DISCRIMINATION IN CONSUMMATED CAR PURCHASES

Ian Ayres[1]

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 in a variety of ways mitigate its impact. Richard Epstein, for example, in critiquing my previous audit studies of new car sales 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.[2]

Epstein is somewhat agnostic (it's "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 nondiscriminating, but James Heckman 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.[3]

---

[1]Townsend Professor, Yale Law School. ian.ayres@yale.edu. This paper updates Chapter 4 of IAN AYRES: PERVASIVE PREJUDICE? UNCONVENTIONAL EVIDENCE OF RACE AND GENDER DISCRIMINATION (2001).

[2]Richard A. Epstein, *Standing Firm, on Forbidden Grounds*, 31 SAN DIEGO L. REV. 1, 52-3 (1994).

[3]James J. Heckman, Detecting Discrimination, 12 J. Econ. Perspectives 101, 103 (1998). The importance of these kinds of general equilibrium concerns has long been recognized in the labor market

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.[4]

---

context, *See, e.g.,* Peter Siegelman, *Racial Discrimination in "Everyday" Commercial Transactions: What do We Know, What Do We Need to Know, and How Can We Find Out?*, in A NATIONAL REPORT CARD ON DISCRIMINATION IN AMERICA: THE ROLE OF TESTING 69 (Michael Fix and Margery Austin Turner, Eds. 1999); Christopher Flinn and James J. Heckman, *Are Unemployment and Out of the Labor Force Behaviorally Distinct States?* 1 J. LAB. ECON. 65 (1983); Harriet Orcutt Duleep and Nadja Zalokar, *The Measurement of Labor Market Discrimination When Minorities Respond to Discrimination*, *in* NEW APPROACHES TO ECONOMIC AND SOCIAL ANALYSES OF DISCRIMINATION (Richard Cornwall and Phanindra Wunnava, eds., 1991). John Yinger, *Cash in Your Face,* 42 J. URB. ECON. 339 (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.

[4]Epstein has also criticized the audit approach for the small proportion of observations in which dealers attempted to accept a tester offer:

> [I]n [Ayres'] 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.

Epstein, *supra* note 1, at 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 as in Chapter 3 of Pervasive Prejudice, Peter Siegelman and I also investigated whether our findings of race and gender discrimination might be linked to the fact that the dealership's 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, however, because acceptances only provide an

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.[5] An important finding of the audit regressions concerns the pervasiveness 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. At most, there were some statistically insignificant point estimates suggesting that black consumers might reduce the amount of race discrimination they encountered by avoiding suburban dealerships or dealerships located in neighborhoods with a low percentage of minority residents. The point estimates of the regression in Table 3.2 of *Pervasive Prejudice* indicated that the race discrimination over final offers was $264 worse in the suburbs. But this result was statistically insignificant.[6] And the same regression also suggested that suburban dealerships generally make final offers that are $221 lower than city dealerships.[7] So blacks could only protect themselves against suburban race discrimination by choosing to patronize urban dealerships that discriminate less but charge on average more. Similarly, the regressions suggested that final offers for black testers in all-white neighborhoods were $540 higher than

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.

[5]See Pervasive Prejudice, at 70-71.

[6]The t-statistic was only 1.22.

[7]The t-statistic was only -1.78.

in all-black neighborhoods. But again, this result was statistically insignificant[8] and combined with a point estimate that final offers at dealerships in all black neighborhoods were generally $118 dollars higher than all white neighborhoods. As Penelopi Goldberg 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.[9]

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.[10]

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 4 hours to negotiate a low markup, but blacks need to negotiate for 8 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

---

[8]The t-statistics were only 1.85 and 1.28.

[9]Pinelopi Goldberg, *Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey*, 104 J. POL. ECON. 622, 643 (1996).

[10]A possible exception to this might be so-called no-haggle dealerships which will be discussed below.

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.[11]

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' decision 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 proceeding evidence of disparate treatment in dealership offers,

[11]The social meaning of separate but equal regimes still can work substantial injury on traditionally subordinated people. See Jed Rubinfeld, *Textualism and Democratic Legitimacy: the Moment and the Millennium*, 66 Geo. Wash. L. Rev. 1085 (1998).

a finding that black purchasers pay more on average than white purchasers suggests that neither (1) purposive searching nor (2) alternative (and possibly more onerous) bargaining nor (3) dropping out of the market eliminates the effects of racially disparate dealership offers.

This paper analyzes five different data sets of consummated transactions. The first part of the paper responds to Pinelopi Goldberg's analysis of the Census Bureau's Consumer Expenditure Survey (CES) of approximately 1300 new car purchasers drawn randomly from the nation.. The second part analyzes analyses of approximately 900 new car purchasers from the Sutherlin Mazda new car dealership in Atlanta. The CES data has the advantage of being based on a national sample of purchasers, but, I will argue, suffers from severe errors in the measurement. The Atlanta data measures more directly the dealerships' underlying profits on financing and the vehicle itself, but is limited to a single dealership in a single city (and, as discussed below, may be subject to various sample selection problems -- because of its production as part of adversarial litigation).

The third part of the paper then analyzes two data sets analyzed in recent articles by David W. Harless and George E. Hoffer and by Fiona Scott Morton, Florian Zettlemeyer & Jorge Silva-Risso.[12] These two papers exploit newly available market research data from *J.D. Powers & Associates* and an

[12]David W. Harless & George E. Hoffer, *Do Women Pay More For New Vehicles?: Evidence from Transaction Price Data*, 92 Am. Econ. Rev. 270 (2002) [hereafter H&H]; Fiona Scott Morton, Florian Zettlemeyer & Jorge Silva-Risso, *Consumer Information and Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?*, Quantitative Marketing and Economics (forthcoming 2003) [hereafter SZS I]. See also Fiona Scott Morton, Florian Zettlemeyer & Jorge Silva-Risso, *The Effect of Information and Institution on Price Negotiations: Evidence from Matched Survey and Auto Transaction Data*, unpublished working paper 2002 [hereafter SZS II].

unnamed firm which provide the kind of direct evidence on vehicle profits that is found in the Atlanta data – but from a potentially less biased source and covering hundreds of different dealerships.

The paper'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 shows that whites pay lower average prices than blacks. Moreover, the size of the differentials are broadly similar. While the racial price differences are not statistically significant in Goldberg's analysis of the CES data, I will argue that this is a function of the noisiness of the data. In the less noisy audit studies and in the less noisy Atlanta purchases data, we observe similar coefficients and smaller standard errors -- which allow us to statistically identify the racial differentials as being statistically significant. The newer data from the market research firms continues to confirm the absence of a gender effect and the general presence of a statistically significant race effect. This paper will discuss the CES and Atlanta datasets in turn as well as the newer market research datasets and then conclude with a meta-analysis of the consummated purchases and audit data sets.

I. GOLDBERG'S ANALYSIS OF THE CES DATA

Spurred in part by the initial publication of the audit data , Pinelopi Koujianou Goldberg used data from the Consumer Expenditure Survey (CES) to undertake the first analysis of racially disparate treatment in consummated transactions. The CES dataset is created by the Bureau of Labor Statistics to compute the Consumer Price Index. Each quarter around 4500 households are interviewed and asked questions about their income and expenditures – including questions about their purchases of cars. For the years 1983 to 1987, Goldberg created a dataset of approximately 1300 reliable new car purchases. She constructed a measure of the discount from the dealer's sticker price and regressed this variable on a host of consumer characteristics including race and gender.

The race and gender coefficients for her core regressions indicate the following racial disparities:

TABLE 1: ESTIMATED PRICE PREMIUM OVER WHITE MALES BY PURCHASER DEMOGRAPHIC GROUP

Demographic Group	National CES Consummated Sales (Goldberg)
White Females	129 (117) [244]
"Minority" Females	426 (525) [28]
"Minority" Males	\$274 (263) [39]
Adjusted R-Squared	.14
Source: Goldberg, Table 2, Col. 1 Standard errors in parentheses Number of observation in brackets. A number of other right hand variables are not reported to save space.	

