

CHAPTER 6

Discrimination in Consummated Car Purchases

Ian Ayres¹

ABSTRACT

Recent studies have examined the question of whether there is racial discrimination in car sales by analyzing newly available data about actual, consummated purchases of cars. This article reviews this research and compares it to my previous findings based on audit studies, in which Black and White testers were sent to car dealerships to negotiate, but not consummate sales. The studies using consummated purchase data contain strikingly similar results to those of the audit studies, and together they provide compelling evidence that Blacks pay substantially more for cars than Whites. The existence of this price differential is consistent with claims of racial discrimination and effectively rebuts claims that the audit studies are flawed because they do not reflect bargaining tactics that minorities could use in traditional, face-to-face car sales that would nullify the impact of discrimination. The new data also reveal that there is a much smaller price differential when Internet referral services are used, which is also consistent with claims of racial discrimination since the race of the buyer is harder to discern over the Internet.

INTRODUCTION

Two noted scholars have questioned whether audit tests of disparate treatment can provide compelling evidence of the economic injury borne by Blacks who in equilibrium might find a variety of ways mitigate its impact. For example, in critiquing my previous audit studies of new car sales, Epstein (1994), has argued that:

[I]n open markets customers are free to select not only their bargaining strategies but also the dealerships they visit. If blacks or women know that they are apt to get a good deal from some small fraction of the market, then they can avoid other, less receptive dealerships and their unattractive offers. How much of the differential found by Ayres would thus have disappeared is hard to say. In addition it may be possible for a buyer to reduce the differentials even further by bringing along a friend, by eliciting a rival offer from another dealer over the telephone. . . . These tactics are of course also open to white males, but given the lower bids that they are able to elicit, they are likely to yield better returns when adopted by others who anticipate that they will be offered higher prices. (pp. 52–53)

¹ This chapter updates Chapter 4 of Ayres (2001).

Epstein is somewhat agnostic (it is “hard to say”) about the extent to which consumers might be able to avoid the effects of discrimination by patronizing dealerships that they know to be non-discriminating, but Heckman (1998) makes a more far-reaching claim:

[Audit evidence of disparate racial treatment] is entirely consistent with little or no market discrimination at the margin. *Purposive sorting within markets eliminates the worst forms of discrimination*. There may be evil lurking in the hearts of firms that is never manifest in consummated market transactions. (p. 103)²

Without benefit of any empiricism, Heckman argues a theoretical possibility as an established fact. In essence, Heckman and Epstein are arguing that victim self-help could produce an equilibrium in which it would be “as if” discrimination did not exist. Blacks would receive (almost) the same prices as Whites.³

However, the audit tests themselves provide powerful evidence that African-Americans cannot protect themselves from the effects of discrimination by merely searching for and shifting their consumption to non-discriminating dealers (Ayres, 2001, pp. 70–71). An important finding of the audit regressions concerns the perva-

² The importance of these kinds of general equilibrium concerns has long been recognized in the labor market context. See, e.g., Siegelman (1999), Flinn and Heckman (1983), and Duleep and Zolotar (1991). Yinger (1997) formalizes the intuition that discrimination by sellers reduces the benefits of additional searches by buyers, causing them to accept higher prices or lower quality than they otherwise would. Yinger applies this methodology to the housing market, and finds that the costs of discrimination are roughly \$4000 per minority household per search.

³ Epstein (1994) has also criticized the audit approach for the small proportion of observations in which dealers attempted to accept a tester offer:

[I]n [Ayres’s] sample there were apparent contracts (i.e., a verbal agreement on the price that was not binding) in only 25 percent of the cases with white males, and in 15 percent of the cases with the remainder of his sample. The market would be in a state of perpetual turmoil if huge percentages of potential buyers were unable to buy cars at all. A technique of testing that leaves so many incomplete transactions cannot be an accurate replica of a functioning market. (p. 34)

This criticism ignores the structure of the audit. The testers were instructed to bargain until the dealership refused to negotiate further or attempted to accept one of their offers. The test focused on the lowest offer the dealer was willing to make—before refusing to negotiate further or in accepting the tester’s predetermined offer. We could have had 100% of the observations end in a non-binding verbal agreement (which would have satisfied Epstein’s definition) if we had merely instructed the testers to accept the dealer’s last and lowest offer at the point the salesperson refused to bargain further.

