Consumer Information and Price Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?*

Fiona Scott Morton
Yale University and NBER

Florian Zettelmeyer
University of California at Berkeley

Jorge Silva-Risso
J.D. Power and Associates

September 2001

*We thank Ian Ayres, Sharon Oster, Daniel Snow, and participants at the 2001 NBER IO Summer Institute for helpful comments. We also had many stimulating discussions with David Levine.

Addresses for correspondence: School of Management, Yale University, PO Box 208200, New Haven CT 06520-8200; Haas School of Business, UC Berkeley, Berkeley CA 94720-1900; J.D. Power and Associates, Power Information Network, 30401 Agoura Road, Agoura Hills, CA 91301. E-mail: fiona.scottmorton@yale.edu, florian@haas.berkeley.edu, jorge.silva-risso@powerinfonet.com

Consumer Information and Price Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?

Abstract

Mediating transactions through the Internet removes important cues that salespeople can use to assess a consumer's willingness to pay. We analyze whether the Internet's obfuscation of individual consumers characteristics affects negotiated equilibrium prices in car retailing. Using a large dataset of transaction prices for new automobiles we find that off-line consumers from census blocks with African-American or Hispanic residents pay more than do other consumers. However, we can explain 70% of this price premium with differences in income, education, and search costs; we find no evidence of statistical race discrimination. In contrast, online minority buyers who use the Internet Referral Service we study, Autobytel.com, pay almost the same prices as do whites. Since members of minority groups who use the Internet may not be representative, we control for selection. We conclude that the Internet is disproportionately beneficial to those who have personal characteristics that put them at a disadvantage in negotiating. African-American and Hispanic individuals, who are least likely to use the Internet, are the ones who benefit the most from it.

1 Introduction

"On the Internet, nobody knows you're a dog."

Before the Internet established itself as an important tool for communication, information search, and purchasing, Peter Steiner foresaw that the emerging medium would create some degree of anonymity for its participants (see his famous 1993 *New Yorker* cartoon above). While such anonymity has garnered much attention in the context of Internet chat rooms, many observers of the commercial uses of the Internet have focused on the opposite effect, namely, that the Internet allows firms to better identify the consumer to whom they are selling. This is because—so the argument goes—firms can learn about consumers' behavior through repeat interaction, enabling "targeted pricing," and "product personalization."

For many products, however, the Internet does not make it easier for firms to identify relevant consumer characteristics. This is because mediating interactions through a computer removes important cues that salespeople can use to determine a consumer's willingness to pay. For example, it is difficult for an online vendor of consumer electronics to recommend products based on how sophisticated a particular consumer appears—a practice that is commonplace in conventional stores. Knowledge of consumer characteristics is even more important in markets in which prices are not posted but instead are negotiated.

In this paper we analyze whether the difficulty of dealers in identifying consumer characteristics and ease of consumers in finding information on the Internet affects equilibrium prices in car retailing—a large industry in which prices are negotiated. Judging from advice on how to negotiate for a new car in publications such as Edmunds.com and Consumer Reports, the price a consumer can obtain depends on her information about dealer costs, prices at other dealers, and alternative car models. The opening price of the salesperson is also important. Her price may also depend on the time of the day, the day of the week, and the consumer's willingness to wait for a deal. Consequently, negotiated prices are likely to be correlated with demographic characteristics of consumers. For example, consumers with better education may find it easier to obtain relevant information. Also, higher income consumers may have a higher valuation of time and may thus be less patient.

To test whether the Internet's obfuscation of consumer characteristics affects equilibrium prices, we first need to establish that *offline* negotiations result in differing car prices depending on individual consumer characteristics. Of particular interest to us is whether characteristics exhibited by consumers of different races and genders can explain variation in new car prices. The two major papers in this literature come to different conclusions. Ayres and Siegelman (1995) run a careful audit with "testers" of different races and genders who are trained to bargain identically. They find that Chicago area car dealers offer black male testers and black female testers prices that are significantly higher, by $1100 and $410 respectively, than those offered to white men. In contrast, Goldberg (1996), using data from the Consumer Expenditure Survey, finds no statistical difference in the mean price paid by white and minority consumers, and thus no evidence of discrimination. In fact, she finds that none of her demographic controls, not just the race and gender indicators, play a role in explaining new car prices. She does find, however, that the variance of prices paid by blacks is higher than that of whites. The 90th percentile black consumer pays more than an equivalent white consumer, while the 10th percentile black consumer pays less. Goldberg (1996) reconciles her findings with those of Ayres and Siegelman (1995) by arguing that dealers should be expected to offer higher initial prices to blacks than to whites, even if the average transaction price were the same.[1]

The existing literature thus leaves unresolved the question of whether women and racial minorities pay higher prices on average than do white males. Furthermore, in light of Goldberg's results, it is not certain that *any* individual consumer characteristics affect car prices. We can answer these questions because we have access to data that are unusually well-suited to this purpose. We have detailed information on approximately 700,000 individual car purchases across neighborhoods with varying demographics.

Without controlling for any other demographic characteristics, we find that black and Hispanic buyers pay on average about 2% more ($460 for the average car) than do white buyers for identical cars. After including neighborhood averages for education, income, wealth, and occupation, the minority premium declines to 1.5% for blacks and 1.1% for Hispanics. Including proxies for search costs diminishes the premium further, to .7%. Thus, 65% of the minority price premium can be attributed to observable individual differences in income, education, and search costs. We find a very small price premium for women, .2% ($45).

[1] For an excellent survey of discrimination in retail markets see Ayres (2002).

We test for the effect of the Internet with data on whether consumers used the Internet referral service Autobytel.com in purchasing their car. Autobytel.com allows consumers to request a price quote from an affiliated dealer without engaging in personal interaction. Consequently, dealers are not exposed to cues that signal a consumer's willingness to pay.[2] In addition, Autobytel.com reduces search costs and provides consumers with information. In line with our hypothesis, we find that the minority premium declines to an insignificant level for buyers who use Autobytel.com.

We conclude, first, that pricing of new cars strongly depends on individual characteristics of car buyers, in particular demographics and search costs. Second, we conclude that the Internet eliminates most variation in new car prices that is due to individual characteristics associated with race and ethnicity. Our results have important policy implications. If use of the Internet is likely to reduce the adverse effects of poor education and income, then the so-called "Digital Divide" is of even greater importance and concern. The very people who seem to most benefit from using the Internet are those who systematically are less likely to have access to it.

This paper proceeds as follows. Section 2 contains a discussion of the likely effect of Autobytel.com on differential pricing. Section 3 is a description of the data. In section 4 we establish that car prices depend on individual consumer characteristics. In section 5, we show that the Internet reduces most of the remaining difference in pricing between ethnic groups. Section 6 is an analysis of whether there is evidence of statistical discrimination, and section 7 concludes the paper.