Goldberg interprets these coefficients as offering no evidence of racial price disparities:

> The results are quite different from the ones reported in Ayres [the pilot study] and Ayres and Siegelman [reported in Chapter 2 of Pervasive Prejudice]. Variables referring to race and sex have no explanatory power.[13]

[13]Goldberg, *supra* note 7, at 624.

Goldberg, however, does not interpret the results as casting doubt on the audit findings of disparate impact. Moreover, Goldberg expressly rejects the Epstein/Heckman hypothesis that purposive searching could eliminate the impact of disparate treatment by a even a large majority of dealerships. As noted above, Goldberg accepts the audit evidence that there were no "safe harbor" dealerships for blacks to patronize in order to avoid discrimination. Rather Goldberg attempts to reconcile the audit and the CES results by proposing a bargaining explanation based on racial differences in the variance of reservation prices:[14]

> [I]f the reservation price distribution for blacks is more spread out than the corresponding distribution for whites, then certain plausible parameter values generate a bargaining process with the following characteristics: (1) First-round offers are higher for blacks than for whites. (2) At the equilibrium, low-value blacks receive lower final offers than low-value whites. (3) High-value blacks receive higher final offers than high-value whites.[15]

Under this theory, dealers may still offer black consumers initially higher prices and still produce an equilibrium where the average prices paid by blacks and whites are the same. Goldberg supports her reservation price variance theory by showing that in the CES data minority purchasers have a statistically significant higher variance of sticker discounts than non-minorities.[16] She also uses 10% and 90% quantile regressions to show that minority purchasers have more dispersed prices than non-minorities. Under the reservation price variance theory, the dealers still engage in disparate racial treatment, but the race conscious selling strategy favors blacks with low reservation prices (relative to whites) while harming blacks

[14]A buyer's reservation price is the maximum amount she would be willing to pay a particular dealership for a particular type of car.

[15]Goldberg, *supra* note 7, at 645.

[16]The variance of minority discounts is 1.45 times larger than the variance of whites. *Id*. at 646.

with relatively high reservation prices.

While Goldberg's attempt to reconcile the two results is admirable, I believe a simpler and more persuasive explanation is that the noisiness of her data and its small sample size did not allow her to identify the statistical significance of positive increments by which minority prices exceeded white prices. Even though Goldberg claims that her results are "quite different" from both the pilot audit and the regressions reported in Pervasive Prejudice (Chapter 2), the size and sign of the coefficients are instead quite similar. Goldberg estimates that white women pay $129 more than white men, while in Pervasive Prejudice (Table 2.1) Siegelman and I estimated a $216 coefficient for the white female differential on final profits -- and the original American Economics Review article (which uses a slightly restricted sample of paired observations) by coincidence reports the exact same differential of $129.

Later in her article, Goldberg claims that "[o]nly the parameter estimates referring to white females seem to be in line with Ayres & Siegelman's findings,"[17] but the coefficients for black females are also of similar magnitude. Goldberg estimates the black female differential to be $426, while Pervasive Prejudice estimated the differential to be $465 (and the original AER estimate was $404). The only difference with regard to the minority female coefficients in these studies is not their sign or size; it is that merely Goldberg's coefficient is not statistically different than zero, while the audit coefficient is. But Goldberg is wrong to characterize this result as being very different from the ours. She certainly cannot reject the hypothesis that her estimate of $426 is the same as either of our audit estimates. Failing to reject that the differential is zero

[17]Goldberg, *supra* note 7, at 641.

does not imply that the true differential is more likely to be zero than the audit estimate. Indeed, her data suggests the opposite is true. Her data provides more evidence in favor of the hypothesis that black women pay $400 more than white males, than for the hypothesis that they pay the same amount.

The only estimated differential that is of a different magnitude concerns black males. Goldberg estimates a $274 differential (above white males), while Pervasive Prejudice 2 reported a $1133 differential (and the AER reported a $1061 differential). But her claim that her results are different than both the AER *and* the pilot study runs into a little trouble, because the pilot study reported a similar black male differential of $283. The difference in the pilot study and the full audit in estimating the black male differential is a troubling issue (and that will be discussed below in connection with the Atlanta results). But what is clear is that the point estimates for minority purchasers are consistently hundreds of dollars above the white males What is crucial for Goldberg's interpretation is that the minority male and minority female differentials are not statistically significant from zero. This is what drives her to characterize her results as "quite different" from the audit study.

But a close analysis of the quality of CES data indicates instead that Goldberg's estimates may not be statistically different than zero simply because her data is too imprecise to accurately measure the racial differentials. Orley Ashenfelter was the first to suggest this alternative interpretation. He concluded that:

> [Her article] basically has results similar to yours–except with much larger sampling errors, which is unsurprising given the noisy data set she uses. Early on I suggested that she send the paper to the AER as a confirmation of your results![18]

The remainder of this section takes a closer look at Goldberg's methodology and her results to argue that

[18]Letter from Ashenfelter to Ian Ayres (July 17, 1996).

my "noisy data" hypothesis is a more persuasive way to reconcile the CES and the audit data.

First, Goldberg tests suffer from aggregating different racial groups together. While the audit study tested for disparate treatment of blacks, the CES data only allowed Goldberg to test whether "minorities" paid higher prices. And for this purpose, a "Minority" was defined to be Black, American Indian, Aleut, Eskimo or Other. Whites, Asians and Pacific Islanders were grouped together as "White" and there is some uncertainty whether Hispanics were included in the minority definition as "Other" or as part of the "White" definition.[19] If different racial groups encounter different degrees of discrimination, then the aggregation of these groups by itself would increase the sampling error of estimating a single "minority" coefficient. Thus, from the get-go there is an important difference in comparing black (male and female) differentials with "minority" (male and female) differentials.

Second, there were very few minorities in the CES data set. While Goldberg's car purchases data set contained a total of 1279 observations, only 3% of the purchasers were minority males and only 2% of the purchasers were minority females. Moreover, it is unclear what fraction of these fractions were actual African Americans. In a footnote, Goldberg reports that: "In other regressions, the minority variable was split into the various race categories reported in the CES, with no change in the results."[20] But with so few observations on the individual races, it hardly surprising that Goldberg was unable to produce statistically significant results. Indeed, the aggregation of multiple races combined with the small number

[19]My information on the racial composition of the CES data was generously provided by Goldberg.

[20]Goldberg, *supra* note 7, at 637.

of minority observations by itself could easily explain Goldberg's inability to produce statistically significant minority coefficients.

Third, Goldberg's data is based on "head of household" surveys that "were collected as long as 3 months after the actual transaction took place, and are subject to errors of recall or memory that add noise to the price and other variables of interest."[21] Although the CES are *derived* from actual transactions, they are not true "transactions prices," but are based on household heads' recollections of what was paid for the car. But the survey respondent may not know, may not remember or may intentionally coverup having paid too much. And as Goldberg acknowledges, her approach assumes that the race and gender of the household head is the same as the race and gender of the person(s) who bargained for the car itself.

Fourth, Goldberg uses a necessarily imprecise measure of dealership profits. Unlike the audit data which used markup estimates (comparing actual dealership offers written contemporaneously by testers with detailed estimates of dealership cost), Goldberg estimates the discount from the sticker price. This is a Herculean task that requires the manipulation of 10 underlying estimates:

Discount = [Base Model List Price + Option List Price + Destination Fees + Dealer Prep Fees + Other Dealer Specific Costs] - [(Down Payment + Principle Borrowed - Extra Charges)/Sales tax rate + Trade-in Allowance]

The first bracketed term on the right-hand side is Goldberg's estimate of the dealer's sticker price and it is plagued by errors in identifying the base model and particular car options (which are often

[21]Peter Siegelman, supra note 2.

unobservable).[22] The second bracketed term is Goldberg's convoluted estimate of the net transaction price which if anything is even noisier than the estimate of dealer sticker price. Goldberg implicitly assumes that both financing and trade-ins were at cost. This is important because 83% of the CES car purchasers financed their cars through the dealership, and 50% traded-in a used car in conjunction with the new car purchase. If the dealers systematically packed transaction profits into minority financing or trade-ins, then Goldberg's results could understate the amount of discrimination. Or more neutrally, if the dealerships sometimes shifted a constant profit into the finance or the trade-in component of the deal, then we might see larger standard errors in the CES than were seen in the audit study (where trade-ins and dealer financing were expressly taken off the table by the tester script).[23]

An analysis of regression results also suggests the noisiness of the data. To begin, Goldberg reports an adjusted R-squared of only 14%, while the audit studies can explain more than twice the variance in the regressor (with adjusted R-squared's ranging between 33% and 44%).[24] And Goldberg's finding that first time buyers receive systematically "better deals" is suspect. Anecdotal evidence and theory strongly

---

[22]Goldberg explains:

> [M]y approach treats unobserved options as part of the error term. To the extent that the purchase of these additional features is uncorrelated with the explanatory variables on the right-hand side, the estimation results are consistent.

