres: Yale Law School, P.O. Box 208215, New Haven, CT 06520 (e-mail: AYRES@MAIL.LAW.YALE.EDU); Siegelman: American Bar Foundation, 750 N. Lake Shore Drive, Chicago, IL, 60611 (e-mail: SIEGELMA@MERLE.ACNS.NWU.EDU). Kathie Heed, Akilah Kamaria, and Darrell Karolyi provided superb assistance with all phases of this project. Roz Caldwell prepared the manuscript with intelligence and good humor. We benefited from helpful comments by Jay Casper, Carolyn Craven, John Donohue, Richard Epstein, William Felstiner, Robert Gertner, James Heckman, and Carol Sanger, as well as substantial input from our colleagues at the American Bar Foundation. We especially want to acknowledge the invaluable advice of Peter Cramton and the sterling assistance provided by Michael Horvath.

though the mean (or median) black customer has a lower reservation price than her white counterpart. Knowing this, dealers might rationally offer higher prices to all black consumers.<sup>1</sup>

The rest of this paper proceeds in three sections. The first explains the audit method used to generate our data and discusses econometric specification. Section II then analyzes the empirical evidence for the existence of race and gender discrimination. Finally, Section III uses some ancillary data to explore the causes of the disparate treatment we found. There is some support for both statistical and animus-based theories of discrimination in the data.

## I. Method

### A. Design of the Study

This study used an audit technique in which pairs of testers (one of whom was always a white male) were trained to bargain uniformly and then were sent to negotiate for the purchase of a new automobile at randomly selected Chicago-area dealerships.<sup>2</sup> Thirty-eight testers bargained for 306

cars at 153 dealerships.<sup>3</sup> Both testers in a pair bargained for the same model of car, at the same dealership, usually within a few days of each other. Unlike most other audit studies, however, we allowed the composition of pairs to vary from audit to audit.<sup>4</sup> Dealerships were selected randomly; testers were randomly assigned to dealerships; and the choice of which tester in the pair would be the first to enter the dealership was also made randomly. The testers bargained at different dealerships for a total of nine car models,<sup>5</sup> following a uniform bargaining script that instructed them to focus quickly on one particular car and start negotiating over it. At the beginning of the bargaining, testers told dealers that they could provide their own financing for the car.

After deciding which car they were going to bargain over,<sup>6</sup> testers waited for an offer

---

Because this study involves deception, it necessarily raises important questions of research ethics (Ayres, 1991; Michael Fix and Raymond Struyk, 1992). We minimized the effects of our tests on sellers by conducting tests at off-peak hours (mid-mornings and mid-afternoons during the week) and by instructing testers to abandon the test if all salespeople were busy with legitimate customers.

<sup>1</sup>In other markets, competition often gives individual sellers an incentive to undermine price discrimination by offering posted prices with lower markups. The general failure of dealerships to opt for posted prices may be attributed to the high concentration of profits in a few car sales. Some dealerships may earn up to 50 percent of their profits from just 10 percent of their sales (Ayres, 1991 p. 854). Committing to posted prices would force dealerships to forgo these high-profit sales. If the extra profits from additional sales at a posted price are less than the forgone profits from selling a few cars at extremely high markups, individual dealers may not have a first-mover advantage in changing from bargained to posted prices. Nevertheless, recent evidence suggests that a move to posted prices for cars may be underway (Jim Mateja, 1992; Frank Swoboda, 1992).

<sup>2</sup>The technique is analogous to "fair housing" tests for discrimination in the real-estate market (John Yinger, 1986). Audit procedures were also used in tests of employment discrimination by Jerry Newman (1978), Shelby McIntyre et al. (1980), and in two recent Urban Institute studies (Harry Cross et al., 1990; Margery Turner et al. (1991). For an analysis of the strengths and weaknesses of this technique, see James Heckman and Siegelman (1992).

<sup>3</sup>We began with 404 tests, but because of discarded tests and scheduling difficulties, ended up deleting one of the observations for 98 audits, leaving us with 306 tests. While the techniques are somewhat more complicated, it is possible to analyze both the paired and unpaired observations together, using a variant of the approaches described here. We conducted extensive tests (see Ayres and Siegelman, 1992) to examine whether our results are in any way sensitive to the exclusion of the 98 "unpaired" observations. We concluded that they are not, and therefore we report only the results from the paired data set in the following analysis.

<sup>4</sup>In other words, rather than matching tester A with tester B for all tests, A was sometimes matched with B, sometimes with C, and so on.

<sup>5</sup>Testers in a pair bargained for the same car model, but the test allowed dealers to systematically steer testers to cars with different options. There is no evidence of this behavior: the average cost of the cars bargained for did not vary significantly by tester type. The nine models included a range from compacts to standard-size cars and included both imports and domestic makes. Human-subjects constraints prevent us from disclosing the identities of the car models.

<sup>6</sup>If they were shown more than one car of the type they were bargaining for, the testers were instructed to choose the car with the lowest sticker price.

from the dealer, or after 5 minutes elicited a dealer offer. Once the dealer made an initial offer, the tester waited 5 minutes and responded with a counteroffer equal to our estimate of the dealer's marginal cost for the car.<sup>7</sup> If the salesperson responded by lowering his or her offer, the test continued, with the tester's second counteroffer derived from the script in one of two ways.

At some dealerships, testers used a "split-the-difference" strategy. In these tests, the tester responded to subsequent dealer offers by making counteroffers that averaged the dealer's and the tester's previous offers. Thus, if a tester's first counteroffer was \$10,000 and the salesperson responded with an offer of \$12,000, the tester's next response would be \$11,000. At other dealerships, the testers used a "fixed-concession" strategy in which their counteroffers (concessions) were independent of sellers' behavior. Testers began, as before, by making their first counteroffer at marginal cost. Regardless of how much the seller conceded, each subsequent counteroffer by the tester increased by 20 percent of the difference between the sticker price and the tester's previous offer.<sup>8</sup>

Under either bargaining strategy, the test ended when the dealer either (i) attempted to accept a tester's offer,<sup>9</sup> or (ii) refused to bargain further. During the course of negotiations, testers jotted down each offer and counteroffer, as well as options on the car and its sticker price. After leaving the dealership, each tester completed a survey describing ancillary details of the test (including the kinds of questions they were asked,

the race and gender of the salesperson with whom they negotiated, etc.).

### B. Controls and Uniformity

The paired audit technique is designed to eliminate as much intertester variation as possible, and thus to insure that differences in outcomes (such as prices quoted) reflect differences in *dealer* rather than tester behavior.<sup>10</sup> We began by choosing testers according to the following criteria:

- (i) *Age*: All testers were between 28 and 32 years old.
- (ii) *Education*: All testers had 3–4 years of postsecondary education.
- (iii) *Attractiveness*: All testers were subjectively chosen to have average attractiveness.

The testers also displayed similar indicia of economic class. Besides volunteering that they did not need financing, all testers wore similar "yuppie" sportswear and drove to the dealership in similar rented cars.

The script governed both the verbal and nonverbal behavior of the testers, who volunteered very little information and were trained to feel comfortable with extended periods of silence. The testers had a long list of contingent responses to the questions they were likely to encounter. If asked, they gave uniform answers about their profession (e.g., a systems analyst at a large bank) and address (a prosperous Chicago neighborhood).

<sup>10</sup>Heckman and Siegelman (1992 p. 188) point out that:

Despite suggestive rhetoric to the contrary, audit pair studies are not experiments or matched pair studies. Race or ethnicity cannot be assigned by randomization or some other device as in... [a classical experiment]. Race is a personal characteristic and adjustments must be made instead on "relevant" observed characteristics to "align" audit pair members.

<sup>7</sup>Estimates of dealer cost were provided by *Consumer Reports Auto Price Service* and *Edmund's 1989 New Car Prices*. As we discuss below, making an initial offer at the dealer's cost reveals some sophistication on the buyer's part.