2 Autobytel.com's effect on differential pricing

Autobytel.com is an independent Internet referral service that offers consumers detailed information about individual cars, including current market conditions and invoice pricing. At any point a consumer may submit a free purchase request that is forwarded to one of Autobytel.com's contracting dealers. The consumer provides her name, address, contact information, and the type of car she is looking for. A salesperson at the dealership contacts the consumer within 48 hours (often much sooner) with a price. While Autobytel.com strongly encourages its contract dealers to set a fixed price, dealers are free to deviate from the initial price offer in response to consumer negotiation.[3] Communication may occur by email or telephone. In this

[2]However, dealers could (imperfectly) identify members of certain minority groups and female buyers by their names.

[3]According to J.D. Power and Associates (2000a), 42% of dealerships claim that their initial price contains no room for further negotiation. 42% give discounts but leave room for negotiation. 14% will quote a discounted price only if the customer insists by e-mail or phone. 2% of dealerships don't give discounted price until the consumer comes to the dealership.

way a consumer may purchase a car without setting foot in the dealership until she picks up the vehicle. Autobytel.com assigns dealers an exclusive territory; any leads generated within that territory are passed on to the dealer in exchange for a dealer subscription fee. As of the year 2000, Autobytel.com contracted with approximately 5,000 of the 22,000 US dealerships.

Car prices are individually negotiated, so there is opportunity for significant price discrimination in the market. The same car sells for different prices because consumers differ in characteristics. The economics literature has focused mainly on patience, search costs, and information as the characteristics that affect negotiated prices (Admati and Perry 1987, Salop and Stiglitz 1977). The Internet is likely to change such price discrimination, first, because consumers can obtain more information, second, because services such as Autobytel.com train dealership salespeople to treat consumers in a uniform manner, and third, because many of the personal characteristics of consumers are no longer observable. However, one can also argue that the Internet might make price discrimination easier since a dealer knows a consumers' name and address prior to offering a price. We discuss these arguments in sequence.

Autobytel.com and other online services allow consumers to determine features and specifications of new cars and also to read reviews. This may narrow down a consumer's search to fewer vehicles, thereby reducing her search costs. In addition, a consumer can learn the invoice price of the vehicle she is interested in. While this is not a perfect measure of the dealer's marginal cost, it is a good measure, and can help the buyer determine dealer surplus.

The manner in which Autobytel.com trains salespeople at contracting dealerships may also contribute to different bargaining outcomes. The "Internet salesperson" is supposed to handle only Internet referrals and not "walk-ins." Also, he is supposed to be compensated on sales volume rather than margin. This decreases the Autobytel.com salesperson's incentive to look for individual characteristics that indicate a weak buyer's bargaining position. In addition, Autobytel.com encourages Internet salespeople to charge a uniform price.

Also, the Internet removes important cues that salespeople can use to determine a consumer's willingness to pay. A salesperson cannot take into account the buyer's clothing, vehicle, or accent as signals of her reservation value or bargaining ability. In addition, a consumer may feel less indebted to an Internet salesperson; he did not invest time to show her the car, explain its features, or to take her on a test drive. Referred consumers may thus be willing to bargain harder.

The preceding arguments suggest that dealers should be less likely to price discriminate for on-line than off-line consumers. Thus, Autobytel.com might help certain types of consumers more than it does others. Consumers who lack information or have characteristics that indicate they are poor at bargaining should benefit the most from Autobytel.com because they benefit more than do other consumers from information, fewer cues about their type, and uniform pricing policies.

However, one can also argue that the Internet actually facilitates price discrimination. This is because a purchase request contains name and address and could thus be used to infer gender, ethnicity, and neighborhood. At a minimum the dealer could look up the average demographics of the consumers' zip code; at a maximum the dealer could purchase individual-level data of the type normally used by direct marketers, and condition on likely ethnicity and gender.

3 Data

Our main data come from J.D. Power and Associates (JDPA). JDPA collects transaction data from a random sample of dealers in the major metropolitan areas in the United States. We have data containing every new car transaction at those dealerships from January 1, 1999 to February 28, 2000. This includes customer information, the make, model and trim level of the car, financing, trade-in information, dealer-added extras, and the profitability of the car and the customer to the dealership. We add to this data census demographic information, measures of dealer competition, and information on whether a consumer submitted a purchase request using Autobytel.com. After dropping observations with missing data, our dataset contains 671,468 transactions at 3,562 dealerships. Summary statistics are in the Appendix.

3.1 Dependent variable

We define *Price* as the price the customer pays for the vehicle, factory installed accessories and options, and dealer-installed accessories contracted for at the time of sale that contribute to the resale value of the car. We subtract the *ManufacturerRebate*, if any, given directly to the consumer. We also subtract what is known as the *TradeInOverAllowance*. This is the difference between the trade-in price paid by the dealer to the consumer and the estimated wholesale value of the trade-in vehicle (as booked by the dealer). We adjust for this amount to account for the possibility, for example, that a dealer may offer a consumer a low price for the new car because the dealer is profiting from the trade-in.

3.2 Measures of race and gender

Our data on race and gender are of two types, census block group level data and individual level data. A "block group" makes up about one fourth of the area and population of a census tract. On average, block groups have about 1100 people in them, and we will refer to them hereafter as census blocks. J.D.Power matches census data from the buyer's address to the transaction record. The census variables that pertain to race are *PctHispanic*, *PctBlack*, and

PctAsian, which measure the percent of residents in a census block that indicate they belong to those groups.

On the individual level, J.D.Power records *Age* directly. Gender and race are coded as what JDPA calls "target" variables. They are created by software programs that analyze the buyer's first and last name. JDPA compares the first name to a list of common female first names and creates a "probably female" variable. This will be our *Female* variable, where a one indicates a female customer. JDPA also looks for common Chinese and Japanese last names. These we combine into an indicator variable called *Asian*. The buyers that JDPA classifies as having Hispanic last names get a value of one in our *Hispanic* indicator variable. Notice, however, that it is not clear that the dealer's perception of "Hispanic" is better captured by the name variable than the census neighborhood variable. This is because of a potential difference between having a Hispanic surname, coming from a Hispanic neighborhood, a persons's self-perception as a Hispanic, and the dealer's perception that a consumer is Hispanic. JDPA also identifies some other races such as Native American and Pacific Islander through common names, but the numbers are so small that we do not use this information.

The median percent black, Hispanic, and Asian in buyers' census blocks are 1.3%, 4.5%, and 2.2% respectively. See Figures 1-3 for the distribution of neighborhood minority percentages for new car buyers (our data). The sample includes buyers from blocks that are 100% Asian and 100% black, but the Hispanic maximum is 55%. 12,150 of our buyers (1.7% of the sample) come from census blocks with greater than 75% black residents. The JDPA name analysis results in 8% of our new car buyers being classified as likely Hispanic, 2% being classified as likely Asian, and 36% as likely female.