Goldberg, supra note 7, at 630. But, of course, if blacks tend to buy no frills cars with fewer unreported options than Goldberg will underestimate black list prices and consequently the amounts of discrimination.

[23]Because the CES data does not disclose the purchaser's residence, Goldberg was also forced to use a tax rate which was the probabilistic (weighted average) tax rate for the region.

[24]It should also be noted that Goldberg's graphs indicate a substantial proportion of her observations having negative discounts – indicating that the car was sold above its original sticker price. This suggests some errors in measurement as well.

suggest that such buyers are easy marks for salespeople. Indeed, analysis of the pilot study suggests that testers who revealed (under direct questioning) that they did not own a car were ask to pay $337 more,[25] but Goldberg estimates that such first time buyers were asked to pay $444 less (and the result was statistically significant). It is also surprising that in such a large data set that none of the financial or demographic variables (value of financial assets, after tax income, college education) were statistically significant. A small part of this may also be due to the author's decision to report White heteroskedastic-consistent errors which are systematically larger than normal OLS estimates and thus tend to reduce the tests of significance.

Goldberg's own "reservation price variance" theory emphasizes that the regression residuals for minority purchasers are larger than for white purchasers. She says that this great variance can explain both why the audits show initial higher offers to minorities and why minorities may not pay more on average. But reservation cost variance theory does not do a good job explaining why discrimination against minorities persisted throughout the bargaining process in the audit tests. The audit data uncovered not just initial higher offers but higher final offers. The tested dealerships -- contrary to the reservation price variance theory -- were not willing under either the split the difference or the fixed concession strategies to let the black testers bargain down to the lower price. And contrary to the prediction of the reservation price variance theory, the dealers were more likely to end the negotiation by refusing to bargain further when bargaining with black testers. But if the dealers believed blacks to have more variable reservation

[25] See Ian Ayres, *Fair Driving: Gender and Race Discrimination in Retail Car Negotiations*, 104 HARV. L. REV. 817, 848 (1991). The result was statistically significant at the 8% level.

prices, we would expect dealerships to be willing to bargain longer with blacks to allow low-valuing blacks to credibly signal their reservation price.[26]

A more convincing interpretation is simply that the differences in the statistical significance of the CES and auditing data arises because we are both measuring the same underlying parameter(s), but that Goldberg is doing so with data that are substantially noisier than the audit studies. Goldberg does acknowledge this possibility:

> The possibility that measurement error is responsible for the standard errors of race and gender parameters cannot be eliminated, yet it is hard to explain why measurement error would make all socioeconomic characteristics insignificant while most of the parameter estimates referring to market, transactions or model specific variables have the expected signs and are highly significant.[27]

As noted above, first buyer dummy does not have the expected sign; the insignificance of the financial variables and the low R-squared are all indicia of measurement errors. But more importantly the imprecision in the very aggregated definition of "minority" when combined with the paucity of observations combine to make the measurement error hypothesis especially compelling.

Still the difference in magnitude of the black male coefficient is troubling. Fortunately, I have gained access to another data source on transaction prices from an individual dealership in Atlanta that allows an alternative look into the actual workings of the retail car market. It is to this data that I now turn.

II. ATLANTA DEALERSHIP TRANSACTION PRICES

[26] It should be also noted that while Goldberg claims her theory works for "plausible" parameter values, she does not fit her data to a model (as was done in Pervasive Prejudice (Chapter 3)).

[27] Goldberg, *supra* note 7, at 642-3.

Just as Goldberg's article in part was a response to my previously published American Economics Review article, the data analyzed in this section was created in part because of a previous publication in Harvard Law Review concerning the original pilot study. An Atlanta attorney, Edward D. Buckley III, in the course of interviewing former car salespeople of Sutherlin Mazda in Conyers, Georgia (a suburb of Atlanta), was told that the dealership was pressuring the salespeople to discriminate against African American customers. Buckley remembered reading popular press accounts of the original Harvard article and went out and read the article itself. Buckley ended up filing a class action law suit on behalf of black customers against the dealership and as part of his investigation in the suit Buckley collected sales and commission receipts from the dealerships' salespeople on over 800 car sales.[28] This transaction data forms the core of this section's analysis.[29]

The sales and commission data has several advantages over the CES data. It expressly states not only the total vehicle profit, but breaks down this profit stemming from dealership financing and the profit coming from the vehicle sale itself. It expressly states the transaction price for the vehicle sale as well -- so it is possible to deduce the dealership's vehicle cost.[30] The data also indicates whether there was a trade-in but (as with the CES data) not the dealership's profit (or loss) on the trade-in. The Atlanta data also has a much higher percentage of black and/or female purchasers: 49% of the purchasers are black

[28] See Tim O'Reiley, *Race-Based Car Pricing Charged: Buyers Attempt Class Action Against Sutherlin*, FULTON COUNTY DAILY REP., May 8, 1996, at 1.

[29] The legal implications of the suit are discussed in Pervasive Prejudice.

[30] Dealership cost = Sale price - vehicle profit.

and 46% of the purchasers are female – so there is much more information about the experience of black purchasers in this sample.

And while the data set has information on sales for a number of years (from 1990 to 1995), it is information on cars sales at only a single dealership in a single city. Moreover, the data was produced by potentially biased salespeople (who were also pursuing a lawsuit against the dealership). The salespeople might have censored the data to produce more discriminatory results. More importantly, salespeople provided the race of the consumers based on their memory. This memory might be faulty or intentionally biased (so as to report all high markup sales as "black").[31] To explore this worrisome possibility, I informally audited the raw data comparing the zip code of the purchaser (which was indicated on the original sale receipt) to the salesperson's recollection of race to see if black and white consumers tended to live in predominantly black and white neighborhoods. While my investigation did not rise to the level of sophisticated geocoding that has been undertaken in other contexts,[32] I did not uncover any worrisome tendencies. It may also be worth noting that the purchasers who were reported as being "black" often had first and last names that are particularly common among African-Americans (in comparison to the "white"

[31]I coded gender based on purchasers first names. However, because some names (e.g., "Pat") do not reliably indicate sex, there were some missing purchaser gender values for 66 out of 898 transactions. In what follows, I have dropped these missing values from the analysis. But I can also report that none of the results change if the additional observations (and a dummy for missing-gender) is included in the regressions.

[32]See, e.g., Commission for Racial Justice, United Church of Christ, *Toxic Wastes And Race in the United States: a National Report on the Racial and Socio-economic Characteristics of Communities with Hazardous Waste Sites* 2-3 (1987); Richard Seltzer, John M. Copacino & Diana Roberto Donahoe, *Fair Cross-section Challenges in Maryland: an Analysis and Proposal*, 25 U. Balt. L. Rev. 127 (1996).

names which more often signaled European ethnic origin).[33] In sum, while there are some dimensions on which the Atlanta data is more limited and possibly more infected by adversarial bias, on net it seems to provide at least as informative data on this issue of disparate racial pricing and probably much more powerful data.