<sup>8</sup>That is, if the car had a sticker price of SP and the tester's last offer was LO, then the tester's next offer would be  $LO + 0.2 \times (SP - LO)$ . Since the gross margin (SP - LO) decreases as the bargaining continues, the fixed-concession strategy produced smaller concessions in each subsequent round.

<sup>9</sup>The testers did not purchase cars. If a salesperson attempted to accept a tester offer, the tester would end the test, saying, "Thanks, but I need to think about this before I make up my mind."

Because selling a car is a more discrete transaction than hiring an employee or renting out an apartment, the task of matching testers is substantially easier in this area.

Before visiting the dealerships, the testers had two days of training in which they memorized the bargaining script and participated in numerous mock negotiations that helped them negotiate and answer questions uniformly. Unlike many other audit studies, the testers did not know that another tester would visit each dealership, or even that the study tested for discrimination.<sup>11</sup>

Despite our efforts to insure uniformity, some differences between testers undoubtedly remained. Two important questions about such residual differences must then be asked: First, are they likely to be correlated with race or gender? If not, the remaining nonuniformity should not influence our conclusion that it is race and gender that generate different outcomes for the testers. Second, are the residual differences large enough to explain the *amount* of discrimination we report below? Although no experiment can eliminate all idiosyncratic differences in tester behavior, we feel confident that the amounts of discrimination we observed cannot plausibly be explained by divergence from the uniform bargaining behavior called for in our script.

### C. *Econometric Specification*

In the analysis that follows, we consider four definitions of the dependent variable. The profit that the dealership would earn on its initial offer provides an especially well-controlled test for discrimination.<sup>12</sup> Be-

cause the initial offer was made by the dealer with relatively little intervention on the tester's part, it is unlikely that differences in first offers reflect differences in testers' abilities to follow our uniform bargaining script. On the other hand, the profit that the dealership would earn on its final offer more closely reflects the price a real consumer would pay. We use both percentage markup over marginal cost and actual dollar profits as dependent variables.

Table 1 presents some simple summary statistics that reveal the overall pattern of discrimination in dealer offers. White male testers were quoted initial offers that were roughly \$1,000 over dealer cost. Offers to black males averaged about \$935 higher than those to white males. Black female testers got initial offers about \$320 higher than those white males received, while white females received initial offers that were \$110 higher. These differences are statistically significant at the 0.05 level, except for white females.

Not surprisingly, the process of negotiation lowered dealers' average offers to all four tester types. However, dealer concessions further increased the disparities between white males and black testers, while only slightly narrowing the gap for white females. Thus, there is a stronger overall pattern of discrimination in final offers than in initial offers: black males were asked to pay \$1,100 more than white males, black females \$410 more, and white females \$92 more. Although the differences in concessions by tester type were not statistically significant, it is striking that black male testers, despite receiving the highest initial offers, got the lowest average concessions (\$290, or 15 percent) over the course of negotiations.

The results in Table 1 are suggestive but do not make full use of the information available from the audits. One improvement

<sup>11</sup>The testers were told only that we were studying how sellers negotiate car sales. For the importance of isolating participants from "experimenter effects" (behavior induced by an unconscious desire to produce the expected results), see Robert Rosenthal (1976).

<sup>12</sup>Profits on the initial offer were calculated as the difference between the dealer's first offer (before any bargaining took place) and our estimate of marginal cost for the car. Marginal cost was in turn derived as follows. We began with an estimate of the dealer's cost for the base model with no options, using data from *Consumer Reports* and *Edmund's*. We then subtracted the sticker price for the base car from the total sticker price (including options), giving the retail cost of the options. We applied an option-specific discount factor (dealer markup on each option, also derived from

*Consumer Reports* and *Edmund's*) to each option price, to get the marginal cost of all the options. The marginal cost of the options was then added to the marginal cost of the base model to give the marginal cost of the car.

TABLE 1—SUMMARY STATISTICS ON PROFITS AND COSTS, BY TESTER TYPE

| Tester type                                       | Initial profit | Final profit | Concession <sup>a</sup> |
|---|----------------|--------------|-------------------------|
| White males (18 testers; 153 observations)        |                |              |                         |
| Mean  | 1,018.7        | 564.1        | 454.6                   |
| Standard deviation                                | 911.3          | 708.0        | (44.6 percent)          |
| Average markup (percentage)                       | 9.20           | 5.18         |                         |
| White females (7 testers; 53 observations)        |                |              |                         |
| Mean  | 1,127.3        | 656.5        | 470.8                   |
| Difference from white male average                | 108.6          | 92.4         | (41.8 percent)          |
| Standard deviation                                | 785.3          | 472.4        |                         |
| Average markup (percentage)                       | 10.32          | 6.04         |                         |
| Black females (8 testers; 60 observations)        |                |              |                         |
| Mean  | 1,336.7*       | 974.9*       | 361.8                   |
| Difference from white male average                | 318.0          | 246.1        | (27.1 percent)          |
| Standard deviation                                | 887.8          | 827.8        |                         |
| Average markup (percentage)                       | 12.23          | 7.20         |                         |
| Black males (5 testers; 40 observations)          |                |              |                         |
| Mean  | 1,953.7*       | 1,664.8*     | 288.9                   |
| Difference from white male average                | 935.0          | 1,100.7      | (14.8 percent)          |
| Standard deviation                                | 1,122.7        | 1,099.5      |                         |
| Average markup (percentage)                       | 17.32          | 14.61        |                         |
| All nonwhite males (20 testers, 153 observations) |                |              |                         |
| Mean  | 1,425.5*       | 1,045.0*     | 380.5                   |
| Difference from white male average                | 406.8          | 481.0        | (26.6 percent)          |
| Standard deviation                                | 973.6          | 989.9        |                         |
| Average markup (percentage)                       | 12.99          | 9.40         |                         |

<sup>a</sup>Average initial profit minus average final profit; average percentage concession is given in parentheses.

\*Significantly different from the corresponding figure for white males at the 5-percent level.

would be simply to regress profits on a vector of variables thought to explain them, including dummy variables for tester race and gender. This ordinary least-squares (OLS) regression will produce unbiased estimates of the race and gender effects, as long as any variables that might be omitted from this equation are uncorrelated with the race or gender of the testers.

These estimates will be inefficient, however, because OLS fails to account for the correlation between errors for the two observations in a given audit (John Yinger, 1986). This correlation arises because there are unobservable variables whose effects are common to both testers in the same audit, including, for example, any factors that are unique to the specific dealership being tested. Since these variables are omitted

from the OLS regression, their effect will be captured in the error term, imparting a correlation between errors at the same dealership.

We therefore exploit the panel structure of the data set, using the fact that we have two observations (one for a white male and one for one of the three other tester types) for each of the 153 audits. To capture the possibility of audit-specific errors we estimate the following fixed-effects model:

$$(1) \quad \Pi_{ai} = \mathbf{X}_{ai}\boldsymbol{\beta} + \mu_a + \varepsilon_{ai}$$

where  $\Pi_{ai}$  is dealer profit on the  $i$ th test ( $i = 1, 2$ ) in the  $a$ th audit ( $a = 1, \dots, 153$ ),  $\mathbf{X}_{ai}$  is a matrix of dummy variables for tester race/gender, a constant,  $\mu_a$ , is an unob-