Table 1: Summary statistics of census race variables

Variable	Median	90th Pctile	Maximum
% Asian	2.2	13	100
% Black	1.3	13	100
% Hispanic	4.5	22	55

To establish the relationship between JDPA race variables and census data, we examine block groups where the percentage of Hispanics is greater than 50%. We tabulate the JDPA indicator variable *Hispanic* for that sample. We find that 62% of these consumers are considered Hispanic by JDPA. This suggests the JDPA procedure does very well at identifying Hispanic consumers. We repeat the test for *Asian* and find that that JDPA considers only 22% of consumers to have Asian names in census blocks where over 50% of residents identify themselves as Asian. This may be because Asian last names are harder to categorize or because they buy fewer cars. We double check the reliability of the indicators by repeating this procedure on blocks with zero *PctHispanic* and *PctAsian*. The results for the second trial yield 2% Hispanic names and .5% Asian names, a reasonable level considering that residents select their racial groups and that marriage may create some ambiguity. We will use the JDPA indicator variables in the remainder of the paper, recognizing that the Asian indicator is somewhat less reliable than the Hispanic indicator. Note that we have no way of checking the accuracy of the "likely female" variable.

The major racial group not identified on the basis of last names is African-American. However, we know the percentage of any given census block that is black. We use the relationship between *PctHispanic* and *Hispanic* and *PctAsian* and *Asian* to infer the effect of being a black customer in addition to living in a minority census block.

3.3 Data on usage of Internet Referral Services

To test for the effect of Internet usage we use purchase requests submitted by consumers on Autobytel.com during 1999. Autobytel.com referred slightly over 2 million customers to dealers. We consider a match between observations from Autobytel.com and JDPA when the geocoded address or phone associated with the referral and the purchase transaction are the same. Each observation in the new dataset is a transaction from the JDPA data, augmented with the information from the Autobytel.com data if there was a match. We have (1) an indicator for Autobytel.com customer (ABT) indicating that the customer who purchased the car submitted a purchase request using Autobytel.com (irrespective of whether this purchase request went to the dealer that sold the car), (2) an indicator for Autobytel.com franchise dealer (*ABTFranchise*) indicating that the dealer who sold the car is an Autobytel.com affiliated dealer, i.e. is under contract with Autobytel.com and receives purchase requests, and (3) an indicator (*SameDealer*) marking cases in which the dealer that sold the car is the same dealer to whom the purchase request was submitted (given that ABT=1).

We restrict ourselves to observations in which an Autobytel.com user purchased a make and model for which she requested a referral. This is to ensure that Autobytel.com consumers received an initial price quote for the purchased automobile without having had to have stepped into the dealership. This eliminates about 3% of observations.

Autobytel.com was the leading Internet Referral Service in 1999.[4] However, since there are online referral services other than Autobytel.com, the customers in the combined dataset who are not identified as using Autobytel.com may have used one of its competitors. This strengthens our test since we will be comparing a group that used Autobytel.com to a group

[4]Autobytel.com had between 45 and 50% market share of online car shopping in 1999 (LA Times, 3/28/2000, "Mergers and Acquisitions Report," Securities Data Publishing 6/12/2000). According to J.D. Power and Associates (2000b), Autobytel.com is the most visited purchase referral site. It is visited by 33% of consumers that researched online to shop for a car, followed by Autoweb.com (18%), and Carpoint.com (17%).

that may include users of competing services.

3.4 Controls

We control for car fixed effects. A "car" in our sample is the interaction of make, model, body type, transmission, displacement, number of doors, number of cylinders, and trim level. We control for 834 "cars" after dropping "cars" with fewer than 300 sales. We do not have information on options that are outside of trim levels, which is why we include the percent deviation of the dealer's *VehicleCost* from the average *VehicleCost* of that car in the dataset. The *VehicleCost* is the retailer's 'net' cost for the vehicle and includes the cost of accessories added by the factory and/or retailer and included in the customer's contract that add to the vehicle's book value. The measure takes into account holdback and includes transportation charges.

To control for time variation in prices we define a dummy *EndOfMonth* that equals 1 if the car was sold within the last 5 days of the month. Dealers who want to meet volume targets for the month often have sales or other inducements to purchase near the end of the month. A dummy variable *WeekEnd* specifies whether the car was purchased on a Saturday or Sunday for the same reason. In addition, we introduce dummies for each month in the 14 month sample period to control for other seasonal effects and inflation.

To control for how "hot" a car is and the dealer's opportunity cost of not selling it, we control for the number of months between when a car was sold and its introduction. Judging by the distribution of sales after car introductions we assign a dummy variable to sales in the first four months, months 5-13, and month 14 and later.

We also control for the competitiveness of each dealers's market. For each dealership we count the number of dealerships of the same nameplate that fall in a zip code that is within a 10 mile radius of the zip code of the focal dealership. We control for cases where one owner owns several franchises. Hence, our *Competition* measure counts only the number of separately-controlled entities. Finally, we control for the 17 regions in which the car was sold.[5]

4 Prices vary with demographics

We begin our analysis by estimating the effect of various demographic measures, including race, gender, and age. We estimate the following specification:

[5] For a more detailed description of many of the variables in the data, see our earlier paper Scott Morton, Zettelmeyer, and Silva-Risso (2001).

$$\ln(Price_i) = \gamma D_i + \beta X_i + \epsilon_i$$

The D matrix contains demographic information about the purchaser in transaction i as described above. The X matrix is composed of transaction and car variables: car, month, and region fixed effects, controls for time variation, competition, car cost, and whether a consumer traded in a vehicle.

4.1 Results

Our first specification includes census demographic information but no JDPA race variables. We expect income and education to be positively correlated but to have the opposite effect on transaction prices. High income indicates a lower elasticity of demand and a higher opportunity cost of time, while high educational levels may make a person a more effective negotiator. Hence, we have few priors on the signs of our census block variables. We find that most are significant: higher income lowers car prices until the average block income reaches $130,000, at which point the effect becomes positive. A higher probability of being a blue-collar worker and high house values lower prices. Coming from a block with a higher percentage of people who have gone to college also lowers prices. Percent professional and executive are insignificant. Home ownership, a proxy for creditworthiness, lowers prices. A higher probability of not finishing high school increases customer price.

We find that women pay more for cars (.2%), as do older consumers (.2% for moving from 20 to 64 years old) and consumers who have a higher probability of being either black or Hispanic (see column 1 of Table 2). A buyer with probability one of being black pays 1.5% more for the equivalent vehicle than does a buyer that has probability zero of being black (an increase of 100% percent black in a census block group). An increase from zero to one in the probability of being Hispanic raises the expected price of a new car by 1.1%. People from census blocks with Asians pay less for new cars; an increase from zero to one in the probability of being Asian lowers the price of a car by about .4%. All age, gender, and race coefficients are significant at the 1% level.