Table 2 reports the results of the simplest regressions testing for racial and gender disparities:

Table 2: Profit and Markup Regressions Detailing Basic Racial/Gender Disparities

Purchaser Type	Vehicle Profit	Finance Profit	Total Profit	Vehicle Markup	Finance Markup	Total Markup
Constant	229.16*** (45.84)	504.69*** (48.95)	526.25*** (64.43)	.0258*** (.0064)	.0431*** (.0052)	.0510*** (.0079)
White Female	-22.94 (74.14)	57.70 (79.32)	9.17 (104.21)	-.0029 (.0103)	.0088 (.0084)	.0027 (.0128)
Black Male	405.04*** (72.32)	470.79*** (72.16)	836.81*** (101.65)	.0534*** (.0100)	.0439*** (.0077)	.0931*** (.0124)
Black Female	504.99*** (67.73)	589.01*** (68.18)	1018.40*** (95.21)	.0614*** (.0094)	.0566*** (.0072)	.1098*** (.0116)
N	831	551	831	822	544	822
R-squared	.09	.15	.17	.08	.125	.138

The regressions show a consistent pattern that white females receive approximately the same deals as white

[33]See, e.g., Richard Delgado, *The Colonial Scholar: Do Outsider Authors Replicate The Citation Practices of The Insiders, but in Reverse?*, 71 Chi.-Kent. L. Rev. 969 (1996) (relying in part on racial inferences of names).

males while black male and black female testers pay substantially higher profits. The total profits of white female purchasers is only $9 more than that of white males – but unsurprisingly this result is not statistically significant from zero. But total profit paid by black males is $837 more, and the total profit paid by black females is $1018 more than that of white males. And these results are highly statistically significant (far beyond the 1% level). The total markup up regressions correspondingly show that black male markups are 9.3% higher and black female markups are 11.0% higher than the markups for white male purchasers.

Figure 1: Percentage of Sales by Total Gross and Purchaser Type

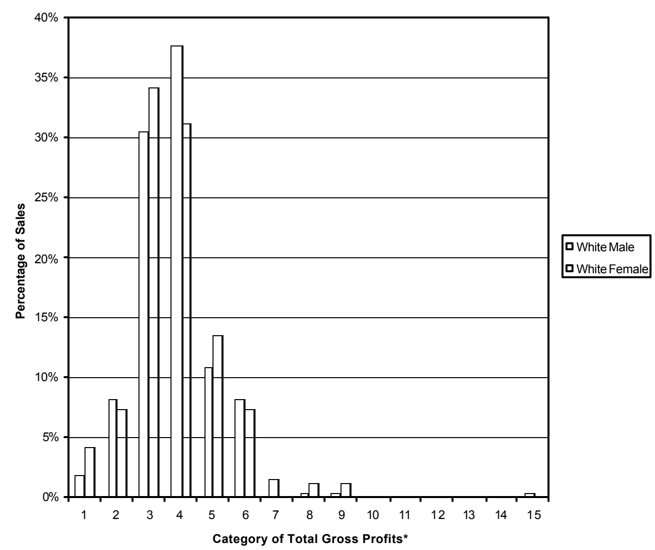
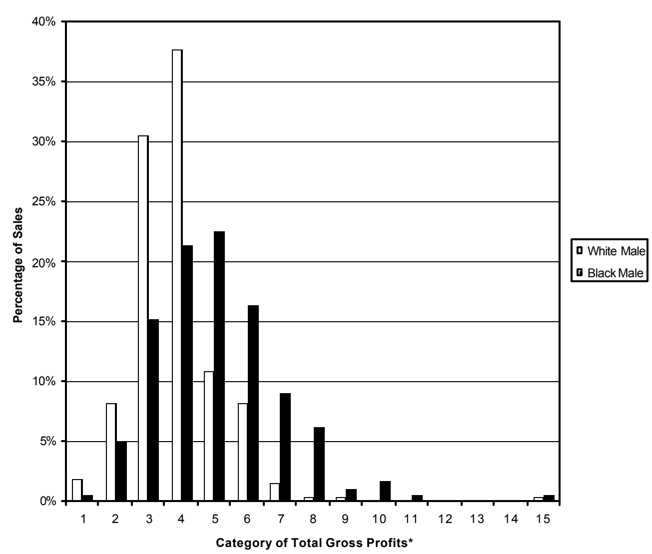
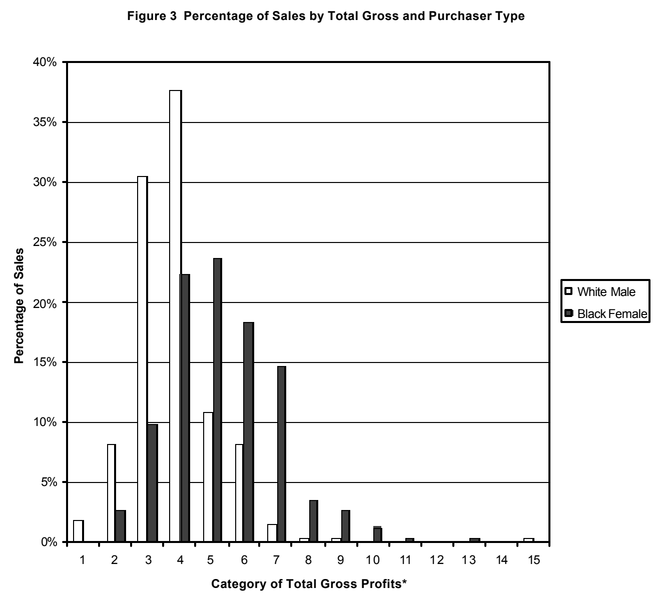


Figure 2: Percentage of Sales by Total Gross and Purchaser Type



The table also shows that the disparate total profits stem from disparities both in vehicle and financing profits. For example, the dealership vehicle profits from Black females are $505 higher than their profits from white male purchasers; and the dealerships financing profits (from those 500 some-odd people who financed) are $589 higher for black female relative to white male purchasers. Again these results are highly significant.[34]



Black purchasers were much more likely to pay high profits. Only 2.2% of white male purchasers were part of the 10% highest profits, while 14.6% of black males and 20% of black female purchasers were represented in this category of most profitable sales. This means that black males were 6.5 times more likely than white males to pay the highest decile of profits. And black females were 8.9 times more likely than whites to pay the highest decile of profits.

The different distributions of profits for the different purchaser types are graphically displayed in histogram bar charts of Figures 1, 2 and 3. Figure 1 dramatically illustrates the similarity of the white male and white female profit "bell curves," while Figures 2 and 3 show a dramatic difference in between the white male and both the black male and black female profit distributions.

The data also supports an "anecdotal" assertion of the initial Harvard article that "at some dealerships up to fifty percent of the profits can be earned on just ten percent of the sales." In this data set, 50% of the dealerships profits come from 17.2% of its most profitable sales. And the profits stemming from white males and white female purchasers are even more concentrated with 12.5% of the most profitable white male purchasers generating 50% of the white male profits and with just 11.6% of the most

[34]Note the total profit coefficients are not merely the sum of the vehicle and financing profit coefficients, because not all purchasers use dealership financing.

profitable white female purchasers generating 50% of the white female profits. Black profits are more uniformly elevated and therefore relatively less concentrated with 21.4% and 25.0% of black male and black female purchasers respectively generating 50% of the profits coming from that tester type.

But the analysis up until now has only explored racial/gender differences in pricing. It is also possible to assess whether other aspects of the transaction are "gendered" or "raced." To this end, I ran logit regressions testing whether the race and gender characteristics of the purchaser were correlated with various aspects of the transaction. The regressions indicate that black men and women were respectively 55% and 37% more likely (than white males) to buy a used car and that black men and black females were respectively 73% and 70% more likely (than white males) to finance their purchases at the dealership. And these differences were statistically significant at (at least) the 5% level, while the white female differential in both comparisons was small and statistically insignificant. There were no race/gender differences in purchasers' propensity to trade-in a car as part of their purchase transaction. And controlling for the used/new status of the purchased car, there were no race or gender differences in the dealership's cost of the cars purchased.[35] So there is no evidence that different types of purchasers were steered to particularly high- or low-cost models.