TABLE 2—OLS AND FIXED-EFFECTS (ONE DUMMY PER AUDIT) REGRESSIONS OF INITIAL AND FINAL PROFITS AND MARKUPS ON RACE AND GENDER DUMMIES AND CONTROL VARIABLES

| Variable                        | Initial dollar profit |               | Final dollar profit |               | Initial percentage markup |               | Final percentage markup |               |
|---------------------------------|-----------------------|---------------|---------------------|---------------|---------------------------|---------------|-------------------------|---------------|
|                                 | OLS                   | Fixed effects | OLS                 | Fixed effects | OLS                       | Fixed effects | OLS                     | Fixed effects |
| Race/gender dummies:            |                       |               |                     |               |                           |               |                         |               |
| Constant                        | 1,014.95*             |               | 607.51*             |               | 0.114*                    |               | 0.072*                  |               |
|                                 | (4.12)                |               | (2.98)              |               | (5.44)                    |               | (4.19)                  |               |
| White female                    | 192.38                | 55.10         | 174.68*             | 129.09        | 0.017                     | 0.007         | 0.014                   | 0.013         |
|                                 | (1.23)                | (0.39)        | (1.35)              | (1.05)        | (1.26)                    | (0.63)        | (1.29)                  | (1.25)        |
| Black female                    | 404.28*               | 281.05*       | 504.64*             | 404.65*       | 0.039*                    | 0.027*        | 0.045*                  | 0.037*        |
|                                 | (2.75)                | (2.13)        | (4.15)              | (3.49)        | (3.09)                    | (2.42)        | (4.45)                  | (3.87)        |
| Black male                      | 1,068.24*             | 1,061.17*     | 1,242.85*           | 1,061.27*     | 0.094*                    | 0.091*        | 0.107*                  | 0.090*        |
|                                 | (6.10)                | (6.56)        | (8.57)              | (7.47)        | (6.31)                    | (6.60)        | (8.78)                  | (7.70)        |
| Controls                        |                       |               |                     |               |                           |               |                         |               |
| SPLIT <sup>a</sup>              | 20.30                 |               | -57.36              |               | -0.02                     |               | -0.02                   |               |
|                                 | (0.15)                |               | (-0.51)             |               | (-1.52)                   |               | (-1.95)                 |               |
| Time <sup>b</sup>               | -1.73                 |               | -2.47               |               | -0.0004*                  |               | -0.0004*                |               |
|                                 | (-0.88)               |               | (-1.52)             |               | (-2.29)                   |               | (2.70)                  |               |
| Experience <sup>c</sup>         | -3.58                 |               | -0.50               |               | 0.00                      |               | 0.00                    |               |
|                                 | (-0.43)               |               | (0.07)              |               | (0.10)                    |               | (0.56)                  |               |
| First <sup>d</sup>              | 203.32                |               | 192.18              |               | 0.01                      |               | 0.01                    |               |
|                                 | (1.69)                |               | (1.93)              |               | (1.30)                    |               | (1.48)                  |               |
| $F_{[3, 298]}$ :                | 12.91*                |               | 26.52*              |               | 14.04*                    |               | 27.98*                  |               |
| Adjusted $R^2$ :                | 0.10                  | 0.44          | 0.19                | 0.43          | 0.11                      | 0.45          | 0.21                    | 0.47          |
| Standard error of the estimate: | 914.35                | 723.2         | 757.1               | 635.6         | 0.078                     | 0.06          | 0.064                   | 0.05          |
| Degrees of freedom:             | 298                   | 150           | 298                 | 150           | 298                       | 150           | 298                     | 150           |
| N:                              | 306                   | 306           | 306                 | 306           | 306                       | 306           | 306                     | 306           |

Note: The numbers in parentheses are  $t$  statistics.

<sup>a</sup>Dummy variable: 1 if tester used a split-the-difference bargaining strategy; 0 otherwise.

<sup>b</sup>Number of days between this test and the first day of testing.

<sup>c</sup>Number of prior tests by this tester.

<sup>d</sup>Dummy variable: 1 if tester was first in the pair; 0 otherwise.

\* Statistically significant at the 5-percent level.

served, mean-zero, audit-specific error term,<sup>13</sup> and  $\varepsilon_{ai}$  is an independent, mean-zero error term.

Including an audit-specific fixed effect transforms each observation into a difference from its audit-specific mean. Thus, the fixed-effects regression (including only the race and gender dummies) is equivalent to a paired-difference estimate (Yinger, 1986).

<sup>13</sup>By definition, the factors that determine  $\mu_a$  are shared by both members of an audit. Thus,  $\mu_a$  must be uncorrelated with the race/gender dummies for audit  $a$ .

## II. Results

### A. Tester Race and Gender Effects

Table 2 reports the results of OLS and fixed-effects (one dummy per audit) regressions explaining raw profits and percentage markups associated with dealers' initial and final offers. Consistent with Table 1, the OLS regressions again suggest that a tester's gender and race strongly influence both the initial and final offers made by sellers.  $F$  tests for the joint significance of the three race/gender dummies (vs. a model with only

control variables and a constant term) are significant in all four of the regressions. However, the size of the race and gender effects is generally somewhat smaller in the OLS estimates than is suggested by the raw comparison of means. As with the raw means, white females are quoted the smallest additional markups over white males, and black males the largest. The white female effect is not significant.

Allowing for audit-specific fixed effects does not change the basic story. There is strong evidence for the presence of heterogeneity among audits (the 153 audit dummies are jointly significant in all four specifications). But controlling for such effects does not have a dramatic influence on either the size or significance of the tester-type dummies. In the fixed-effects regressions, black males receive initial offers that generate dealer profits \$1,100 (9 percentage points, or 81 percent) higher than those received by white males, with the disparity unchanged for final offers. While discrimination against black males does not increase in the final offers, this group still receives the smallest average concession (in both absolute and percentage terms). For black females, a gap of \$280 in initial offers widens to just over \$400 (3 percentage points higher markup) in final offers. Initial offers to white females are \$55 higher than to white males, with final offers differing by \$130. This amounts to about 1.7 percentage points of additional markup beyond the 11 percent quoted to white male testers. The estimated coefficients are significant for the black testers, although not for white females.

### B. Control Variables

Our confidence in the methodology is supported by the finding that the variables testing whether the study was adequately controlled produced coefficients that were neither large nor statistically significant. We were concerned about possible secular trends in the car market because the tests were carried out over a period of  $4\frac{1}{2}$  months. The regressions do indicate that there was a slight downward trend in car prices over the period covered by our tests; but given our

testing procedures, this trend should not be correlated with race or gender and is therefore innocuous.<sup>14</sup> Another concern was that a tester's experience—the number of previous tests he or she had conducted—might influence the bargaining outcomes. The tables provide no evidence of any such experience effect.

We also examined whether the dealership's experience with the first tester affected its treatment of the second tester in the pair, as could happen, for example, if the seller learned that a test was taking place. (The two testers in a pair rarely negotiated with the same salesperson; and dealers never gave any indication that they suspected our testers were not bona fide buyers. Both of these facts suggest that the probability of discovery should have been low.) The order effect, captured by the FIRST dummy, was never statistically significant in any of the regressions in Table 2. Its magnitude, however, was surprisingly large, with the first tester asked to pay a \$200, or 1 percentage point, higher markup than the second.<sup>15</sup>

The regressions in Table 2 also control for a bargaining-strategy effect. Buyers did slightly better with the split-the-difference strategy than with fixed concessions. However, this effect was quantitatively small and statistically insignificant.

<sup>14</sup>Since many car salespeople are paid on a commission basis, and since there are weekly and monthly quotas, we wanted to allow for possible day-of-week and week-of-month effects. In alternative specifications (not shown), we tested for these effects. They were uniformly small and insignificant, with the exception that dealerships' profits tended to be lower on Fridays. A referee suggested that this might be explained by dealers' inferences about consumers' propensity to engage in additional search. A consumer shopping on Friday may be more likely to visit other dealerships during the weekend, and dealers may therefore offer lower prices at the beginning of the weekend to forestall this additional search.

<sup>15</sup>One possible explanation is that sellers quoted lower prices to subsequent buyers because the failure to complete a sale to the first tester caused them to believe that demand conditions were worse than they had expected. While theoretically plausible, it seems unlikely that the "learning" effect from a single failed sale could explain so substantial a price decrease.