Our second specification includes the JDPA race variables and is reported in column 2 of Table 2. The coefficients of *Asian* and *Hispanic* are statistically significant and have the same sign as in the census specification. Including these variables reduces the size of the census block coefficients in each case. The coefficient on *Hispanic* is .5% while the coefficient of *PctHispanic* falls to .7%. This results in almost the same total effect as in the previous specification. Adding an indicator variable for *Asian* raises the total effect of being Asian; this racial group pays 1%

Table 2: Regressions for results section†

	(1)	(2)	(3)	(4)	(5)
Dep. Variable ln(price)	Full Sample	Full Sample	> 75% or < 2% Black	> 75% or < 2% Black	Full Sample
%Black	0.00015 (0.00001)**	0.00015 (0.00001)**			0.00014 (0.00001)**
%Hispanic	0.00011 (0.00001)**	0.00007 (0.00001)**	0.00003 (0.00001)*	0.00003 (0.00001)*	
%Asian	-0.00004 (0.00001)**	-0.00001 (0.00001)	-3.08e-06 (0.00001)	-3.06e-06 (0.00001)	-0.00001 (0.00001)
Asian		-0.00966 (0.00043)**	-0.00854 (0.00058)**	-0.00854 (0.00058)**	-0.00966 (0.00043)**
Hispanic		0.00505 (0.00028)**	0.00539 (0.00038)**	0.00539 (0.00038)**	0.00593 (0.00033)**
%Black> 75			0.01366 (0.00062)**	0.01297 (0.00087)**	
%Black> 75*Female				0.00127 (0.00116)	
Hispanic*Female					-0.00152 (0.00054)**
Female	0.00206 (0.00014)**	0.00209 (0.00014)**	0.00194 (0.00018)**	0.00190 (0.00018)**	0.00221 (0.00014)**
CustomerAge	0.00004 (0.00001)**	0.00005 (0.00001)**	0.00003 (0.00001)**	0.00003 (0.00001)**	0.00005 (0.00001)**
Age> 64	-0.00168 (0.00030)**	-0.00168 (0.00030)**	-0.00145 (0.00037)**	-0.00145 (0.00037)**	-0.00171 (0.00030)**
MedianHHIncome	-1.63e-07 (1.39e-08)**	-1.67e-07 (1.39e-08)**	-1.77e-07 (1.71e-08)**	-1.77e-07 (1.71e-08)**	-1.47e-07 (1.37e-08)**
$(\text{MedianHHInc.})^2$	1.26e-12 (7.58e-14)**	1.25e-12 (7.57e-14)**	1.23e-12 (9.11e-14)**	1.23e-12 (9.11e-14)**	1.17e-12 (7.52e-14)**
%CollegeGrad	-0.00003 (0.00001)**	-0.00003 (0.00001)**	-0.00001 (0.00001)	-0.00001 (0.00001)	-0.00003 (0.00001)**
%<HighSchool	0.00004 (0.00001)**	0.00003 (0.00001)*	0.00003 (0.00002)*	0.00003 (0.00002)*	0.00006 (0.00001)**
%HouseOwn.	-0.00003 (4.53e-06)**	-0.00003 (4.53e-06)**	-0.00002 (0.00001)**	-0.00002 (0.00001)**	-0.00004 (4.37e-06)**
%Professional	0.00005 (0.00001)**	0.00005 (0.00001)**	0.00002 (0.00002)	0.00002 (0.00002)	0.00004 (0.00001)**
%Executives	-1.34e-06 (0.00001)	7.70e-07 (0.00001)	-0.00001 (0.00002)	-0.00001 (0.00002)	-0.00001 (0.00001)
%BlueCollar	1.77e-06 (0.00001)	2.35e-06 (0.00001)	0.00001 (0.00001)	0.00001 (0.00001)	0.00001 (0.00001)
%Technicians	0.00005 (0.00003)	0.00004 (0.00003)	-0.00001 (0.00004)	-0.00001 (0.00004)	0.00004 (0.00003)
MedianHouseVal.	-2.73e-08 (1.28e-09)**	-2.58e-08 (1.28e-09)**	-2.38e-08 (1.60e-09)**	-2.38e-08 (1.60e-09)**	-2.71e-08 (1.26e-09)**
EndOfMonth	-0.00345 (0.00015)**	-0.00345 (0.00015)**	-0.00356 (0.00020)**	-0.00356 (0.00020)**	-0.00346 (0.00015)**
Weekend	0.00112 (0.00016)**	0.00111 (0.00016)**	0.00102 (0.00020)**	0.00102 (0.00020)**	0.00111 (0.00016)**
VehicleCost	0.88183 (0.00134)**	0.88155 (0.00134)**	0.88112 (0.00171)**	0.88112 (0.00171)**	0.88158 (0.00134)**
AnyTrade	0.00309 (0.00014)**	0.00309 (0.00014)**	0.00334 (0.00018)**	0.00334 (0.00018)**	0.00306 (0.00014)**
Competition	-0.00022 (0.00004)**	-0.00022 (0.00004)**	-0.00021 (0.00004)**	-0.00021 (0.00004)**	-0.00020 (0.00004)**
Constant	10.01657 (0.00131)**	10.01676 (0.00131)**	10.03879 (0.00177)**	10.03879 (0.00177)**	10.01632 (0.00131)**
Observations	650850	650850	386155	386155	650850
R^2	0.97	0.97	0.98	0.98	0.97

* significant at 5%; ** significant at 1%. Robust standard errors in parentheses.

† Unreported are car, month, region, and model recency fixed effects
Cell sizes in column 2: *Asian* 13030, *Hispanic* 53847; column 3: *%Black>75* 11205; column 4: *%Black>75**Female 6134; column 5: Hispanic*Female: 18491.

less than others on average, in contrast to -.4% on the basis of the census data alone. These results suggest that—were it to exist—an indicator for African-American would be statistically and economically significant and reduce the coefficient on *PctBlack*, but that it would not change the overall impact of race on price. It also suggests that the census block information picks up some, but not all of the race indicator effect.

In interpreting the coefficients, there are two marginal effects of interest. One is the difference between probability zero and probability one of being a particular race. The other is the price premium for a targeted minority in an average census block. This is obviously a much smaller number, since, for example, the average census block has 1% black residents. The next two specifications show that the zero to 100% interpretation is more appropriate. We restrict the sample to buyers from two types of census blocks: those with less than two percent black residents, and those with more than 75% black residents. The leaves about 386,000 out of the original 650,000 transactions in the sample. We then generate a new indicator *Black* that is a one if the customer is from a census block where more than 75% of people are black. The coefficient on this variable is 1.4% (see column 3 of Table 2). Notice that this coefficient is extremely close to 100 times the *PctBlack* coefficient, or 1.5%.