And, in Chapter 3 of Ayres (2001), Peter Siegelman and I 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. We found that sessions ending in attempted acceptances had an approximately \$400 lower final profit than those that ended in a refusal to bargain (and this result was statistically significant). The size of this acceptance effect, however, was the same for all testers. But the fact that sellers are more likely to accept offers from White males actually biases our estimates *against* finding discrimination because acceptances provide only 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.

siveness of discrimination across dealerships. The regressions suggest that Black consumers could not protect themselves by patronizing minority, and/or women-owned dealerships or by striving to bargain with minority and/or women salespeople. As Goldberg (1996) concluded:

[The audit] experiment offers some direct evidence on the issue [of whether purposive sorting by testers mitigated the effect of dealership disparate treatment]: testers in the controlled experiment visit various dealerships in the Chicago area, some of which are located in poorer areas or in black neighborhoods, yet there is no evidence of any difference in the treatment minorities receive in such locations. (p. 643)

In short, our testing of over 400 dealerships in Chicagoland found pervasive prejudice. There were no statistically significant “safe harbor” dealerships where our testers could confidently go and uniformly bargain for a deal as good as White testers received.⁴

Still the audit does not exclude the possibility that minority purchasers might have been able to get a better deal at the same dealership if they had employed a different bargaining strategy. It is important to emphasize that forcing these minorities to use a different type of negotiation might itself represent an important type of race discrimination. This is especially true if the alternative path to a good deal was significantly more onerous. If Whites need only bargain for four hours to negotiate a low markup, but Blacks need to negotiate for eight hours, then a finding that Blacks in equilibrium paid the same for cars would not mean that Blacks were not injured by the dealerships’ disparate treatment. However, it might as a theoretical matter be possible that the dealership does not require more onerous bargaining but merely different bargaining by Blacks and Whites: Whites may be penalized if they speak in a Black voice and vice versa. This “separate but equal” possibility would still be a form of disparate racial treatment, but the harms from discrimination would be more contestable.⁵

Accordingly, it is useful to test whether Blacks pay more in actual consummated transactions. Actual sales of course are not controlled tests, and while multivariate regressions might control for a handful of purchaser characteristics, it is, as a practical matter, impossible to econometrically control after the fact for the myriad of different ways that purchasers might bargain. Therefore, an analysis of consummated transactions does not provide independent evidence of disparate treatment. A finding that Blacks pay more than Whites does not—by itself—indicate that dealerships engage in race-based bargaining. Dealerships might bargain uniformly with all potential customers—conceding at a uniform rate against all potential purchasers—with White customers on average holding out for better deals than Black customers. Instead, a finding of disparate transaction prices would be at a minimum evidence that the dealerships’ decisions to haggle (as opposed to the no-haggle policies of Saturn and others) have a disparate impact on Black purchasers. But when combined with the preceding

⁴ A possible exception to this might be so-called no-haggle dealerships that will be discussed below.

⁵ The social meaning of separate but equal regimes still can work substantial injury on traditionally subordinated people. See Rubinfeld (1998).

evidence of disparate treatment in dealership offers, a finding that Black purchasers pay more on average than White purchasers suggests that neither (1) purposive sorting, nor (2) alternative (and possibly more onerous) bargaining, nor (3) choosing not to purchase eliminates the effects of racially disparate dealership offers.

This chapter extends the meta-analysis first done in Ayres (2001) to compare seven different datasets of consummated and audit transactions:

- (1) The Chicago “Pilot” Audit (Ayres, 1991): An audit analysis of six testers (three White males, one Black female, one Black male, one White female) at approximately 90 Chicagoland dealerships.
- (2) Chicago Full Audit (Ayres, 2001): An audit analysis of 38 testers (18 White males, 7 White females, 5 Black males, and 8 Black females) at 242 Chicagoland dealerships.
- (3) National Consumer Expenditure Survey (CES) Consummated Sales (Goldberg, 1996): A consummated sales analysis of national CES data (67 minority observations, 1,212 White observations).
- (4) Atlanta Consummated Sales (Ayres, 2001): A consummated sales analysis of litigation-generated data from a single dealership (over 800 observations, approximately half were Black purchasers).⁶
- (5) J.D. Powers 2000 National Consummated Sales (Harless and Hoffer, 2002): A consummated sales analysis of national J.D. Powers data (4,030 observations).
- (6) National Consummated Sales—Traditional Bargaining (Scott Morton, Zettelmeyer, and Silva-Risso, 2003): A consummated sales analysis of national J.D. Powers-like data of traditional dealership sales (671,468 observations).
- (7) National Consummated Sales—Autobytel Purchases (Scott Morton et al.): A consummated sales analysis of national J.D. Powers-like data of Autobytel referred sales. (Over 20,000 observations of which approximately 7,300 were African-American purchasers.)