TABLE 3—OLS FIXED-EFFECTS (ONE DUMMY PER AUDIT) AND GLS (RANDOM-EFFECTS) REGRESSIONS OF FINAL PROFITS AND MARKUPS ON RACE AND GENDER DUMMIES, AUDIT-SPECIFIC EFFECTS

| Variable  | Final dollar profit |                    |                     | Final percentage markup<br>(actual coefficients $\times 100$ ) |                 |                 |
|---|---------------------|--------------------|---------------------|--|-----------------|-----------------|
|   | OLS                 | Fixed effects      | GLS                 | OLS  | Fixed effects   | GLS             |
| Constant  | 564.09*<br>(9.16)   | —                  | 564.09*<br>(9.11)   | 5.18*<br>(9.94)  | —               | 5.18*<br>(9.89) |
| White female  | 92.40<br>(0.76)     | 129.10<br>(1.04)   | 103.83<br>(0.96)    | 0.86<br>(0.84)   | 1.26<br>(1.25)  | 1.00<br>(1.12)  |
| Black male  | 1,100.68*<br>(8.13) | 1061.27*<br>(7.47) | 1,088.40*<br>(8.87) | 9.43*<br>(8.24)  | 8.96*<br>(7.70) | 9.26*<br>(9.12) |
| Black female  | 410.79*<br>(3.54)   | 404.65*<br>(3.49)  | 408.88*<br>(3.96)   | 3.71*<br>(3.78)  | 3.68*<br>(3.87) | 3.70*<br>(4.35) |
| Adjusted $R^2$ :  | 0.18                | 0.43               |                     | 0.18   | 0.47            |                 |
| $N$ :   | 306                 | 306                | 306                 | 306  | 306             | 306             |
| Likelihood-ratio test<br>(fixed effects vs. OLS):                             |                     | 325.12*            |                     |  | 345.30*         |                 |
| Breusch-Pagan test <sup>a</sup><br>(random effects vs. OLS, $\chi^2_{[1]}$ ): |                     |                    | 14.52*              |  |                 | 18.90*          |

Note: The numbers in parentheses are  $t$  statistics.

<sup>a</sup>Reject OLS in favor of random effects for large values of the test statistic.

\*Significantly different from zero at the 5-percent level.

### C. Robustness

The differences in prices quoted to the various tester types found in Table 2 are robust to a variety of alternative specifications and nonparametric tests.

1. *Fixed versus Random Effects.*—It is possible to compare the fixed-effects specification (one dummy variable for each of the 153 audits) described earlier with a random-effects (generalized least-squares [GLS]) specification in which each audit's error term is treated as a random draw from a common distribution. Table 3 presents such comparisons for the final profit and final markup equations, focusing on the tester-type variables (which vary within an audit).

Like the fixed-effects estimates, the GLS estimates indicate heterogeneity across audits: a Breusch-Pagan test indicates that the estimated variance of the audit-specific error term is significantly greater than zero in the random-effects specification. However, controlling for this heterogeneity (with either the random- or fixed-effects specifica-

tions) did not affect the size or significance of the tester-type coefficients.

2. *Individual-Tester Effects.*—Because we have multiple observations for each of the 38 testers (and testers were not paired with a single, fixed partner), we can also test for the presence of individual-tester effects. To do this, we simply reorganize the panel data by individual testers (for example,  $\Pi_{it}$  = dealer profit on the  $i$ th test for the  $t$ th tester) and compute a standard random-effects regression.<sup>16</sup>

<sup>16</sup>Note that the training and selection of the testers were designed to eliminate as much intertester variation as possible. Thus, we would expect to find little or no evidence of individual-tester effects in our data. For reasons described above, however, we cannot test for the presence of individual-tester effects that are correlated with testers' race or gender.

A fixed-effects specification with one dummy variable for each individual tester is equivalent to subtracting off the tester-specific mean for each variable. This means that any variables that do not vary over time for each individual tester (including the tester race and gender dummies) are indistinguishable from the

Rerunning the four specifications of Table 2, we found little evidence that significant individual-tester effects were present. For three of the four regressions, we found that the estimated variance of the individual-tester effects was less than zero, indicating that the individual effects were not statistically significant. (Positive variance estimates are not guaranteed in a finite sample, even though the variance is estimated consistently [William Greene, 1991 p. 493]). A Breusch-Pagan test statistic of 16.71 in the final dollar-profit regression indicates that there was some heterogeneity across testers (i.e., the variance of the individual-specific error was significantly greater than zero). The estimated coefficients in this regression, however, were virtually identical to the OLS and GLS estimates presented in Table 2, and all of the black-tester coefficients remained significant.

3. *Attempted Acceptances versus Refusals to Bargain Further.*—We also investigated whether our findings of race and gender discrimination might be linked to the fact that the dealerships' final offers were sometimes refusals to bargain further and sometimes acceptances of tester offers.<sup>17</sup> By adding an ACCEPT dummy (= 1 if the seller attempted to accept a tester offer) to the regressions of Table 2, we found that sessions ending in attempted acceptances had a \$400 lower final profit than those that ended in a refusal to bargain ( $t$  statistic of 4.20). The size of this acceptance effect,

---

individual-tester fixed effect and cannot be used, thus making the individual-tester, fixed-effect model inappropriate for examining race and gender effects (Jerry Hausman and William Taylor, 1981).

<sup>17</sup>In a parallel effort, we examined whether our results were affected by the fact that sellers sometimes made unsolicited initial offers and sometimes needed to have offers elicited by the testers. Logit regressions indicated that dealers were less likely to make an unsolicited initial offer to white males than to other tester types, but that this difference was not statistically significant. Solicited initial offers were significantly larger than unsolicited initial offers, but there was no statistical difference in final offers between tests that began with elicited initial offers and those that began with unsolicited offers by the seller.

however, was the same for all testers: interacting the ACCEPT dummy with the tester type yielded small and insignificant coefficients.<sup>18</sup>

Dealers' willingness to offer lower prices to white males was reflected in a greater willingness to continue bargaining until an acceptable offer was made. When the tester was a white male, 25.6 percent of the tests ended in attempted seller acceptances; this figure was only 14.9 percent for the other tester types. The fact that sellers are more likely to accept offers from white males actually biases our estimates *against* finding discrimination, however, because acceptances only provide an upper bound for sellers' reservation prices. That is, in those cases where dealers attempted to accept an offer from a white male tester, the dealers might have been willing to make an even lower offer, which would have increased our measure of discrimination. Overall, our findings of discrimination do not seem to be sensitive to the fact that most negotiations did not end in an attempted acceptance.

4. *Nonparametric Tests for Race and Gender Effects.*—If race or gender were unrelated to the prices quoted to testers, we would expect that the benchmark white male testers would get lower offers than their audit partners half the time, while doing worse than their counterparts in the remaining half of the tests. As Table 4 indicates, however, this was not the case. Overall, white males did better than others in roughly two-thirds of the paired tests (for both initial and final offers). A likelihood-ratio test reveals that the differences from 50 percent were all statistically significant at the 5-percent level.

The disparities are even larger in dollar terms: in tests in which white male testers received the lower final offer, they did \$897 better than their counterparts on average. Where the nonwhite males did better, they

<sup>18</sup>The ACCEPT variable may not be exogenous in these regressions, because higher profitability may cause the dealer to accept a tester's offer.

TABLE 4—PROPORTION OF TESTS IN WHICH WHITE MALE OBTAINED THE BETTER RESULT

| Test  | Percentages     |               |
|---|-----------------|---------------|
|   | Initial profits | Final profits |
| White males vs. all others<br>(153 pairs)   | 68.0            | 66.7          |
| White males vs. white females<br>(53 pairs) | 58.4            | 56.6          |
| White males vs. black males<br>(40 pairs)   | 87.5            | 85.0          |
| White males vs. black females<br>(60 pairs) | 63.3            | 61.7          |

Notes: All values are significantly different from 50 percent at the 1-percent level using a likelihood-ratio test ( $\chi^2_{[1]}$ ); in 43.5 percent of the tests, white males received an *initial* offer that was lower than the *final* offer made to the nonwhite male tester.

beat the white male by only \$167. Perhaps even more startling, in 43.5 percent of the tests, white males received an initial offer that was lower than the *final* offer made to their audit-mate. That is, without any negotiating at all, 43 percent of white males obtained a better price than their counterparts achieved after an average of 45 minutes of bargaining. Wilcoxon signed-ranks tests (Morris DeGroot, 1986 pp. 573–76) similarly reveal that the median final and initial profits with white males were significantly lower than those with the other tester types. This suggests that white males did better on average not simply because a few of them received very low offers, but because the entire distribution of offers to white males was lower than for the other tester types.