There was, however, interesting evidence of steering purchaser types to particular sales people. A Chi-squared test with 9 degrees of freedom equaling 42.4756 strongly rejects the hypothesis that the race and gender of purchasers is independent of the salesperson's identity. To further investigate a possible

[35] Dealer's cost was calculated as "sales price" minus "vehicle profit."

motive for such steering I reran the total profit regression of Table 2 adding a "Specialization Percentage" variable which equals the percentage of a particular purchaser type encountered by the particular salesperson. For example, if a black female purchased from the second salesperson (referred to in later regressions as "S2") then the "Percentage" variable would take on a value of 41% because 41% of this salesperson's customers were black females. I also created four additional percentage variables – by interacting the purchaser type dummies with the "Percentage" variable. The results of these regressions are reported in Table 3:

Table 3: Test of whether salespeople who specialize in negotiating with a particular type of purchaser garner higher profits		
Regressands	Total Profit	Total Profit
Constant	-4.34 (220.05)	372.91 (431.63)
White Female	224.12* (134.39)	-319.33 (779.42)
Black Male	1094.26*** (99.55)	558.05 (497.74)
Black Female	1016.34*** (123.85)	758.48 (573.89)
Specialization Percentage	1571.83** (623.50)	–
White Male x Specialization Percentage	--	454.26 (1264.43)
White Female x Specialization Percentage	--	2399.50 (3206.11)
Black Male x Specialization Percentage	--	1037.32 (1656.52)
Black Female x Specialization Percentage	--	2121.29*** (821.95)
N	831	831
R-squared	.17	.17

These regressions suggest that purchasers dealing with a salesperson who specialized in your type of customers produced systematically higher profits for the dealership. The $1572 coefficients on

"Specialization Percentage" in the first regression indicates that buying from a salesperson who makes 10% more of his/her sales to purchasers of your race and gender (than the average salesperson) is likely to pay a $157.2 higher profit.[36] The regression reported in the right-hand column breaks down the result by specific purchaser type. There we see that the purchaser-specific "Specialization Percentage" coefficients are all estimated to be positive, but that only the steering coefficient for black females is statistically significant.[37]

To get an even more particularized view of the determinants of dealership profit, I ran a series of nested regressions adding a series of additional variables to the purchaser type characteristics. In the simpler regressions (reported in columns 1, 3, and 5 of Table 2), I added variables concerning core transaction characteristics. To test whether dealership profits were related to the dealership's underlying cost (in a potentially non-linear way), I included both the dealership's cost and the dealership's cost squared.[38] I also included dummy variables indicating whether the consumer was buying a used (=1) or new (=0) car; whether the transaction included a trade-in sale (=1) or not (=0); and whether the purchaser used dealer financing (=1) or not (=0).[39]

[36]Because all of the salespeople were African-American men, it was impossible to test, as in the original pilot study, whether purchasers are steered to salespeople of their own race and gender who consequently give them systematically worse deals. See Ayres, *supra* note 21, at 833.

[37]The steering result is pronounced and correlates sufficiently with the purchasers race and gender to render these purchaser type coefficients statistically insignificant.

[38]If dealerships charged a fixed markup that was less than 50%, then we would expect the coefficient on the squared term to be zero and the coefficient on the linear term to be less than 1.

[39]The finance dummy was dropped from the regressions using finance profits as the left-hand side variable (reported in columns 3 and 4) – because these indicators are perfectly correlated with the constant term.

In the fuller regressions (reported in columns 2,4 and 6 of Table 4), I tested whether the dealership earned different profits on used cars produced by different manufacturers; whether the dealership made different profit on different makes of its new cars; and whether specific salespeople tended to earn different profits. I also included a number of time related dummies to control for day, week, week of month and year dummy variables. Following Goldberg, I also included a number of variables interacting the quarter a new car was sold with a dummy variable indicating the car's relative model year. The idea being -- for example -- that new cars of next model year sold in the 4th quarter might sell with a higher profit rate than cars of the same or previous model year that are sold in the 4th quarter.

Table 4: Determinants of Dealer Profits						
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variables	Vehicle Profit	Vehicle Profit	Finance Profit	Finance Profit	Total Profit	Total Profit
Constant	-188.73 (142.88)	-55.95 (327.44)	-13.98 (177.64)	640.60* (361.84)	-636.53*** (183.86)	25.32 (427.78)
Purchaser Type						
White Female	-12.21 (72.92)	-41.56 (74.35)	48.20 (79.81)	92.16 (83.59)	7.58 (93.83)	-11.67 (97.13)
Black Male	324.76*** (71.61)	459.95*** (70.58)	452.80*** (73.33)	637.05*** (74.46)	632.35*** (92.15)	865.47*** (92.20)
Black Female	450.26*** (67.21)	328.98*** (73.69)	590.48*** (69.09)	433.95*** (77.08)	850.24*** (86.49)	611.43*** (96.27)
Transaction Characteristics						
Dealer Cost	.0082 (.01584)	.0166 (.0178)	.0532*** (.0195)	.0533** (.0228)	.0397* (.0204)	.0460** (.0233)
Dealer Cost Squared	-2.26e-08 (4.34e-07)	-3.80e-07 (5.12e-07)	-1.21e-06** (5.14e-07)	-1.20e-06* (6.21e-07)	-7.16e-07 (5.59e-07)	-1.06e-06 (6.68e-07)
Trade-In (= 1)	140.59** (55.12)	148.30** (57.84)	37.55 (55.82)	17.49 (59.55)	170.58** (70.94)	166.56** (75.57)
Used (=1)	416.57*** (73.88)	516.07*** (106.70)	185.73** (74.77)	189.01* (114.14)	555.93*** (95.06)	660.83*** (139.41)
Dealer Financing (=1)	229.22*** (55.69)	192.16*** (58.19)	--	--	959.13*** (71.66)	936.82*** (76.02)
Manufacturer Dummies (Mazda Excluded)						
Other Japanese Manuf. (=1)	--	-86.23 (167.13)	--	-222.61 (183.11)	--	-234.80 (218.36)
Big 3 Manuf. (=1)	_	-209.26 (169.61)	--	-4.78 (191.86)	--	-265.03 (221.60)

Other Make (=1)	–	166.52 (329.45)	--	512.99 (350.34)	--	430.60 (430.42)
Model (Miscellaneous Models Excluded)						
Mazda 626 (=1)	--	-190.93 (157.36)	--	-109.65 (177.15)	--	-257.22 (205.59)
Mazda 929 (=1)	–	107.10 (199.86)	--	155.87 (213.89)	--	199.06 (261.12)
Mazda BXX00 (=1)	--	-55.95 (179.95)	--	1.01 (197.30)	--	-110.42 (235.10)
Mazda MPV (=1)	--	-47.33 (184.01)	--	9.44 (204.50)	--	-42.10 (240.41)
Mazda MX3 (=1)	--	-27.33 (193.78)	--	-105.51 (203.40)	--	-120.91 (253.18)
Mazda MX6 (=1)	--	-154.46 (200.91)	--	-151.051 (221.18)	--	-204.03 (262.49)
Mazda Miata (=1)	–	-15.98 (184.42)	--	-18.67 (197.62)	--	-49.09 (240.95)
Mazda Protege (=1)	--	-139.30 (163.56)	--	54.06 (180.50)	--	-114.05 (213.69)
Salesperson Dummies (Salesperson 1 Excluded)						
Salesperson2 (=1)	--	26.61 (111.18)	--	-1.22 (106.00)	--	-3.14 (145.26)
Salesperson 3 (=1)	--	-111.18 (76.74)	--	74.84 (78.37)	--	-80.94 (100.25)
Salesperson 4 (=1)	--	65.68 (98.25)	--	177.70* (101.29)	--	191.04 (128.36)
Salesperson 5 (=1)	--	-72.47 (75.39)	--	21.75 (78.30)	--	-48.63 (98.50)
Year Dummies (1990 excluded)						