This chapter’s thesis is that the consummated transaction datasets are consistent with audit study findings that dealerships offer higher prices to Black consumers. Both the CES and Atlanta data on actual purchases show that Whites pay lower average prices than minorities (for CES data) and, more specifically, Blacks (for Atlanta data). Moreover, the sizes of the differentials are broadly similar. While the racial price differences are not statistically significant in Goldberg’s (1996) analysis of the CES data, Ayres (2001) showed that this insignificance was a function of the noisiness of the data. In the less noisy audit and Atlanta purchases data, we observe similar coefficients and smaller standard errors—which allow us to identify the racial differentials as statistically significant. The newer data from the market research firms continue to confirm the absence of a gender effect and the general presence of a statistically significant race effect.

The first part of this chapter analyzes the datasets examined in recent articles by Harless and Hoffer (2002); Scott Morton et al. (2003); and Scott Morton, Zettelmeyer,

⁶ An extended discussion of the first four datasets is included in Ayres (2001).

and Silva-Risso (2002). These articles exploit newly available market research data from J.D. Powers & Associates and an unnamed research firm in the automotive industry, which can provide the direct evidence on vehicle profits. The second part includes a meta-analysis of the seven datasets.

1. THREE NEW TRANSACTION STUDIES USING DEALER MARKET RESEARCH DATA

In the last two years, two new published articles (Harless and Hoffer, 2002; Scott Morton et al., 2003), and one unpublished study (Scott Morton et al., 2002) have exploited transaction data from J.D. Powers & Associates and an unnamed research firm in the automotive industry. These data include crucial information about vehicle profits for thousands of dealerships instead of just the single Atlanta dealership analyzed in Ayres (2001).⁷ While this type of data has presumably been available to manufacturers for years, the recent academic access makes this a particularly exciting time to conduct automobile pricing studies—especially because the negotiation equilibrium may dramatically change with the resurgence of no-haggle dealerships on the one hand and the rise of Internet sales, referral services and research on the other.

1.1. The Harless and Hoffer Study

The first published study based on the new J.D. Powers data was by Harless and Hoffer (2002). Harless and Hoffer analyzed 4,030 purchases made from more than 2,300 dealerships (part of the Powers Information Network) in February 2000. The authors admirably emphasize a number of weaknesses with their data:

First, the database does not include the race of the buyer. Hence, omitted variable bias is possible. If black males pay more than black females (as suggested in Ayres and Siegelman) then our model will not be as likely to detect a difference in profit between men and women. Second, to protect the privacy of dealers and customers, the database reports information in “cells” containing at least three transactions. Our dataset contains information on 4,030 transactions, but these 4,030 transactions are grouped in 414 cells with the cells containing from three to 74 transactions. Hence, our results are subject to the problem of ecological correlation (the possibility of drawing incorrect conclusions about individual outcomes from aggregated data), but the extent of this problem should be slight since the number of transactions per cell (median = 7) is quite small. Third, the J.D. Powers database reports dealer information only for the approximately 2,300 dealers who have been recruited to be included in the system. We cannot claim that this represents a random sample of dealers. While acknowledging that the sample of dealers is not random, it is noteworthy that vehicle manufacturers find the information sufficiently valuable to pay tens of thousands of dollars a month for access to the database. (Harless and Hoffer, 2002, pp. 271–272)

⁷ The datasets themselves may be less biased because the source was not involved in ongoing litigation. The data also include a wealth of information on other sources of profit (trade-in and finance) that are ripe for further analysis.