Two conclusions emerge from this analysis. First, both final and initial offers display large and significant differences in outcomes by race and gender. For black males, the final markup was 8–9 percentage points higher (24 percent vs. 15 percent) than for white males; the equivalent figures are 3.5–4 percentage points for black females and about 2 percentage points for white females. Second, the results are robust. The magnitude and significance of the race and gender effects under various alternative

specifications, combined with the insignificance of the individual-tester effects, reinforce our confidence in these conclusions.

### III. The Sources of Discrimination

In this section we try to explain the race and gender discrimination uncovered in our testing. It is particularly difficult to distinguish between competing hypotheses without an explicit model of how bigotry or asymmetric information might influence sellers' bargaining behavior. Either animus or statistical inference might cause sellers to make higher take-it-or-leave-it offers to some groups. But when sellers can make alternating offers over time, as occurs during the purchase of a new car, the consequences of animus or asymmetric information become much murkier.<sup>19</sup>

Thus, our results should not be read as explicit tests of the two theories of discrimination. Instead, we simply explore the effects of some plausible covariates on the level of discrimination, as well as considering some ancillary evidence from other research. We conclude that the dealerships' disparate treatment of women and blacks may be caused by dealers' statistical inferences about consumers' reservation prices, but the results do not strongly support any single theory of discrimination.

#### A. Animus-Based Discrimination

Discrimination might be caused by the bigotry of a dealership's owners, employees, or customers. In this view, the higher prices paid by minorities and women serve to compensate the bigoted market participants for having to associate with the victims of discrimination (Gary Becker, 1957).

In Table 5, we report regressions (analogous to Table 2) testing whether the race

<sup>19</sup>For a bargaining-theoretic analysis of discrimination in the sale of new cars that shows how different animus-based and statistical theories disparately affect a seller's equilibrium negotiation strategy, see Narasimhan Srinivasan and Kuang-Wei Wen (1991) or Ayres (1994).

TABLE 5—OLS AND GLS (RANDOM EFFECTS, AUDIT-SPECIFIC ERRORS) REGRESSIONS OF INITIAL AND FINAL PROFITS/MARKUPS ON RACE/GENDER DUMMIES AND OTHER VARIABLES

| Variable   | Initial dollar profit |                | Final dollar profit |                | Initial percentage markup |                | Final percentage markup |                |
|--|-----------------------|----------------|---------------------|----------------|---------------------------|----------------|-------------------------|----------------|
|  | OLS                   | Random effects | OLS                 | Random effects | OLS                       | Random effects | OLS                     | Random effects |
| <b>Race/gender dummies:</b>                            |                       |                |                     |                |                           |                |                         |                |
| Constant   | 1,704.46*             | 1,713.50*      | 1,343.17*           | 1,338.14*      | 0.19*                     | 0.19*          | 0.148*                  | 0.147*         |
|  | (6.52)                | (5.88)         | (6.18)              | (5.71)         | (8.49)                    | (7.61)         | (8.40)                  | (7.78)         |
| White female   | 130.12                | 106.77         | 102.97              | 108.13         | 0.014                     | 0.012          | 0.009                   | 0.01           |
|  | (0.98)                | (0.90)         | (0.93)              | (1.05)         | (1.20)                    | (1.16)         | (1.11)                  | (1.24)         |
| Black male   | 985.19*               | 1,007.99*      | 1,156.55*           | 1,135.94*      | 0.084*                    | 0.086*         | 0.098*                  | 0.096*         |
|  | (6.59)                | (7.49)         | (9.32)              | (9.75)         | (6.61)                    | (7.60)         | (9.74)                  | (10.15)        |
| Black female   | 267.14*               | 272.16*        | 366.52*             | 375.65*        | 0.026*                    | 0.027*         | 0.033*                  | 0.034*         |
|  | (2.08)                | (2.40)         | (3.44)              | (3.81)         | (2.41)                    | (2.79)         | (3.91)                  | (4.29)         |
| <b>Neighborhood variables:</b>                         |                       |                |                     |                |                           |                |                         |                |
| Income $\times 10^{-3}$                                | 3.66                  | 2.98           | -0.21               | -0.26          | 0.0002                    | 0.0002         | -0.0001                 | -0.0001        |
|  | (0.58)                | (0.42)         | (-0.04)             | (-0.05)        | (0.43)                    | (0.30)         | (-0.20)                 | (-0.16)        |
| Suburb <sup>a</sup>                                    | -188.08               | -165.09        | 162.76              | -156.73        | -0.014                    | -0.012         | -0.01                   | -0.009         |
|  | (1.20)                | (-0.97)        | (-1.25)             | (-1.14)        | (-1.04)                   | (-0.84)        | (-0.91)                 | (-0.83)        |
| Black tester $\times$ suburb                           | 166.29                | 107.64         | 170.35              | 165.42         | 0.022                     | 0.018          | 0.021                   | 0.021          |
|  | (0.60)                | (0.43)         | (0.74)              | (0.76)         | (0.94)                    | (0.85)         | (1.13)                  | (1.18)         |
| Minority-owned dealer <sup>b</sup>                     | 102.25                | 77.53          | 138.70              | 135.49         | 0.004                     | 0.0006         | 0.009                   | 0.008          |
|  | (0.34)                | (0.25)         | (0.55)              | (0.53)         | (0.16)                    | (0.02)         | (0.44)                  | (0.40)         |
| Black tester $\times$ minority-owned dealer            | 62.80                 | 224.86         | -54.59              | -2.80          | 0.025                     | 0.04           | 0.008                   | 0.013          |
|  | (0.13)                | (0.52)         | (-0.13)             | (-0.01)        | (0.60)                    | (1.12)         | (0.26)                  | (0.43)         |
| Percentage black in neighborhood                       | -50.62                | -83.20         | 152.60              | 151.43         | -0.002                    | -0.006         | 0.015                   | 0.015          |
|  | (-0.18)               | (-0.27)        | (0.63)              | (0.60)         | (-0.09)                   | (-0.21)        | (0.77)                  | (0.74)         |
| Black tester $\times$ percentage black in neighborhood | -675.91               | -641.54        | -600.92             | -607.22        | -0.058                    | -0.054         | -0.051                  | -0.053         |
|  | (-1.20)               | (-1.26)        | (-1.28)             | (-1.37)        | (-1.22)                   | (-1.25)        | (-1.35)                 | (-1.47)        |
| <b>Seller interactions:</b>                            |                       |                |                     |                |                           |                |                         |                |
| Tester black male, seller white female                 | 380.74                | 378.81         | 175.65              | 130.66         | 0.035                     | 0.036          | 0.019                   | 0.016          |
|  | (0.93)                | (0.98)         | (0.52)              | (0.39)         | (0.10)                    | (1.08)         | (0.69)                  | (0.60)         |
| Tester black male, seller black male                   | 322.08                | 236.48         | 126.65              | 116.65         | 0.023                     | 0.017          | 0.009                   | 0.01           |
|  | (0.79)                | (0.61)         | (0.37)              | (0.35)         | (0.67)                    | (0.52)         | (0.34)                  | (0.39)         |
| Tester black female, seller white female               | 115.15                | 182.18         | -73.29              | -85.86         | 0.025                     | 0.029          | 0.006                   | 0.004          |
|  | (0.31)                | (0.52)         | (-0.24)             | (-0.29)        | (0.79)                    | (0.10)         | (0.25)                  | (0.18)         |
| Tester black female, seller black male                 | 271.72                | 329.59         | 227.13              | 229.96         | 0.022                     | 0.027          | 0.023                   | 0.022          |
|  | (0.77)                | (0.98)         | (0.78)              | (0.80)         | (0.74)                    | (0.10)         | (0.95)                  | (0.96)         |
| Tester white female, seller white female               | -389.30               | -76.35         | 69.65               | 162.79         | -0.026                    | -0.005         | 0.007                   | 0.011          |
|  | (-0.87)               | (-0.18)        | (0.19)              | (0.45)         | (-0.69)                   | (-0.14)        | (0.23)                  | (0.38)         |
| Tester white female, seller black male                 | 106.31                | 263.29         | 168.43              | -126.48        | 0.005                     | 0.019          | -0.015                  | -0.013         |
|  | (0.35)                | (0.93)         | (-0.66)             | (-0.52)        | (0.19)                    | (0.79)         | (-0.78)                 | (-0.67)        |
| Adjusted $R^2$ :                                       | 0.26                  |                | 0.33                |                | 0.27                      |                | 0.39                    |                |
| Standard error of the estimate:                        | 827.8                 |                | 687.2               |                | 0.071                     |                | 0.056                   |                |
| Degrees of freedom:                                    | 281                   |                | 281                 |                | 281                       |                | 281                     |                |
| N:   | 306                   |                | 306                 |                | 306                       |                | 306                     |                |

Notes: Numbers in parentheses are  $t$  statistics. See the text for explanation of the interpretation of the interaction effects. Note that all the regressions also include dummy variables for eight of the nine car models we tested, omitting the least expensive car.