1991 (=1)	--	386.39 (355.15)	--	-918.69* (496.515)	--	-125.21 (463.99)
1992 (=1)	--	158.54 (253.99)	--	-593.67** (256.73)	--	-239.36 (331.83)
1993 (=1)	--	212.70 (236.01)	--	-627.55*** (238.52)	--	-256.99 (308.34)
1994 (=1)	--	-118.00 (236.89)	--	-556.72** (9.53)	--	-531.42* (309.49)
1995 (=1)	--	-210.60 (258.01)	--	-345.25 (274.07)	--	-576.63* (337.08)
(Calendar Year) Quarter Dummies Interacted With Relative Model Year (1st Quart. x Same Model Year Excluded)						
1st Quart. x Prev. Model Year	--	16.517 (124.83)	--	-59.7282 (129.76)	--	-10.38 (163.09)
2nd Quart. x Same Model Year	--	51.15 (89.84)	--	-45.92 (93.10)	--	8.84 (117.38)
2nd Quart. x Prev. Model Year	--	137.49 (210.28)	--	208.01 (244.33)	--	196.34 (274.72)
3rd Quart. x Same Model Year	--	-101.29 (91.24)	--	14.78 (90.96)	--	-116.46 (119.21)
3rd Quart. x Prev. Model Year	--	-48.04 (199.98)	--	180.63 (202.06)	--	37.71 (261.27)
4th Quart. x Same Model Year	--	57.00 (103.81)	--	33.83 (110.72)	--	79.75 (135.62)
4th Quart. x Prev. Model Year	--	-119.20 (341.37)	--	208.94 (438.00)	--	16.98 (445.99)
4th Quart. x Next Model Year	--	121.74 (127.98)	--	169.99 (138.43)	--	220.90 (167.20)
Day of Week Dummies						

Monday	--	-112.28 (92.19)	--	-97.27 (92.26)	--	-187.62 (120.45)
Tuesday	_	70.71 (97.82)	--	-215.37** (104.31)	--	-81.82 (127.80)
Wednesday	_	-39.96 (95.32)	--	-67.17 (98.46)	--	-115.69 (124.54)
Thursday	--	-38.98 (105.82)	--	-174.07 (111.64)	--	-129.79 (138.25)
Friday	--	-43.95 (103.01)	--	-166.61 (110.92)	--	-148.69 (134.59)
Saturday	--	-50.28 (79.56)	--	-195.21** (82.60)	--	-168.92* (103.94)
Week of Month Dummies						
Week 2	--	-105.50 (96.38)	--	38.27 (99.39)	--	-52.50 (125.92)
Week 3	--	-34.00 (90.89)	--	87.07 (94.30)	--	19.34 (118.75)
Week 4	--	-21.15 (92.00)	--	-25.10 (94.83)	--	-32.45 (120.19)
Week5	--	-24.89 (97.78)	--	-47.02 (101.92)	--	-29.39 (127.75)
Week 6	_	-380.46* (205.31)	--	-232.26 (221.76)	--	-506.06* (268.24)
N	818	816	542	541	818	816
Adj. R-squared	.1525	.1927	.1629	.1900	0.3484	.3598

Notes: Standard Errors in parentheses
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

The regressions show that even after controlling for other attributes of the transaction there is a robust tendency for African-American customers to pay significantly more than white customers. The coefficients on the black male and black female variables are positive (ranging from $325 to $865) and statistically significant at the 1% level in all regressions. The racial disparities are quite robust, but note that the ordering is not. In the fuller regression the black male premium (over white males) exceeds the black female premium, but in the shorter regression (without time and manufacturer controls) this result was reversed. The profits of white female purchasers continue to be statistically indistinguishable from white males. The primary picture from this data set seems to be that race and not gender is the important determinant of dealership profit. This relative unimportance of gender might, however, be caused by purchasers bargaining as part of (single-race) heterosexual couples.[40]

The dealership's cost of the vehicle does not have a significant impact on the size of the dealership profit. The linear term suggests that adding an extra thousand dollars to the dealer's cost of a car only increases its expected profit by $8.2 to $16.6 (in columns 1 and 2). However, the profits from dealership financing are positively related to the dealership's cost of car. A car costing the dealership an extra $1000 leads to more than an extra $50 profit. Although we do not observe the amount of principle borrowed, it may be fair to infer that the larger the principle the larger the expected dealership profit. The squared term on dealership costs in the finance profit regressions is statistically significant and negative indicating that the

[40]Gender does, however, have some statistical influence. In each of the 6 regressions, a test of the hypothesis that the black female coefficient and black male coefficient are equal was strongly rejected. But as noted above the ordering the gender effect is not stable.

effect of cost on profits diminishes as the cost increases. But the effect remains positive for the relevant range of costs in our data set.[41] The final regressions (reported in columns 5 and 6 of Table 4) show that dealership cost has a positive influence on dealership profit but this effect is only statistically significant at the 10% level and is probably caused by the aforementioned financing effect.

The dealership financing variable is of particular interest. One might expect vehicle profits to be lower on sales that the dealership financed. As Goldberg quite reasonably hypothesized: "One would expect dealers to offer special [vehicle] discounts [on deals which they also financed] since they make additional profits through the higher interest rates charged or the cuts they receive as agents for the loan providers. . . . [T]he dealers have incentives to lower the nominal purchase price of the car in exchange for a higher future return on the loan."[42] However, columns 1 and 2 of Table 4 indicate that vehicle profits are systematically higher on transactions in which dealers also provide financing.[43] In retrospect, it seems that customers who finance their car purchases are also more willing to pay higher prices. Signaling a willingness to finance at the dealership may also signal a willingness to pay more for the car itself. But as noted above this result may be related to the customer's race, as black purchasers systematically were more likely to finance.

A similar effect appears with regard to trade-ins and the sale of used cars. Consumers who trade-in a car pay systematically higher vehicle profits and higher total profits. Columns 1 and 2 (of Table 4) suggest

[41]The costs are positively related to profit for all costs below $86,000.

[42]Goldberg, supra note 7, at 634, 638.

[43]This is different than Goldberg's finding of larger discounts on transactions with dealership financing.

that vehicle profits are between $140 and $150 higher when there is a trade-in; and Columns 5 and 6 indicate that total profits are between $166 and $170 higher. This might be caused by dealership's giving above-market trade-in allowances in exchange for higher profits on the car sale itself. Or it might be that consumers who trade in their car have a higher willingness to pay. Since we cannot observe the trade-in allowance or the profit on the trade-in transactions, we cannot distinguish between these hypotheses.

The dealership profits on the sale of used cars were systematically higher than those on new cars and surprisingly large -- ranging between $556 and $661 more in columns 5 and 6. Moreover, this result is at least marginally significant in all 6 regressions suggesting that dealerships earn both higher vehicle profits and higher financing profits when a used car is involved. The financing result might in part be caused by a higher risk of lending to used car purchasers -- so that the dealership demands a higher risk-adjusted return. The racial disparities for black male and black female purchasers (and their statistical significance) are robust to dropping all the observations in which a trade-in occurred or in which used cars were sold. In other words, even if we restrict our attention only to purchases without trade-in or only to purchases involving new cars, we still find that blacks paid significantly higher profits.

Neither the manufacturer (make) nor the model dummies were statistically different than zero in any of the regressions (reported in columns 2, 4 and 6). However, the three manufacturer coefficients in the finance regression (column 4) were jointly different than zero at 5% level of significance – possibly indicating that the dealership disparately matched financing incentives of different manufacturers to remain competitive.[44]

[44]Other manufacturers, however, predominantly offer finance incentives on their new cars and the non-Mazda observations were predominantly for used cars.

The salesperson indicator variables were not jointly different from each other in any of the three regressions. However, the total profits of Salesperson 4 were significantly higher than those of Salespersons 3 or 5 (respectively, at the 5 and 10% level of significance). Transactions which the dealership financed produced statistically higher finance profits for the dealership when Salesperson 4 was involved relative to Salesperson 1 or Salesperson 2. This financing result is somewhat surprising since the financing terms are generally not negotiated by the salespeople but by specialized "F & I" (finance and insurance) representatives in the "back office" of the dealership.[45]

The year dummies suggest that financing profits were significantly higher in 1990 than in later years but that financial profits were generally increasing over time after 1991.[46] Total profits were decreasing year by year throughout the data – with total profits from observations in the final two years (1994 and 1995) being significantly lower than profits associated with 1990, This suggests a general increase in competition throughout the period under analysis.