To my mind the even bigger concern with the study is that the researchers could not observe the gender of who bargained for the car. Gender inferences were based on a probabilistic inference about the name of the titled purchaser:

Buyers are classified as female or male as determined by a PIN proprietary probabilistic name program. If there are two or more buyers, the sex of the first recorded purchaser is used; we assume in such cases that the first name listed indicates the person who took the lead in negotiations. Misclassification would bias our results against finding differences between male and female buyers. Increasing our confidence in the accuracy of classification of sex by name, however, is that in 20 percent of all transactions the name could not be unambiguously assigned to a sex (and these cases are excluded from our sample). (Harless and Hoffer, 2002, p. 274)

Thus, if a man bargained but a woman was the titled owner, the data would report the transaction as being “female.” Given these data limitations, it should not be surprising that Harless and Hoffer did not find a statistically significant gender disparity. The study found that women paid an average of \$29 (or 1.9% higher mean gross profit) more for new cars than men, but the disparity was insignificant ($p = 0.47$) (2002, pp. 275–276, tab. 2, col. 2). Still, the finding of a small and statistically insignificant gender disparity conforms with the results of both the previous audit and purchase studies.

1.2. The Scott Morton, Zettelmeyer and Silva-Risso Study (2003)

Scott Morton et al. (2003) exploited an even larger dataset from an unnamed market research firm (which I would bet is also J.D. Powers). They analyzed 671,468 transactions at 3,562 dealerships concerning purchases made between January 1, 1999 and February 28, 2000. Unlike the Harless and Hoffer (2002) study, Scott Morton et al. observed individual transactions. But like Harless and Hoffer, Scott Morton et al. did not observe the gender or race of the bargainers. Like Harless and Hoffer, they inferred the gender of the titled purchaser by making probabilistic inferences about the purchaser’s first name. They made inferences about the purchaser’s race by exploiting census data about the racial composition of the “block group” (on average 1,100 people) where the purchaser resided. So this initial Scott Morton et al. study suffers from the same problems as Harless and Hoffer concerning gender identification. But the racial inference is not likely to be as problematic, if we believe that is less likely for cross-racial bargaining to occur (e.g., for a White to bargain on behalf of Black purchaser) than cross-gender bargaining. However, there can still be selection bias (for example, if Whites in a Black neighborhood are more likely to purchase a car).

With these limitations, Scott Morton et al. (2003) find that Black purchasers are expected to pay \$456 more than White purchasers,⁸ and this result is statistically

⁸ Literally a purchaser from an all black block group is predicted to pay \$456 more than a purchaser from an all white block group.

significant (p. 83, tab. 5, col. 1). Hispanic purchasers fared even worse—paying an estimated (and statistically significant) \$523 more than Whites, while Asian purchasers fared the best, paying \$218 less than Whites (again statistically significant). In contrast, female purchasers are predicted to pay just \$48 more than male purchasers (statistically significant).⁹

The Scott Morton et al. (2003) study is also important because it is the first and only analysis of whether Internet referral services tend to reduce the racial disparities now repeatedly found in traditional negotiations. Specifically, they tested whether purchasers who used Autobytel, the largest Internet referral service at the time, received systematically different deals. In particular, they tested whether there were fewer racial disparities because the service may have negotiated better terms with the dealership and because the dealership may have had a weaker racial signal than in face-to-face transactions.¹⁰

Their results are striking. They found that Autobytel users paid approximately \$273 less (1.2%) than non-Autobytel users (Scott Morton et al., 2003, p. 86, tab. 6, col. 1). Moreover, they found a marked decline in racial disparities: Black users were estimated to pay only \$68 more than White users (compared to the \$456 differential found in traditional sales); female users were estimated to pay \$21 more than male users (compared to the \$48 differential found in traditional sales); and Hispanic users were estimated to pay \$285 *less* than Anglo users (compared to the \$523 differential found in traditional sales) (Scott Morton et al., 2003, p. 86, tab. 6, col. 1).

Scott Morton, Zettelmeyer, and Silva-Risso have already followed up their excellent study (2003) with an unpublished second study (2002) which pairs the same type of vehicle profit information used in their first study (but limited to California sales in April and May of 2002) with responses from a survey that they mailed to 5,200 consumers—eliciting information about how informed the bargainer was and how he or she bargained (Scott Morton et al., 2002). The resulting dataset (after accounting for non-responses and incomplete surveys) had 1,507 observations.