<sup>a</sup>Dummy variable = 1 if dealership in suburb; 0 otherwise.

<sup>b</sup>Dummy variable = 1 if minority-owned dealer; 0 otherwise.

\* Significantly different from zero at the 5-percent level.

and gender of the dealership's owner, employees, or customers influenced the amount of discrimination.<sup>20</sup> One conclusion emerges from Table 5: the neighborhood effects have virtually no power in explaining discrimination. Moreover, the tester-type dummies still have roughly the same size and significance level when the neighborhood effects are included.<sup>21</sup> This finding, along with some ancillary information, argues against standard animus-based theories as the primary explanation for the discrimination we observed.

1. *Owner Animus*.—If owners are the source of animus, black-owned dealerships should presumably exhibit less discrimination against blacks than dealerships owned by whites.<sup>22</sup> The regressions in Table 5 con-

<sup>20</sup>To simplify interpretation of the interaction terms, we follow the procedure suggested by John Yinger (1986 p. 888). Consider, for example, the interaction of a white female tester with a black male seller. Let SBM be a dummy variable which is 1 if the seller is a black male, and let TWF be a dummy that is 1 if the tester is a white female. Then the normal interaction specification would be  $TWF \times SBM$ . Instead, we use

$$TWF \times \left( SBM - \sum_{i=TWF} SBM_i / N_{TWF} \right).$$

That is, we subtract the average value of SBM for all white female testers (simply the percentage of all white female tests in which the seller was a black male). This subtraction allows us to interpret the TWF coefficient as the average level of discrimination facing white female testers, while the interaction term represents the marginal effect of facing a black male seller.

<sup>21</sup>These regressions also contain dummy variables for eight of the nine car models (omitting the least expensive car). In other regressions, we investigated whether tester groups encountered different dealer behavior when bargaining for foreign or luxury cars. We included variables interacting the tester dummies with dummies indicating whether the car model was domestically produced and whether the car was an economy or standard/luxury model. These interaction coefficients were uniformly small and insignificant, indicating that the pattern of discrimination does not depend on the model class or whether the car is manufactured domestically.

<sup>22</sup>Because there are so few black-owned dealerships (we were able to identify only nine in our sample), they might be able to "free-ride" on the market discrimination by charging their black customers a price that

tain a dummy variable MINOWN for minority-owned dealerships, as well as an interaction dummy (equaling 1 when both the tester and the owner were black). None of these coefficients was significant in any of the regressions, indicating that the seller's race did not influence the bargaining outcome and that black testers did not fare better at black-owned dealerships.

Bigoted owners should also be more likely to discriminate against their own employees (with whom they presumably have to associate closely over an extended period of time) than against their customers. Given that owners are willing to hire nonwhite and nonmale *salespeople* (who comprised nearly one-fourth of those encountered in our tests),<sup>23</sup> it seems implausible that they would need a \$500 higher markup to compensate for selling to *customers* who are not white males.

2. *Employee Animus*.—Employees—in this case, salespeople—rather than dealership owners are another possible source of animus. Again, however, the magnitudes of the discrimination observed do not seem consonant with this source. For example, it is difficult to imagine that the \$1,000 additional markup to black males represents compensation to white salespeople for the disutility of having to spend 45 minutes negotiating with them. The interaction effects in the regressions of Table 5 test whether the gender and race of the salesperson affected the amount of discrimination. The coefficients were uniformly insignificant,

---

reflects the discriminatory premium at white-owned dealerships. It is also theoretically possible that black owners dislike dealing with blacks, but at a minimum this would implicate a nontraditional form of discrimination. We used the "Black Pages," (an analogue of the Yellow Pages which lists firms that are more than 50-percent black-owned), supplemented by a City of Chicago listing of minority-owned businesses as sources. Listing in either source is voluntary, so we may have excluded some black-owned dealerships. We were unable to find any female-owned dealerships in analogous sources for women.

<sup>23</sup>Of the 306 tests, 23.2 percent involved salespeople who were not white males.

casting doubt on employee animus as a source of discrimination. Black testers did no worse when buying from white salespeople, nor did women get worse deals when the salesperson was male,<sup>24</sup> as one would expect if salesperson animus were the motive for discrimination. In addition, we would expect that bigoted salespeople would want to spend less time with non-white-male testers than with white males. In fact, however, salespeople spent nearly 13-percent longer negotiating with the “minority” testers than with the white males, which casts doubt on salesperson animus as the source of price differences.<sup>25</sup> We do not want to exclude this hypothesis completely, however, because our testers did report some instances of explicitly hostile racist or sexist language from the sellers.<sup>26</sup>

3. *Customer Animus.*—Fellow-customers should also be considered as a source of

<sup>24</sup>In fact, an earlier pilot study (Ayres, 1991) found that women and blacks did *worse* when negotiating against a salesperson of the same race/gender and got their best deals, on average, from white males. The earlier study also found that testers were more likely to “draw” a salesperson of the same race and gender as themselves (who then tended to charge higher prices). We detected no evidence of steering in these data, however. The race and gender of testers and sellers appeared to be independent of each other, as the following table demonstrates:

| Tester       | Seller     |              |            |              |
|--------------|------------|--------------|------------|--------------|
|              | White male | White female | Black male | Black female |
| White male   | 123        | 11           | 17         | 2            |
| White female | 37         | 4            | 11         | 1            |
| Black male   | 46         | 6            | 7          | 1            |
| Black female | 29         | 5            | 6          | 0            |

( $N = 306$  observations;  $\chi^2_{[9]} = 5.64$ ,  $p = 0.78$ ).

<sup>25</sup>The average test by a white male lasted 36.2 minutes. The average for the other testers was 40.8 minutes. Although the 4.6-minute difference is small, a  $t$  test reveals that the two means are significantly different at the 0.001 level. The shorter negotiations with white males could be explained by a dealer preference for wasting the time of minority buyers.

<sup>26</sup>Dealers made hostile race- or gender-based statements in about 4 percent of the tests.

animus against black or women shoppers. For example, a dealership might charge more to black or female consumers if their presence in the showroom made it less likely that others (whites, men) would shop there. The evidence for this kind of discrimination is mixed. First, concerns about the reactions of other customers should lead dealers to shepherd blacks and women out of the showroom as rapidly as possible (to avoid their being seen by other potential customers). Yet it was white male testers, rather than blacks or white women, who had the shortest average negotiating sessions. In addition, the customer-based theory implies that the extent of customer prejudice should vary by neighborhood: black testers buying in a white neighborhood (where most other customers are likely to be white) should do worse than in a black neighborhood.<sup>27</sup> Table 5 does indicate that black testers shopping in black neighborhoods may receive lower offers. The coefficients on these interaction variables were large (roughly  $-\$500$ ), but were poorly measured ( $t$  statistics between  $-0.9$  and  $-1.3$ ).<sup>28</sup>

In sum, animus-based theories of discrimination do not find support in our analysis of dealership characteristics. Consistent with Becker’s analysis of the effects of competition on discrimination, the evidence for owner and employee animus is the weakest. The large but insignificant coefficients for

<sup>27</sup>Neighborhoods are defined by what the City of Chicago (1992 p. 1) calls “Community Areas.” Each has 30,000–60,000 people. “Community areas are defined by the city of Chicago as groups of census tracts. They were first identified in the 1930’s by the Social Science Research Council of the University of Chicago. They correspond roughly to informally recognized neighborhoods such as Lakeview and Hyde Park. The boundaries have changed very little since their inception....”