An analysis of the calendar quarters indicates, as in Goldberg's analysis, that the amount of the total profits that dealerships earn did not vary significantly by quarter or by relative model year. For some quarters, the profits on the "same year models" were higher than profits on the "previous years model" (as in the 1st and 4th quarters), but in other quarters the result was reversed. In the fourth quarter, as predicted,

[45]Still, ex post, one could hypothesize a number of scenarios that could explain the result. Salesperson 4 might be steering particular types of customers who were likely to pay higher profits when financing (i.e., customers of used cars). Or something about the way that Salesperson 4 negotiated for the underlying vehicle might have softened the purchasers for their subsequent financing negotiation.

[46]The year dummies were jointly different from each other in all three regressions at the 5% level.

I found that total profits on the next model year were higher than on the same or previous model year, but this difference was not statistically significant. And in each of the regressions, I could not reject the hypothesis that the coefficients were jointly equal to zero.

The day of the week dummies suggest that cars purchased on Sunday had higher total profits than other days of the week (because the incremental coefficients are negative for all other days) and that this difference was increasing from Tuesday onward through Saturday, which had estimated profits that were $169 less than Sunday. The Saturday-Sunday increment was the only pairwise difference that was statistically significant.[47] It may be that dealerships concede more when they believe that customers have greater opportunity to search. When negotiating with a potential consumer on a Friday or Saturday, both the customer and the dealer may realize that the consumer can easily look elsewhere the next day if they do not reach a deal – whereas during Sunday negotiations the customer and dealer may both realize that it will be difficult for the consumer to take time off during the week or negotiate after a long work day.[48]

Finally, I tested whether the dealership was likely to earn different profits during different weeks of the month. The week of month variable was constructed on a Sunday to Sunday basis -- so that a month beginning on a Saturday could have six weeks (4 full weeks and a partial week at beginning and end). Interestingly the regression suggests that the dealership charged lower prices in the 6th week -- more than $506 lower total profit than in the excluded 1st week of the month. This is consistent with anecdotes that

[47]A joint test failed to reject that the hypothesis that day of the week effects were equal to zero.

[48]The fourth column unexpectedly finds that financing deals were less profitable on Tuesday. This is unexpected, and even ex post I cannot think of a cogent explanation.

manufacturer incentives sometimes motivate dealers at the end of the month to increase volume to make a particular quota. But this result was only marginally significant (at the 10% level) and was not robust to alternative specifications.

While this extended discussion of the right-hand side variables is interesting in and of itself, for this paper, the regressions underscore that -- controlling for a host of transaction characteristics -- the finding that blacks pay more remains a robust and statistically significant result. While the regressions in Table 4 are much more controlled than the previous regression comparing the profits of different purchaser types (reported in Table 2), these regressions still cannot be interpreted by themselves as evidence of disparate treatment. It is still very possible that black and white consumers behaved differently. But a major the goal of the exercise is to show -- counter to Heckman's and Epstein's arguments -- that disparate behavior by black consumers does not mitigate the harms of disparate treatment uncovered in the audit studies.

Finally, to get a further purchase on the causes of the disparate dealership profits for blacks and whites, I reran the regression from column 6 of Table 4 adding a number of variables interacting a dummy equal to one if the purchaser was African American with the previously included transaction characteristics (such as dealership cost, financing, trade-in, and used) as well as interaction with the salesperson dummies. The results are reported in Table 5:

Table 5: Regression of Total Profit on Previous Explanatory Variables and Black Interaction Effects

Explanatory Variables	Results	Additional Explanatory Variables	Results
Constant	249.12 (461.30)		
White Female	-1.77 (96.32)		
Black Male	347.95 (399.05)		
Black Female	110.00 (399.98)		
Dealer Cost	.0345 (.0296)	Black x Dealer Cost	.0087 (.0426)
Dealer Cost Squared	-1.01e-06 (8.04e-07)	Black x Dealer Cost Squared	22.88e-07 (1.18e-06)
Trade-In (= 1)	191.05* (105.65)	Black x Trade-In (= 1)	-74.72 (150.14)
Used (=1)	652.89*** (164.52)	Black x Used (=1)	-96.72 (197.19)
Dealer Financing (=1)	767.18*** (98.64)	Black x Dealer Financing (=1)	456.85*** (149.34)
Salesperson 2 (=1)	-469.50** (206.21)	Black x Salesperson 2 (=1)	899.53*** (286.74)
Salesperson 3 (=1)	-158.03 (131.47)	Black x Salesperson 3 (=1)	207.75 (194.81)
Salesperson 4 (=1)	150.80 (216.39)	Black x Salesperson 4 (=1)	42.32 (263.13)
Salesperson 5 (=1)	-9.56 (128.95)	Black x Salesperson 5 (=1)	-125.07 (193.64)
N		816	
Adj. R-squared		.3774	

Notes: Standard Errors in parentheses
- * Significant at the 10% level.
- ** Significant at the 5% level.
- *** Significant at the 1% level.

The make, model, quarter, day of week, and week of month dummies used in Table 4 were included in this regression but are not reported. As before, they were statistically insignificant.

A joint test of black interaction terms listed in the right-hand column strongly reject the hypothesis that the interaction terms are equal to zero. This suggests that the purchaser's race does not merely affect the profits as a fixed markup but depends on other aspects of the transaction. Two of the interactions are highly significant (at the 1% level). While white customers who finance with the dealership tend to pay $767 more than white customers who don't use dealership financing, the regression suggests that black customers who finance with the dealership pay an additional $457. This suggests, for example, that a black male who financed could expect to pay $1572 (348+767+457) more than a white male who did not use dealership financing.

The regression also suggests that Salesperson 2 exhibited the greatest racial disparity in pricing. A white male purchasing from Salesperson 2 could expect pay total profits that were $1718 ($900 +470+ $348) lower than a black male purchasing from the same salesperson. When combined with evidence that blacks were systematically steered toward Salesperson 2 (discussed above), this is consistent with a story that Salesperson 2 specialized in discriminating against charged black customers.

In an unreported regression, I added to the regression in Table 5 interactions of a female dummy variables to the same transaction characteristics that had black interactions. The interactions were uniformly insignificant with one exception. The total profits for women who financed were $245 more than for men who financed and this result was marginally significant (at the 10.5% level). Still the overall picture for the Atlanta data is non-gendered pricing – in that white men and women paid similar profits and that black men and black women paid similar profits – at least, in comparison with the larger and more robust racial differences.

III. THREE NEW TRANSACTION STUDIES USING DEALER MARKET RESEARCH DATA

In the last two years, two new articles (H&H, SZS I)[49] have been written exploiting transaction data from *J.D. Powers & Associates* and an unnamed research firm in the automotive industry. This data includes crucially information about the vehicle profits analyzed in the data set – but for thousands of dealerships instead of the single Atlanta dealership.[50] 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 be in the dramatic change with the resurgence of no-haggle dealerships on one hand and the rise of Internet sales, referral services and research on the other.

A. *The Harless and Hoffer Study*

The first published study based on the new J.D. Powers data was by Harless and Hoffer [H&H]. H&H analyzed 4030 purchases made from more than 2300 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.

[49]See *supra* note 12.

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

> 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 data set 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.[51]

But to my mind the even bigger concern with the study is that the researchers could not observe the gender(s) 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).[52]

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 H&H did not find a statistically significant gender disparity. The study found that found that women paid an average of $29 (or 1.9% higher

---

[51] H&H, *supra* note 12, at 271-72.

[52] *Id.* At 274.

mean gross profit) more for new cars than men, but the disparity was insignificant ($p = 0.47$).[53] Still, the finding of a small and statistically insignificant gender disparity conforms with the results of both the previous audit and purchase studies.