The survey allows the authors to identify the gender and race of the real bargainers much more accurately than before. We learn for example that women purchasers were more likely to bring a man along to bargain than vice versa (69% of women versus 48% of men). While the paper estimates racial price results, they are generally insignificant because (as the authors emphasize) there are so few minorities in their final sample (for example, only 3.4% of the purchasers self-identified as African-American). But, even with much better gender information, the paper finds gender disparities identical

⁹ These disparities should be interpreted as evidence of disparate racial and gender impacts. When the authors control for a variety of non-racial and non-gender characteristics concerning the transaction, the racial disparities decrease. For example, the differential for black purchasers drops to \$342. However, the gender disparity remains \$48 (Scott Morton et al., 2003, p. 76, tab. 2, col. 1).

¹⁰ The race of a purchaser in an Autobytel transaction might still be inferred from the purchaser's name and address and, at times, from telephone conversations. Also Autobytel may not be able to protect purchasers from racial disparities in negotiating the price of a trade-in that must be done on a face-to-face basis.

to Scott Morton et al. (2003)—to wit, that women paid about a half a percent more than men in traditional bargains.

2. META-ANALYSIS OF THE SEVEN STUDIES

There now exist six different datasets to help assess whether Blacks are discriminated against in car purchasing. It is useful to take stock of the overall message of these data. While it is essential to undertake the micro-analysis of these data, it is also useful to step back and assess the broad contours of discrimination. Table 1 attempts just this task by summarizing the bottom line race/gender pricing differences. Of course, because the datasets come from such disparate sources, important caveats are in order. In interpreting this table, the reader should keep in mind the following:

- (1) The transactions differ: the audit testers solicited offers, but did not purchase cars; the Goldberg (1996), Atlanta (Ayres, 2001), Harless and Hoffer (2002) and Scott Morton et al. (2003) datasets include consummated transactions.
- (2) The price measures differ: the audit studies numbers are based on the profits implicit in the dealers' final offers; Goldberg's study is based on imputed differences in discounts from the sticker price; the Atlanta study is based on differences in total profit (including financing but excluding trade-in profit); and the Harless and Hoffer and Scott Morton et al. studies are based on just vehicle profits.
- (3) The controls differ: the audit testers used a uniform bargaining strategy and were controlled on a host of verbal and non-verbal dimensions while the completed transaction data has no *ex ante* control, and we lack basic information about how purchasers bargained, which makes it impossible to control *ex post* with regressions.
- (4) The racial groups differ: the audit, Atlanta, and Scott Morton et al. studies are tests of Black/White disparities (with the Scott Morton et al. study also examining disparities for Hispanics and Asians), while the Goldberg study is a test of "minority"/"non-minority" disparities, and the Harless and Hoffer data contain no racial information.
- (5) The geographic areas differ: the audit data come from Chicago; the Goldberg, Harless and Hoffer, and Scott Morton et al. studies are based on a nationwide sample; and the Atlanta data are of course from Atlanta.
- (6) The time periods differ: the pilot audit study was completed in 1989; the full audit study was completed in 1990; the Goldberg data covered transactions completed in 1983–1987; the Atlanta data covered transactions completed in 1990–1995; the Harless and Hoffer data covered February 2000; and the Scott Morton et al. study covered 1999 and the first 2 months of 2000.

Still, with all these caveats in mind, a global comparison of the race/gender differentials reveals striking similarities. The differentials for both Black males and Black females (i.e., the amount by which their profits exceeded the profits of White males)

Table 1. Meta-analysis of estimated dollar price premium over White males in four studies of markups on new cars, by demographic group

| Demographic group | White females | Black females | Black males | Adjusted <i>R</i> -squared |
|---|-------------------------------|--------------------------------|--------------------------------|----------------------------|
| Chicago “Pilot” Audit (Ayres, 1991) | \$220 (129) [21] | \$1,013*** (124) [23] | \$283*** (136) [18] | 0.37 |
| Chicago Full Audit (Ayres, 2001) | 216* (116) [53] | 465*** (103) [60] | 1,133*** (122) [40] | 0.33 |
| National CES Consummated Sales (Goldberg, 1996) | 129 (117) [244] | 426 (525) [28] | 274 (263) [39] | 0.14 |
| Atlanta Consummated Sales (Ayres, 2001) | -11 (97) [164] | 865*** (92) [224] | 611*** (96) [178] | 0.36 |
| J.D. Powers Feb. 2000 Sales (Harless and Hoffer, 2002) | 29 (38.9) [2,015]† | | | 0.76 |
| Scott Morton et al. (2003) | 47.9*** (3.2) [241,729] | 503.7*** (12.1) [14,383] | 445.8*** (11.6) [38,514] | 0.97 |
| Weighted Average | 47.8*** (3.7) [244,226] | 509.7*** (14.7) [14,718] | 457*** (12.4) [38,789] | |
| Scott Morton et al. (2003) (Autobytel purchases) | 20.5 (13.6) [6,800] | 88.9 (66.4) [204]‡ | 68.4 (65.0) [306]‡ | 0.98 |

Row 1, Ayres (1991).