<sup>28</sup>John Yinger (1986) concludes that, in the housing market, discrimination against blacks is motivated by realtors’ perceptions that other renters or house buyers would disapprove of having a black neighbor. Interestingly, Yinger finds, as we do, that black females encounter substantially less discrimination than black males (p. 891). Because of the discrete nature of automobile purchases, animus by fellow consumers strikes us as inherently more plausible in the housing context than in the automobile showroom.

black testers in black neighborhoods might raise some concern, however, that customer animus may play a role in the higher prices charged to blacks.

### B. *Statistical Theories*

“Statistical discrimination” is based not on a psychological distaste for associating with blacks or women, but rather on sellers’ use of observable variables (such as race or gender) to make inferences about a relevant but unobservable variable (Edmund Phelps, 1972). In the labor market, productivity is the relevant unobservable variable. 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. If sellers believe, for example, that women are on average more averse to bargaining than men, it may be profitable to quote higher prices to women customers.<sup>29</sup>

Dealers might also make racial or gender inferences about the expected costs of contracting—including, for example, the expected costs of default on car loans. The script attempted to eliminate cost-based statistical discrimination by having all testers volunteer early in the negotiations that they were providing their own financing. This disclosure indicated that the dealer should not have to bear a risk of buyer default.<sup>30</sup>

<sup>29</sup>As one car salesman turned consumer advocate (Darrell Parrish, 1985 p. 3) wrote:

...Salesmen... categorize people into “typical” buyer categories. During my time as a salesman I termed the most common of these the “typically uninformed buyer”... [In addition to their lack of information, these] buyers tended to display other common weaknesses. As a rule they were indecisive, wary, impulsive and, as a result, were easily misled. Now take a guess as to which gender of the species placed at the top of this “typically easy to mislead” category? You guessed it—women.

<sup>30</sup>Even though all testers volunteered that they did not need financing, dealers might disparately assess the credibility of this information depending on the gender and race of the tester. If statistically valid, this inference could form the basis for cost-based statistical discrimination.

The plausibility of revenue-based statistical discrimination as an explanation for our results is heightened by the fact that salespeople have their own term for a kind of statistical discrimination, which they call “qualifying the buyer.” “Qualifying” is the process of estimating how much the buyer is willing/able to pay on the basis of direct observation (how the buyer is dressed, what kind of car she is currently driving) and answers to questions the seller asks (“How did you get to the dealership?” “Have you visited other dealerships?”). This section looks at possible causes of statistical discrimination and examines whether they are consistent with the evidence.

1. *Search Costs.*—Sellers might perceive that race and gender are related to buyers’ search costs for several reasons. For example, black consumers might have higher search costs because they are less likely than whites to own a car at the time they are shopping for a new one (and therefore might have more difficulty traveling to multiple dealerships) (Fred Mannerling and Clifford Winston, 1991 p. 98).<sup>31</sup>

---

Profit-maximizing dealers might also make inferences about the profits from ancillary sales, so statistical discrimination could also be caused by dealers’ inferences about the likelihood of repeat purchases, referrals, or repair service. To dampen the importance of such inferences, testers initially volunteered to salespeople that they were moving out of the state within a month. However, having more than one tester make this representation at a single dealership increased the likelihood that dealers would suspect a test, and so this was discontinued.

<sup>31</sup>There is a large, uneven, and largely dated marketing literature which does seem to support the notion that “[v]ariation in prepurchase search behavior is related to racial differences” (Carl Block, 1972 p. 9). Laurence Feldman and Alvin Starr (1968 pp. 216–26) also conclude that there are differences in search behavior by race, although they find that these diminish after controlling for income. For a survey of studies examining differences in car ownership rates by race, see Raymond Bauer and Scott Cunningham (1970 pp. 157–60), who conclude that blacks are less likely to own a car than whites (even when controlling for income). For a theoretical examination of some possible effects of differences in search costs on the ability of sellers to discriminate, see Robert Masson (1973).

Our data provide some evidence that sellers considered search costs to be (differentially) important for different groups of testers. Non-white-male testers were more than 2.5 times as likely as white males to be asked how they got to the dealership, suggesting that dealers show particular interest in determining whether non-white-males have substantial opportunities to search. In addition, testers who revealed that they did not own a car were asked to pay \$127 more while those who indicated that they had visited other dealerships saved \$122 (although these results were only significant at a 20-percent level).<sup>32</sup>

The evidence does not uniformly support a search-cost explanation, however. If sellers are sensitive to buyers' search costs in setting prices, we would expect black testers to receive better deals (relative to whites) in the suburbs than at urban dealerships. By traveling the substantial distance from the center city to the Chicago suburbs (where very few blacks live), blacks are in effect signaling to suburban dealers that they are willing and able to undertake an extensive search for a car. Contrary to this theory, however, the coefficient for blacks negotiating in suburbs was a positive \$64 for initial offers and \$264 for final offers (*t* statistics of 0.25 and 1.22, respectively). Moreover, initial and final offers for black testers in all-white neighborhoods were \$675 and \$600 higher than in all-black neighborhoods (*t* statistics of 1.20 and 1.28). The presence of black customers in white neighborhoods did not signal a willingness to search that translated into lower dealer offers.<sup>33</sup>

<sup>32</sup>These figures were derived by constructing dummy variables for the tests in which this information was revealed and by then rerunning the final-profit regression in Table 1. Testers revealed this information, however, only when dealers asked, and the decision to ask might not be exogenous to the dealer's final offer.

<sup>33</sup>Customer animus and search-cost theories of discrimination may not be independent. Neighborhoods with few black residents may also have stronger animus against black customers. This animus might swamp (or at least confound) the signaling effect and cause blacks bargaining in such neighborhoods to be quoted higher prices.

2. *Consumer Information.*—Race or gender may also be correlated with buyers' information about the car market. For example, a recent study by the Consumer Federation of America (1991) found that 37 percent of respondents did not believe that the sticker price on a car was negotiable. More important for our purposes, there were wide differences in consumer knowledge by race and gender. Sixty-one percent of blacks surveyed believed the price was not negotiable, while only 31 percent of whites believed this. Women were more likely than men to be misinformed about the willingness of dealers to bargain, although the disparities were not as great as between blacks and whites.<sup>34</sup>

Sellers in our study may have been motivated in part by such informational disparities in quoting higher prices to blacks and women. Dealers were somewhat more likely to volunteer information about the cost of the car to white males than to the other testers, possibly because they believed that white males already had such information.<sup>35</sup> More significantly, dealers made their first offer at the sticker price to 29 percent of nonwhite males, but initially offered the sticker price to only 9 percent of white male testers ( $X^2_{[1]} = 25.9$ ). This suggests that sellers believed white males were more knowledgeable than other testers about the possibility of bargaining over sticker prices.

3. *Bargaining Costs.*—Another type of statistical discrimination might focus on

<sup>34</sup>Because the survey respondents were not limited to people who were actually interested in buying a car, the survey may overstate racial differences in information among the car-buying public (if nonbuying blacks are relatively less informed than nonbuying whites). George Moschis and Roy Moore (1981 p. 261), however, find "differences in consumer knowledge, skills and attitudes between blacks and whites" controlling for actual purchase behavior.