B. *The Scott Morton Zettlemeyer and Silva-Risso Study*

SZS exploited an even larger data set 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 H&H study, SZS observed individual transactions. But like H&H, SZS did not observe the gender or race of the bargainers. Like H&H, they inferred the gender of the titled purchaser by making probabilistic inferences about the purchaser's first name. And they made inferences about the purchaser's race by exploiting census data about the racial composition of the ""block group" (on average 1100 people) where the purchaser resided. So this initial SZS study suffers from the same problems as H&H 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 (say, for a white to bargain on behalf of black purchaser) than cross-sexual 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, SZS find that black purchasers are expected to pay $456 more than white

[53] *Id.* At 275, & 276 Table 2, second column.

purchasers,[54] and this result is statistically significant.[55] Hispanic purchasers faired even worse – paying an estimated (and statistically significant) $523 more than whites, while Asian purchasers faired the best paying $218 less then whites (again statistically significant). In contrast, female purchasers are predicted to pay just $48 more than male purchasers (statistically significant).[56]

The SZS 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, it 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.[57]

Their results are striking. They found that *Autobytel* users paid approximately $273 less (1.2%) than non-*Autobytel* users.[58] Moreover, they found a marked decline in racial disparities – with African American

[54]Literally a purchaser from an all black block group is predicted to pay $456 more than a purchaser from an all white block group.

[55]See SZS I, *supra* note 12, at 27 (Table 5, col. 1).

[56]These disparities should be interpreted as evidence of disparate racial impacts. When the authors control for a variety of non-racial characteristics concerning the transaction, the racial and gender disparities drop to $342 (but the gender disparity remains $48). *Id.* At 24 (Table 2, col. 1).

[57]The race of purchasers in *Autobytel* transactions might still be inferred by residence and at times by telephonic inferences. 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.

[58]*Id.* At 28 (Table 6, col. 1).

users paying only $68 more than white users (compared to $456 differential found in traditional sales); female users paid only $21 more (compared to the $48 differential found in traditional sales); and Hispanic users were estimated to pay about $285 *less* than Anglo users (compared to the $523 more they had to pay in comparison to Anglos in traditional sales).[59]

SZS have already followed up their excellent study with an unpublished second study 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 5200 consumers – eliciting information about how informed the bargainer was and how he or she bargained.[60] The resulting data set (after accounting for non-responses and incomplete surveys) had 1507 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 vs. 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 the data, with much better information about gender of bargainer finds identical gender disparities – to wit, that women paid about a half a percent more than men in traditional bargains.

III. META-ANALYSIS OF THE SEVEN STUDIES

There now exist six different data sets to help assess whether blacks are discriminated against in car

[59]*Id.*

[60]SZS II, *supra* note 12.

purchasing. And 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 6 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:

(1) the transactions differ: the audit testers solicited offers, but did not purchase cars; the Goldberg, Atlanta, H&H and SZS 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 data is based on differences in total profit (including financing but excluding trade-in profit); and the H&H and SZS data is 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 nonverbal dimensions while the completed transaction data has no ex ante control and we lack basic information about how purchasers bargained, making it impossible to control ex post with regressions.

(4) the racial groups differ: the audit, Atlanta, and SZS studies are tests of black/white disparities, while the Goldberg study is a test of "minority"/"non-minority" disparities and the H&H data contains no racial information.

(5) the geographic areas differ: the audit data comes from Chicago; the Goldberg, H&H and SZS studies are based on a nationwide sample; and the Atlanta data is 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 H&H data covered February 2000; and the SZS study covered 1999 and the first two 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) are uniformly positive and substantial (ranging from $274 to $1133). Moreover, the black male and black female differentials are highly significant in four of the five data sets. Also in four of five datasets, black women are estimated to pay more then black men. And for heuristic purposes, if we combine the observations from these six 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.[61] Thus while the differentials uncovered by Goldberg 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 335 black female observations) is not sufficient to render the global analysis

[61]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.

statistically insignificant.

Indeed, several things are striking when comparing the Goldberg differentials with the differentials from the other data sets. Goldberg's standard errors for black males and black females are more than twice (and sometimes 3 and 4 times the size of) the comparable differentials, and her R-squared is less than half all of the other regressions. These are strong indications in the noisiness of her data. But the size of the differentials themselves are not so far from at least some of the other data sets. Goldberg's black female differential of $426 is similar to the $465 differential from the full audit study. And Goldberg's black 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 male differentials are near $300 and two others are more than twice this amount ($611 and $1133). Three of the black female differentials are in the $450-$500 range, but the other two are roughly twice this amount ($865 and $1013). These higher differentials are not only statistically distinct from Goldberg's estimates, but 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 tables far outweigh the dissimilarities.

Table 6: Meta-Analysis of Estimated 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 R-Squared
Chicago "Pilot" Audit (Ayres)	220 (129) [21]	1013*** (124) [23]	283*** (136) [18]	.37
Chicago Full Audit (Ayres/Siegelman)	216* (116) [53]	465*** (103) [60]	$1133*** (122) [40]	.33
National CES Consummated Sales (Goldberg)	129 (117) [244]	426 (525) [28]	$274 (263) [39]	.14
Atlanta Consummated Sales	-11 (97) [164]	865*** (92) [224]	611*** (96) [178]	.36
JD Powers Feb 2000 Sales (H&H)	29 (38.9) [2015]†			.76
SZS I	47.9*** (3.2) [241729]	503.7*** (12.1) [14383]	445.8*** (11.6) [38514]	.97
Weighted Average	47.8*** (3.7) [244226]	509.7*** (14.7) [14718]	457*** (12.4) [38789]	
SZS I (autobytel purchases)	20.5 (13.6) [6800]	88.9 (66.4) [204]‡	68.4 (65.0) [306]‡	.98

Notes and Sources:

Row 1, Ian Ayres, Harvard Law Rev.

Row 2, Pervasive Prejudice, Chapter 2, Table 2.1 (Col. 2)

Row 3, Pinelopi Goldberg, *Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey*, 104 J. POL. ECON. 622 (1996), Table 2, Col. 1 and Table 5. Goldberg's tested for differences between "Minority" and "Non-Minority" purchasers. See supra note 15 and accompanying text.

Row 4, This paper, Table 4 (Col. 6)

Row 5, H&H, *supra* note 12, Table 2, Col. 1

Row 6, SZS I, *supra* note 12, Table 5, Col. 1

Row 8, SZS I, *supra* note 12, 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

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, her analysis does not contradict, but adds marginal confirmation to this result. Counter to the conjectures of Heckman and Epstein, additional search and/or alternative bargaining strategies seen in traditional negotiations do not eliminate or even importantly 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 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 referals may ultimately help confirm the Heckman/Epstein. The last row of Table 6 report the analogous differentials found in the SZS I 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, SZS II 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. While the pentration 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.

The results of this paper's analysis of the Atlanta dataset concerning racial disparities in financial profits are also supported by yet another litigation-generated analysis. A class action suit against General

Motor's credit division, General Motors Acceptance Corp. (GMAC), has uncovered that financing profits for African American borrower's are systematically higher than for white borrowers.[62] 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 are $377 higher than the profits. These are not quite as high as the racial differentials uncovered in the Atlanta dataset (which as shown in Table 4 range from $453 to $637), but are nonetheless highly significant.

The plaintiffs in this litigation are using the Equal Credit Opportunity Act (ECOA),[63] which allows plaintiffs to bring racial disparate impact suits in lending. In this case, the plaintiffs class alleges that the financing companies decision to allow dealerships to negotiate have an unjustified disparate impact on African American borrowers. This statute has provided a new weapon to attack not just racial disparities in excess

[62] Coleman v. General Motors Acceptance Corp., No. 3-98-0211 (M.D. Tenn. 2000). I have been retained as a plaintiff's expert in a number of cases challenging the disparate racial impacts of dealership markups.

[63] 15 U.S.C. §§ 1691-1691f (1994 & Supp. IV 1998); see also 12 C.F.R. § 202.1; Interagency Policy Statement, 1994 WL 128417.

interest charged – but also to attack racial disparities in the underlying purchase price of the car (the principal of the loan).

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, for example, has been a leading critic of audits for not adequately controlling for unobserved variables – factors other than race or gender but correlated with these traits which might offer non-discriminatory explanation for the audit results.[64] 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 calls a "double bind")[65] comes however when researchers produce an effective test where blacks and whites (men and women) *do* behave the same. Then comes Heckman 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 don't. If you don't 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

[64] Heckman, supra note 2.

[65] Margaret J. Radin, *The Pragmatist and the Feminist*, 63 S. Cal. L. Rev. 1699 (1990).

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.