To see if this procedure replicates the JDPA indicator variable, we repeat it for Asians and compare the coefficient on our indicator variable to the -.97% in column 2. We define the new Asian indicator variable using bounds of 0.5% and 75%. The coefficient on our constructed variable is -1.2%. This is quite close to the sum of the coefficients on the JDPA Asian indicator and the *PctAsian**100, which total -1.1%. However, it is larger than the effect we would estimate by taking 100 times the *PctAsian* coefficient of -.006. These experiments lend support for the interpretation of the percentage coefficients as representing the effect of a buyer changing from being minority with zero probability to 100% probability. (We do not create a new Hispanic indicator because the maximum *PctHispanic* is only 55% and thus too low to create an equivalent variable.)

Our estimates are much smaller than those of Ayres and Siegelman (1995), whose testers find unexplained minority premia of $410 (female) and $1100 (male). We investigate whether our data show the same relationship between minority female and minority male prices. Column 5 of Table 2 shows that the interaction of *Black* with *Female* is positive, insignificant, and additionally, only 0.13%, or about $30 on the average car. The coefficient on *Black* remains fairly steady at 1.3%. Hispanic women appear to pay a little bit less than Hispanic men. The coefficient is again, tiny, at -.15%.

We also run run 10% and 90% quantile regression to see if the variance in minority prices is greater than that of white prices, as found in Goldberg (1996) (see Table 3). We find that a buyer that has a probability one of being black vs. a buyer that has a zero probability of being black pays only .7% more in the 10 percent regression but 2.5% more in the 90% quantile regression. For Hispanics we combine the effect captured in the census and the JDPA variable and find that members of this group pay .26% more in the 10% regression but 1.9% more in the 90% quantile regression. For Asians we also combine the effect captured in the census and the JDPA variable and find that they pay 1.6% less in the 10% regression and the same as whites in the 90% quantile regression. In contrast to Goldberg (1996), we find that blacks and Hispanics pay on average more than whites, even at the low end of the price distribution. Our results on gender are not consistent with Goldberg (1996)'s findings. Females pay .21% more in the 10% regression and .28% more in the 90% quantile regression. While these differences are very small they indicate that the variance of female's reservation price distributions is not smaller than that for males.

We are concerned that our results might be driven by a small group of consumers from poor neighborhoods, so we investigate whether our result holds when we restrict the sample to buyers who live in "good neighborhoods." We repeat our specification restricting the sample to buyers from census blocks with above average educational or income levels. The results are not reported. Neither the coefficient on *PctBlack*, *PctAsian*, or *Female* changes when the sample is restricted to buyers from census blocks where 32% or more of residents have a college education. *PctHispanic* declines to 1% from 1.3%. We find very similar results when we restrict the sample to buyers that reside in blocks with average incomes above the mean of $57,000. The coefficient on *PctBlack* is unchanged at 1.5% and remains significant. The coefficients on Hispanics add to 1.1%, while *PctAsian* remains at -.9%. These results indicate that our basic finding is not driven by one end of the income distribution.

Unlike Goldberg (1996), we find that demographics explain variation in average transaction prices. Our results are directionally consistent with Ayres and Siegelman (1995) but somewhat smaller. We find that black buyers pay about 1.5% more than white buyers, while Hispanic buyers pay a 1.1% premium. Hispanic women pay a little bit below the Hispanic average. We also find that all women pay about .2% more than males (vs. 1.7% in Ayres and Siegelman (1995)). Finally, we confirm Goldberg (1996)'s finding that the transaction price variance is larger for minorities than for whites, consistent with a higher reservation price variance for minorities.

4.2 Explanations

Since Ayres and Siegelman (1995) hold constant the age, education, attractiveness, and bargaining strategy of their testers, they say that their results cannot be driven by differences between consumers other than race and gender. Goldberg (1996) also concludes that demographic dif-

Table 3: Quantile regressions†

	(1)	(2)	(3)
Dep. Variable ln(price)	.1 Quantile	.9 Quantile	Median
%Black	0.00007 (0.00001)**	0.00025 (0.00001)**	0.00012 (4.57e-06)**
%Hispanic	1.53e-06 (0.00002)	0.00010 (0.00002)**	0.00009 (0.00001)**
Hispanic	0.00259 (0.00043)**	0.00861 (0.00054)**	0.00441 (0.00024)**
%Asian	-0.00010 (0.00002)**	0.00014 (0.00002)**	-0.00002 (0.00001)**
Asian	-0.00591 (0.00071)**	-0.01432 (0.00093)**	-0.00867 (0.00046)**
Female	0.00215 (0.00023)**	0.00277 (0.00029)**	0.00144 (0.00013)**
CustomerAge	0.00005 (0.00001)**	0.00006 (0.00001)**	0.00004 (0.00001)**
Age> 64	-0.00184 (0.00049)**	-0.00129 (0.00061)*	-0.00177 (0.00028)**
MedianHHIncome	-1.53e-07 (2.29e-08)**	-2.04e-07 (2.91e-08)**	-1.45e-07 (1.34e-08)**
$(\text{MedianHHInc.})^2$	1.24e-12 (1.26e-13)**	1.39e-12 (1.62e-13)**	1.09e-12 (7.75e-14)**
%CollegeGrad	-0.00001 (0.00002)	-0.00008 (0.00002)**	0.00001 (0.00001)
%<HighSchool	0.00005 (0.00002)*	0.00005 (0.00003)	0.00002 (0.00001)*
%HouseOwn.	-0.00002 (0.00001)*	-0.00006 (0.00001)**	-0.00001 (4.18e-06)**
%Professional	0.00004 (0.00002)	0.00006 (0.00003)*	0.00005 (0.00001)**
%Executives	4.04e-06 (0.00002)	-0.00001 (0.00003)	1.39e-06 (0.00001)
%BlueCollar	9.19e-07 (0.00002)	-0.00002 (0.00002)	0.00003 (0.00001)**
%Technicians	-0.00008 (0.00006)	0.00016 (0.00007)*	-0.00001 (0.00003)
AnyTrade	-0.00007 (0.00023)	0.00586 (0.00028)**	0.00313 (0.00013)**
Competition	-0.00086 (0.00006)**	0.00063 (0.00007)**	-0.00034 (0.00003)**
Constant	9.95829 (0.00224)**	10.07810 (0.00282)**	10.01282 (0.00127)**
Observations	650850	650850	650850

* significant at 5%; ** significant at 1%
Standard errors in parentheses

† Unreported are *EndOfMonth*, *WeekEnd*, *VehicleCost*, car, month, region, and model recency fixed effects

ferences don't affect pricing of automobiles. In contrast, our study finds that demographics other than race can partly explain why minorities pay higher prices. As a baseline regression, we estimate coefficients for African-American and Hispanic buyers that are not conditional on any demographic data except race and gender. We expect the minority premia to rise since minority status is correlated with the demographics that predict higher prices (less education, less home ownership). We find that without controlling for these buyer characteristics, black and Hispanic buyers pay 2.0% ($460) and 2.3% ($530) more for their vehicles, respectively (see column 1 in Table 4). This contrasts with 1.5% and 1.1%, respectively, when income, education, occupation, and wealth are controlled for. If these straightforward differences between consumer type explain 25% to 50% of the price premium paid by minorities, could there be other differences between members of minority groups and whites that can explain the remaining price premium? The following section explores this question.