<sup>35</sup>White males were given unsolicited cost information in 55 percent of their tests, while all other testers were given such information in 48 percent of tests. The difference was not statistically significant at the 5-percent level, however.



buyers' aversion to conducting negotiations. Some buyers experience the process of bargaining as costly, while others apparently derive pleasure from the give-and-take of negotiations. The key question in the present context is whether these bargaining costs are correlated with race or gender. We have no direct evidence on this point,<sup>36</sup> but several findings suggest that dealers made it procedurally more difficult to purchase a car for "minority" testers. First, non-white-male testers were more often asked to sign purchase orders (40.2 percent vs. 27.6 percent for white males;  $X^2_{[1]} = 7.14$ ) and to put down a deposit (37.7 percent vs. 25.6 percent for white males;  $X^2_{[1]} = 6.78$ ). "Minority" testers were also much more likely than white males to be "bumped," that is, to have the dealership manager raise a salesperson's offer (7.0 percent vs. 1.5 percent;  $X^2_{[1]} = 7.66$ ). Forcing non-white-male testers to overcome these additional procedural hurdles might have been one way dealers tried to take advantage of what they perceived as the higher aversion to bargaining of "minority" testers relative to white males.

#### IV. Conclusion

In negotiations for more than 300 new cars, Chicago car dealers offered black and female testers significantly higher prices than the white males with whom they were paired, even though all testers used identical bargaining strategies.<sup>37</sup> This race and

gender discrimination is a robust result, and the magnitude of discrimination is large enough that it cannot credibly be attributed to residual nonuniformities between types of testers.

It is much more difficult to explain discrimination than to document its existence. The evidence that dealers occasionally used racist or sexist language, combined with the large (but statistically insignificant) savings that black testers encountered in black neighborhoods, justify lingering concerns about the possibility of animus or stereotypical discrimination. However, dealership conduct may be better explained as a form of statistical discrimination in which dealers use race and gender as a proxy for the customer's reservation price.

It may be that simple theories of discrimination fail to capture the mutually reinforcing nature of multiple causes. In the end, it may prove impossible to parse out the various elements of animus and rational inferences from irrational stereotypes. No single theory may be adequate to explain the observed discrimination against black men, black women, and white women. Whatever its causes, however, the discrimination uncovered in this study stands squarely in the face of earlier analyses that reject the possibility that discrimination can persist in a competitive market.

#### REFERENCES

- Ayres, Ian. "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations." *Harvard Law Review*, February 1991, 104(4), pp. 817-72.
- \_\_\_\_\_. "Further Evidence of Discrimination in New Car Negotiations and Estimates of Its Cause." Mimeo, Yale Law School, 1994.
- Ayres, Ian and Siegelman, Peter. "Race and Gender Discrimination in Bargaining for a New Car." Working paper, American Bar Foundation, Chicago, 1992.
- Bauer, Raymond and Cunningham, Scott. *Studies in the Negro market*. Cambridge, MA: Marketing Science Institute, 1970.
- Becker, Gary. *The economics of discrimination*. Chicago: University of Chicago Press,

<sup>36</sup>The social-psychology literature does not seem to have reached a firm conclusion on whether women bargain differently from men or blacks from whites (Jeffrey Rubin and Bert Brown, 1975). A recent newspaper article (Warren Brown, 1990) interviewed dealers who suggested that middle-class black males associated needing to bargain with poverty and were thus reluctant to bargain.

<sup>37</sup>We should stress that these results may not be generalizable to a random sample of car buyers. Our testers employed a uniform bargaining strategy. If this strategy is atypical of the population at large, or if black and white consumers bargain differently, then the equilibrium amount of discrimination could be higher or lower than we observed.

- 1957; 2nd. Ed., 1975.
- Block, Carl.** "Prepurchase Search Behavior of Low-Income Households." *Journal of Retailing*, Spring 1972, 48(1), pp. 3-15.
- Brown, Warren.** "Who Gets the Best Deals on Wheels? Chicago Study Finds It's White Males." *Washington Post*, 14 December 1991, p. F1.
- City of Chicago.** "Chicago Statistical Abstract." Mimeo, 1992.
- Consumer Federation of America.** "U.S. Consumer Knowledge: The Results of a Nationwide Test." Washington, DC: Consumer Federation of America, 1990.
- Cross, Harry; Kenney, Genevieve; Mell, Jane and Zimmerman, Wendy.** *Employer hiring practices: Differential treatment of Hispanic and Anglo job seekers.* Urban Institute Report 90-4, Washington, DC, 1990.
- DeGroot, Morris H.** *Probability and statistics*, 2nd Ed. Reading, MA: Addison-Wesley, 1986.
- Feldman, Laurence and Starr, Alvin.** "Racial Factors in Shopping Behavior," in K. Cox and B. Enis, eds., *A new measure of responsibility for marketing.* Chicago: American Marketing Association, 1968, pp. 205-24.
- Fix, Michael and Struyk, Raymond.** "An Overview of Auditing for Discrimination," in Michael Fix and Raymond Struyk, eds., *Clear and convincing evidence.* Washington, DC: Urban Institute Press, 1992, pp. 1-67.
- Greene, William.** *Econometric analysis.* New York: Macmillan, 1990.
- Hausman, Jerry and Taylor, William.** "Panel Data and Unobservable Individual Effects." *Econometrica*, November 1981, 49(6), pp. 1377-98.
- Heckman, James and Siegelman, Peter.** "The Urban Institute Studies: Their Methods and Findings," in Michael Fix and Raymond Struyk, eds., *Clear and convincing evidence.* Washington, DC: Urban Institute Press, 1992, pp. 187-258.
- Mannering, Fred and Winston, Clifford.** "Brand Loyalty and the Decline of American Automobile Firms." *Brookings Papers on Economic Activity*, Microeconomics 1991, pp. 67-103.
- Masson, Robert.** "Costs of Search and Racial Price Discrimination." *Western Economic Journal*, June 1973, 11(2), pp. 167-86.
- Mateja, Jim.** "Escort Sticker Shock: One Price Sells 'em All." *Chicago Tribune*, 13 March 1992, p. C1.
- McIntyre, Shelby; Moberg, Dennis, J. and Posner, Barry Z.** "Preferential Treatment in Preselection Decisions According to Race and Sex." *Academy of Management Journal*, December 1980, 23(4), pp. 738-49.
- Moschis, George and Moore, Roy.** "Racial and Socioeconomic Influences on Adolescent Consumer Behavior," in *Proceedings of the American Marketing Association*, Chicago: American Marketing Association, 1981, pp. 261-65.
- Newman, Jerry.** "Discrimination in Recruitment: An Empirical Analysis." *Industrial and Labor Relations Review*, October 1978, 32(1), pp. 15-23.
- Parrish, Darrell.** *The car buyer's art.* Bellflower, CA: Book Express, 1985.
- Phelps, Edmund.** "The Statistical Theory of Racism and Sexism." *American Economic Review*, September 1972, 62(4), pp. 659-61.
- Rosenthal, Robert.** *Experimenter effects in behavioral research*, 2nd Ed. New York: Irvington, 1976.
- Rubin, Jeffrey and Brown, Bert.** *The social psychology of bargaining and negotiation.* New York: Academic Press, 1975.
- Salop, Steven and Stiglitz, Joseph.** "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion." *Review of Economic Studies*, October 1977, 44(3), pp. 493-507.
- Srinivasan, Narasimhan and Wen, Kuang-Wei.** "Price Negotiations in New Car Purchases: A Game Theoretic Experimental Approach." Mimeo, School of Business Administration, University of Connecticut, 1991.
- Stigler, George.** "The Economics of Information," in *The organization of industry.* Chicago: University of Chicago Press, 1968, pp. 171-90.
- Swoboda, Frank.** "The Sacking of a Sales Force: Fla. Dealer Sells Twice as Many

- Cars After Firing Showroom Staff.” *Washington Post*, 27 February 1992, p. B9.
- Turner, Margery; Fix, Michael and Struyk, Raymond.** *Opportunities denied, opportunities diminished: Discrimination in hiring.* Washington, DC: Urban Institute Press, 1991.
- Yinger, John.** “Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act.” *American Economic Review*, December 1986, 76(5), pp. 881–98.