Row 2, Ayres (2001), Chapter 2, Table 2.1, Col. 2.

Row 3, Goldberg (1996), Table 2, Col. 1 and Table 5. Goldberg tested for differences between “Minority” and “Non-Minority” purchasers.

Row 4, Ayres (2001), Table 4, Col. 6.

Row 5, Harless and Hoffer (2002), Table 2, Col. 1.

Row 6, Scott Morton et al. (2003), Table 5, Col. 1.

Row 8, Scott Morton et al. (2003), Table 6, Col. 2.

Standard errors in parentheses. Number of observations in brackets.

*Significantly different from zero at the 10% level.

**Significantly different from zero at the 5% level.

***Significantly different from zero at the 1% level.

†Number of female transactions assumed to be half of total sample size (omitted in article).

‡Cell size within Autobytel transactions assumed proportional to cell size over the full dataset.

are uniformly positive and substantial (ranging from \$274 to \$1,133, not including the Autobytel dataset). Moreover, of the five datasets addressing race in the traditional car sales market (i.e., excluding Harless and Hoffer, 2002 and Scott Morton et al., 2003, Autobytel data), four contain Black-male and Black-female differentials that are highly significant. Also in four of these same five datasets, Black women are

estimated to pay more than Black men. And for heuristic purposes, if we combine the observations from these five different sources, we find that on average Black men pay (or are finally asked to pay) \$457 more than White men, and that Black women are asked to pay \$510 more than White men. These weighted average differentials, when compared to the combined standard deviation, are also statistically significant.¹¹ Thus, while the differentials uncovered by Goldberg (1996) in the CES data continue to be statistically insignificant, the high standard deviation applied to her small number of observations (28 out of a total of more than 14,000 minority-female observations) is not sufficient to render the global analysis statistically insignificant.

Indeed, several things are striking when comparing the Goldberg (1996) differentials with the differentials from the other datasets. Goldberg's standard errors for minority males and minority females are more than twice (and sometimes three and four times) the size of the comparable differentials, and her *R*-squared is less than half all of the other regressions. These are strong indications of the noisiness of her data. But the sizes of the differentials themselves are not so far from some of the other datasets. Goldberg's minority-female differential of \$426 is similar to the \$465 Black-female differential from the full audit study. And Goldberg's minority-male differential of \$274 is similar to the Black-male differential estimated in the pilot audit.

Still it must be admitted that the sizes of the differentials—while robustly positive—do vary. Two of the Black-minority-male differentials are near \$300 and two others are more than twice this amount (\$611 and \$1,133). Three of the Black-minority-female differentials are in the \$450–\$500 range, but the other two are roughly twice this amount (\$865 and \$1,013). These higher differentials are not only statistically distinct from Goldberg's estimates (Goldberg, 1996), but also a joint test rejects the null hypothesis of equal means. Still, given the important differences in the ways these data were produced and analyzed, I believe the similarities of the table far outweigh the dissimilarities.

The global analysis shows that on net Blacks pay substantially more than Whites in both audit testing and consummated transactions. Once we appreciate the noisiness of Goldberg's data (Goldberg, 1996), her analysis does not contradict, but adds marginal confirmation to this result. Counter to the conjectures of Heckman (1998) and Epstein (1994), additional search and/or alternative bargaining strategies seen in traditional negotiations do not eliminate or even significantly mitigate the amounts of discrimination discovered in the initial audits. Customers often do not have sufficient information to take such self-help measures, and, given the recalcitrance that dealerships across

¹¹ The combined standard deviation was computed to be:

$$\frac{\sum_{i=1} s_i^2(n_i - 1)}{\sum_{i=1} n_i - 4},$$

where s_i = the standard deviation of the i th dataset and n_i = the number of observations of the i th dataset.

the board showed in the audit testing, it is not clear that effective self-help measures currently exist.