Minorities may buy at dealers with higher cost: Minorities might pay more than other groups if the dealerships from which they buy have higher cost. This may be because they are located in locations with higher costs of inputs and real estate. We examine this hypothesis by running a price specification with a franchise fixed effect. For reasons of computation we have to restrict the number of car fixed effects and therefore lose about 36,000 observations. We find slightly larger race and gender coefficients (see column 2 in Table 4) in this specification. Hence, the minority premium is not due to purchases at higher cost dealerships.[6]

Minorities may have an aversion to bargaining: If societal factors lead minorities and women to be less effective at bargaining or to dislike the bargaining process more, then they may be more likely to pay higher prices. Since bargaining is easiest for consumers when they can take their business to a competitor, the payoff from being a skilled bargainer should be lower in a competitive market. Hence, if the premium paid by minorities is due to an aversion to bargaining, this premium should be smaller in more competitive markets. To analyze this conjecture we interact our minority and gender measures with our measure of competition. We find that the interaction between race and market structure has little effect on price for African-Americans; the higher prices paid by blacks are the same in monopoly dealership and competitive markets (see column 3 of Table 4). However, Hispanics do statistically worse in more competitive markets, which is counter to theory. We find no evidence that minorities pay more because they have an aversion to bargaining, however.

Minorities may face higher search costs: Given that minorities are less likely to own a car when shopping for a new car, they are also more likely to face above average search costs (Mannering

[6]Because this procedure limits the sample, leaves us unable to study the effects of market structure, and strains available computing power, we do not use the specification throughout the paper.

Table 4: Regressions for explanations section†

	(1)	(2)	(3)	(4)	(5)
Dep. Variable ln(price)	Full Sample	Franchise Fixed Effects	Full Sample	Full Sample	No Financing
%Black	0.00020 (0.00001)**	0.00019 (0.00001)**	0.00013 (0.00001)**	0.00019 (0.00001)**	0.00012 (0.00001)**
%Hispanic	0.00023 (0.00001)**	0.00011 (0.00002)**	0.00001 (0.00001)	0.00014 (0.00001)**	0.00007 (0.00002)**
%Asian	-0.00010 (0.00001)**	-0.00001 (0.00001)	-0.00002 (0.00002)	-0.00004 (0.00001)**	-0.00002 (0.00002)
Hispanic		0.00439 (0.00038)**			0.00323 (0.00061)**
Asian		-0.01088 (0.00061)**			-0.00692 (0.00068)**
Female	0.00206 (0.00014)**	0.00337 (0.00019)**	0.00227 (0.00022)**	0.00205 (0.00014)**	0.00287 (0.00025)**
CustomerAge		0.00011 (0.00001)**	0.00004 (0.00001)**	0.00004 (0.00001)**	0.00002 (0.00001)*
Age> 64		-0.00607 (.00040)**	-0.00169 (0.00030)**	-0.00170 (0.00030)**	0.00294 (0.00044)**
%Black> 75*Comp.			4.22e-06 (2.00e-06)*		
Hispanic*Comp.			0.00003 (2.91e-06)**		
Asian*Comp.			-0.00001 (3.83e-06)		
Female*Comp.			-0.00007 (0.00006)		
AnyTrade	0.00336 (0.00014)**	0.00337 (0.00019)**	0.00307 (0.00014)**	0.00431 (0.00018)**	0.00811 (0.00024)**
%Black> 75*AnyTrade				-0.00011 (0.00001)**	
Hipanic*AnyTrade				-0.00007 (0.00001)**	
Competition	-0.00037 (0.00003)**		-0.00049 (0.00005)**	-0.00022 (0.00004)**	-0.00017 (0.00007)*
Constant	10.00998 (0.00108)**	10.84324 (1.38E+11)	10.01735 (0.00131)**	10.01614 (0.00131)**	10.06766 (0.00253)**
Observations	650850	683129	650850	650850	159819
R^2	0.97	0.95	0.97	0.97	0.98

* significant at 5%; ** significant at 1%
Robust standard errors in parentheses

† Unreported are *EndOfMonth, WeekEnd, VehicleCost*, car, month, region, and model recency fixed effects. In addition, columns 2,3 and 4 include *MedianHHIncome, (MedianHHInc.)*^{2}*, %Executives, %BlueCollar, MedianHouseVal., %HouseOwn., %CollegeGrad, %<HighSchool, %Professional*, and *%Technicians.*

and Winston 1991). Collecting basic information about features, prices, and availability for that vehicle may be much more difficult without a car. Higher prices would result because minorities cannot comparison shop as easily. To examine whether higher search costs explain our estimated race premia, we add an interaction between *PctBlack*, *PctHispanic* and an indicator variable that is one if a customer traded in a vehicle at the dealer. This allows us to analyze whether minorities that owned a vehicle faced similar search costs as average members of the majority, and therefore paid less of a race premium. The results in column 4 of Table 4 show that this is indeed the case. Consumers of all races who trade in a vehicle pay a small premium for that convenience.[7] Among consumers that did not trade in a vehicle, a buyer that has a probability one of being African-American pays 1.9% more for the equivalent vehicle than a buyer who has zero probability. The race premium declines to .7% for black consumers who have traded in a vehicle.[8] The result also holds for the Hispanic indicator variable.[9] This suggests that higher search costs when buying a vehicle may be responsible for a large part of the price premium paid by minorities.

In conclusion, we find that the minority premium of 2.0% or 2.3% (when no demographics are in the regression) declines to .7% price premium described above, when differences between these groups of consumers are included. In particular we find that minorities seem to pay higher prices because on average they face higher search costs.

5 The effect of the Internet

The use of Autobytel.com varies with the racial composition of a census block. The mean use of Autobytel.com in the data is 3.1%.[10] At 2.8%, women are almost equally likely to use the service. Census group blocks with *PctHispanic* above 25% have a usage rate of 1.5% while the same statistics for African-American and Asian blocks are 1.7% and 4.1%, respectively. We use *PctMinority* (the sum of *PctHispanic* and *PctBlack*) to analyze the total percentage of minorities in a census block. Census blocks where *PctMinority* is greater than 75 have only a 1% use of Autobytel.com.

[7]The dealer can switch the plates, the owner does not have to clean, advertise, and recondition the car, etc.

[8]Our base specification already controls for education, income, and whether a consumer traded in a vehicle.

[9]It is possible that buyers with a trade-in are richer or more highly educated, but we have included interactions of these variables in unreported specifications and the marginal effect is not as high as that of the trade-in, we conclude that the trade-in itself must be important. We also try to roughly control for the value of the trade-in by including its booked dollar value as a determinant of ln(price). If trade-in margins are proportional, a higher value trade-in will result in a consumer paying a higher net price for her new vehicle. We find this to be the case, however, the race and trade-in coefficients do not change at all (unreported).

[10]The overall Autobytel.com use is closer to 7% before we drop consumers who buy a different car.

5.1 Result

We begin with a specification that includes an indicator variable ABT that is one if the car buyer submitted a purchase request using Autobytel.com. We also include an indicator variable *AbtFranchise* for Autobytel.com network dealers, and *SameDealer* to identify when the customer both used Autobytel.com and purchased from the referred dealer. The specification is a follows:

$$\ln\left(Price_i\right) = \alpha_1 ABT_i + \alpha_2 ABTfranchise_i + \alpha_3 SameDealer_i + \gamma D_i + \beta X_i + \epsilon_i$$

Column 1 of Table 5 shows that Autobytel.com users pay about 1.2% less than other customers. The first effect is the main Autobytel.com discount of almost 1%. Users are also sent to an ABT dealer for an additional .5% discount (which they would get by chance with 1/3 probability since ABT dealers sell 1/3 of all cars), resulting in additional savings of .32%. Finally, consumers with a referral pay .2% more with 1/4 probability (since 1/4 of Autobytel.com customers buy from the referred dealer), so we add .05%, for a total of about 1.2%.[11]

The inclusion of the Autobytel.com variables does not change our estimates of the price difference paid by female and minority buyers. In the previous section we presented preliminary evidence that people with high search costs pay more for cars. Since Autobytel.com also lowers search costs, we investigate if women and minorities gain disproportionately from using Autobytel.com. Since we have established that these groups pay above average prices, they should benefit more than do other consumers from information, anonymity, and uniform pricing policies.

We take the basic and minority indicator specifications from the previous section and interact race and gender with the Autobytel.com indicator. Column 2 in Table 5 shows that the coefficient on *PctBlack**ABT, is -1.1% and significant. This substantially offsets the *PctBlack* coefficient of +1.5%. The ABT coefficient declines in magnitude because some of the effect is reflected in the interaction. The female interaction coefficient is very small but also negative. Women who use Autobytel.com pay a lower premium, by about $25, than other women. This specification suggests that Autobytel.com helps African-Americans and women recover a substantial part of the price premium they would otherwise pay. The *PctHispanic**ABT has a coefficient of -2.5%, which more than makes up for the premium of 1.2% we estimate for Hispanics. The interaction coefficient seems too large and leaves Hispanics paying 1% below average. This raises questions of the endogeneity of Autobytel.com use.

[11]The higher price associated with purchasing from the Autobytel.com dealer the customer was referred to may reflect that the people who buy from their referred dealer have a high disutility of continuing to price-shop.

Table 5: Regression for Autobytel.com results[†]

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable ln(price)	Full Sample	Full Sample	.1 Quantile	.9 Quantile	> 75% or < 2% Minority	Full Sample
ABT	-0.00927 (0.00033)**	-0.00637 (0.00049)**	-0.00339 (0.00085)**	-0.01255 (0.00108)**	-0.00883 (0.00043)**	-0.00901 (0.00034)**
ABTFranchise	-0.00464 (0.00015)**	-0.00465 (0.00015)**	-0.00421 (0.00025)**	-0.00478 (0.00032)**	-0.00393 (0.00020)**	-0.00465 (0.00015)**
SameDealer	0.00183 (0.00060)**	0.00169 (0.00059)**	0.00443 (0.00115)**	-0.00239 (0.00152)	0.00161 (0.00076)*	0.00182 (0.00060)**
%Black	0.00015 (0.00001)**	0.00015 (0.00001)**	0.00008 (0.00001)**	0.00025 (0.00001)**		0.00014 (0.00001)**
%Hispanic	0.00007 (0.00001)**	0.00012 (0.00001)**	0.00004 (0.00002)*	0.00020 (0.00002)**	0.00008 (0.00001)**	
Hispanic	0.00506 (0.00027)**					0.00589 (0.00027)**
%Asian	-0.00001 (0.00001)	-0.00003 (0.00001)**	-0.00013 (0.00001)**	0.00011 (0.00002)**	-0.00003 (0.00001)**	-0.00004 (0.00001)**
Asian	-0.00955 (0.00042)**					
ABT*%Black		-0.00012 (0.00003)**	-0.00010 (0.00006)	-0.00022 (0.00007)**		
ABT*%Hispanic		-0.00025 (0.00004)**	-0.00008 (0.00007)	-0.00046 (0.00009)**		
ABT*%Asian		-0.00007 (0.00003)*	0.00011 (0.00006)	-0.00028 (0.00007)**		
%Black> 75					0.01352 (0.00062)**	
ABT*%Black> 75					-0.00837 (0.00419)*	
ABT*Hispanic						-0.00759 (0.00145)**
Female	0.00208 (0.00014)**	0.00208 (0.00014)**	0.00213 (0.00022)**	0.00270 (0.00029)**	0.00190 (0.00017)**	0.00210 (0.00014)**
ABT*Female		-0.00117 (0.00058)*				
CustomerAge	0.00005 (0.00001)**	0.00004 (0.00001)**	0.00004 (0.00001)**	0.00005 (0.00001)**	0.00003 (0.00001)**	0.00005 (0.00001)**
Age> 64	-0.00165 (0.00029)**	-0.00165 (0.00029)**	-0.00181 (0.00048)**	-0.00131 (0.00061)*	-0.00142 (0.00036)**	-0.00170 (0.00029)**
AnyTrade	0.00312 (0.00013)**	0.00312 (0.00013)**	-0.00008 (0.00022)	0.00591 (0.00028)**	0.00335 (0.00017)**	0.00315 (0.00013)**
Competition	-0.00030 (0.00003)**	-0.00030 (0.00003)**	-0.00099 (0.00006)**	0.00057 (0.00007)**	-0.00025 (0.00004)**	-0.00029 (0.00003)**
Constant	10.01885 (0.00139)**	10.01875 (0.00139)**	9.97673 (0.00235)**	10.06839 (0.00297)**	10.04059 (0.00175)**	10.01826 (0.00139)**
Observations	671468	671468	671468	671468	398566	671468
R^2	0.98	0.98	0.3405	0.2819	0.98	0.98

[*] significant at 5%; ** significant at 1%. Robust standard errors in parentheses.