However there is some evidence that Internet referrals may ultimately help confirm the Heckman/Epstein conjecture. The last row of Table 1 reports the analogous differentials found in the Scott Morton et al. (2003) Autobytel data. The racial and gender disparities are markedly lower and not statistically significant (even though there was a substantial sample size and *R*-squared). Moreover, Scott Morton et al. (2002) found that women and minorities were more likely to use the Autobytel service. Together this suggests that the emerging practice of researching and shopping via the Internet (where race and gender characteristics may be less knowable by sellers) may prove to reduce the persistent racial disparities found in traditional auto sales. However, while the penetration of Internet sales has been phenomenal, it still represents a small proportion of the overall market and hence provides only a limited (but rapidly growing) confirmation of the Heckman/Epstein conjecture.

This evidence of racial disparities in consummated vehicle pricing is also consistent with similar analyses of racial disparities in finance profits. Both the Ayres (2001) analysis of the Atlanta dataset and the more recent litigation-generated analysis of major automotive lenders suggest that Black borrowers are much more likely to have their interest rate marked up above their risk-adjusted rate of interest. A class action suit against General Motor's credit division, General Motors Acceptance Corp. (GMAC), has uncovered that financing profits for (African-American) borrowers are systematically higher than for White borrowers (*Coleman v. General Motors Acceptance Corp.*, 2000).¹² The basic facts of the suit can be easily summarized (albeit in a slightly stylized fashion). When a GM dealer approaches GMAC about financing a particular purchaser and passes on core information about the financial risk of lending to such a car buyer, GMAC responds by telling the dealership the minimum amount of interest rate (the risk-adjusted market rate) that the dealer can charge. But GMAC also allows the dealership to negotiate a higher and more profitable interest rate up to some maximum amount. The dealer and GMAC split the profits on any excess interest that the dealer can negotiate.

A report of plaintiffs' expert Marc Cohen shows that, controlling for a host of other variables, the excess profit on loans to Black consumers is \$377 higher than for White consumers. These are not quite as high as the racial differentials in financial profits uncovered in the Atlanta dataset (which range from \$453 to \$637), but are nonetheless highly significant.

The plaintiffs in this litigation are using the Equal Credit Opportunity Act (ECOA),¹³ which allows plaintiffs to bring racial disparate impact suits in lending. In this case, the plaintiff class alleged that the financing company's decision to allow dealerships to negotiate had an unjustified disparate impact on African-American

¹² I have been retained as a plaintiff's expert in this case as well as a number of other cases challenging the disparate racial impacts of dealership markups.

¹³ See also 12 C.F.R. §202.1; Interagency Policy Statement, 1994 WL 128417.

borrowers. This statute has provided a new weapon to attack not just racial disparities in excess interest charged—but also to attack racial disparities in the underlying purchase price of the car (the principal of the loan).

3. CONCLUSION

In closing, it is appropriate to comment on the catch-22 created by the Epstein/Heckman critique. Discrimination tests are often plagued by difficulties of creating “similarly situated” comparisons. Heckman (1998), for example, has criticized some audits for not adequately controlling for unobserved variables—factors other than race or gender but correlated with these traits that might offer non-discriminatory explanation for the audit results. Defendants in discrimination suits *always* claim that their behavior was not predicated on the plaintiff’s race or sex but on some other characteristic.

The catch-22 (or what Margaret Radin (1990) calls a “double bind”) comes however when researchers produce an effective test where Blacks and Whites (men and women) *do* behave the same. Then comes Heckman (1998) claiming that the result is uninteresting because it does not prove that Blacks (and/or women) might not have protected themselves by behaving differently than White men. Thus, as a researcher you are damned if you do and damned if you do not. If you do not adequately assure uniform tester behavior, you will be criticized for not proving disparate treatment. If you do adequately assure uniform tester behavior, you will be criticized for not proving that trivial self-help could have mitigated the harms of the seller’s disparate treatment. By combining an analysis of both audit testing and consummated transactions, I hope to have at least partially responded to both of these criticisms. While it is still true that controlled testers who undertook slightly more aggressive search or bargaining strategies might have been able to mitigate the types of discrimination found, the empiricism put forward presents a strong prima facie case for the propositions that (1) a broad array of new car dealerships discriminate on the basis of race and (2) consumer self-help does not simply solve the problem.