[†] Unreported are *EndOfMonth, WeekEnd, VehicleCost, MedianHHIncome, (MedianHHInc.)*^{2}, *%Executives, %BlueCollar, MedianHouseVal., %HouseOwn., %CollegeGrad, %<HighSchool, %Professional, %Technicians*, car, month, region, and model recency fixed effects

Cell sizes: Column 2: *ABT*Female* 6800. Column 5: *%Black>75* 11,205, *ABT*%Black>75* 98. Column 6: *ABT*Hispanic* 780.

5.2 Endogeneity

Our findings suggest that minorities and women gain disproportionately from using Autobytel.com. However, it is unlikely that these minorities and women are "average." If they share some unobserved characteristic that makes them use Autobytel.com but that also affects price, then our estimates do not reflect the causal effect of the Internet referral service. Notice, however, that regardless of whether the coefficient on ABT is driven by selection or causation, for the interaction of Autobytel.com and race to be significantly different from zero due to selection, there must be an *additional* selection effect operating for specific races. This would be the case if, for example, disadvantaged minorities who manage to locate and use Autobytel.com are even more aggressive about price than are the non-minorities who choose to use Autobytel.com. We can test whether this is the case by analyzing whether the ABT*%*Black* effect is bigger at the bottom end of the price distribution than at the top. If black buyers who have managed to achieve a price in the 10th percentile of the price distribution are unusually aggressive (and therefore unusually likely to use Autobytel.com), the effect of ABT for this group should be high. In contrast, minority buyers in the 90th percentile of the price distribution should not be aggressive and therefore the effect of Autobytel.com should be smaller for them. A quantile regression with the variable ABT*%*Black*, ABT*%*Hispanic* and ABT does not show this pattern (see columns 3 and 4 in Table 5). The interaction coefficients at the 10th percentile are -.010 and -.008, respectively (p values= 0.07 and 0.25), while the same coefficients at the 90th percentile are -.022 and -.046, respectively (p values= 0.01), the reverse of what would be expected under this explanation. These results cast doubt on whether an additional selection effect is operating for minorities.

In the remainder of this section, first, we discuss whether our results could be driven by the way that we measure whether a consumer belongs to a minority group, and second, we econometrically control for a selection effect (Heckman 1979).

5.2.1 Measurement of "minority"

A possible explanation of the preceeding results is that our OLS estimates are driven by middle class minorities who live in white neighborhoods. These people are more likely to use Autobytel.com and are also likely to pay low prices because of their high socio-economic status. However, our measure of minority is not at the individual level; instead, we measure the proportion of minorities in a census block group. Thus, for example, a few middle-class black consumers who live in a heavily white census block will show up as "almost white" in our data and cannot therefore be driving the minority results. In fact we have the reverse problem; consumers that we classify as "minority" may not be minorities. If white residents of a heavily minority block have some unobservable individual characteristic that leads them to use the Internet (for example, higher education), they will have a higher propensity to use Autobytel.com and pay lower prices. We may thus be wrongly interpreting the effect of "being white" in a minority neighborhood as "using Autobytel.com" in a minority neighborhood.

To see whether high-education whites in minority neighborhoods could be driving our finding that minorities benefit disproportionately from using Autobytel.com, we obtained data on education by race at the census block level from 1990. A census block only contains about 100 people on average. We examine heavily minority blocks (in regions in our main dataset), to see if white residents of those tracts have higher educational levels than their black neighbors and might thus be more likely to use Autobytel.com.

On the contrary, we find that in tracts with a black population of greater than 50%, black residents are more highly educated than are their white neighbors. For the median such tract, the percentage of blacks with some college education (associate, bachelors, graduate) is 2% greater than is the same statistic for whites. This difference increases to 5% in the median tract with a white population of less than 25%. In addition, in the large majority of cases in which no member of a group has any college education, that group consists of whites living in heavily minority tracts. In both groups, the average percentage of residents with some college or more is 25-30%. This suggests that the non-minorities in the block-groups we focus on are not educationally advantaged relative to their neighbors and are unlikely to be the ones using Autobytel.com. This is consistent with demographers Denton and Massey who find that "middle-class blacks are forced to live in neighborhoods of much poorer quality than whites with similar class backgrounds."[12]

We can avoid the measurement problem altogether when we interact the indicator variable *Hispanic* and *Black* with Autobytel.com. We find that Autobytel.com eliminates 60% of the race premium for blacks and all of the race premium for Hispanics (see *Black*, *ABT*Black* and *Hispanic*, *ABT*Hispanic* in columns 5 and 6, Table 5). We conclude therefore that it is unlikely that we are misidentifying our minority consumers.

5.2.2 Selection Model

Consider the following set of equations where B is an individual specific characteristic that is unobserved and forms part of the error term.

$$ABT_i = \gamma Z_i + \alpha B_i + \mu_i = \gamma Z_i + \epsilon_{1\,i} \tag{1}$$

[12]Page 814 in Denton, Nancy and Douglas Massey (1988) "Residential Segregation of Blacks, Hispanics, and Asians by Socioeconomic Status and Generation," *Social Science Quarterly*, 69 (4) December 1988 pp.797-817

$$\ln(Price_i) = \phi ABT_i + \beta X_i + \delta B_i + \nu_i = \phi ABT_i + \beta X_i + \epsilon_{2\,i} \tag{2}$$

Suppose that B is a desire and ability to bargain. This desire leads the buyer to use Autobytel.com to strengthen her bargaining position, leading to positive alpha and a negative delta. The result is positively correlated error terms across the two equations. Since ABT is on the right hand side in equation 2, it will be correlated with this equation's error term. In this scenario the estimated coefficient on ABT will be more negative than the true coefficient. A female consumer, for example, who is very interested in collecting information will be more likely both to use Autobytel.com and to bargain for a lower price from the dealer. Consequently it would be incorrect to treat the lower price as having been caused by Autobytel.com.

To control for unobserved heterogeneity we develop a selection equation. We use the demographic variables to predict use of Autobytel.com. We identify the system with instruments that affect the underlying cost or benefit of using Autobytel.com but that are not correlated with the unobserved characteristic or price. Our first instrument comes from the CPS Internet and Computer Use Supplement (2000). Since familiarity with the Internet increases the chance of using Autobytel.com to shop for a car, we use the percentage of people that use the Internet at work in a city, *PctInternetWork*. Use of the Internet at work is determined by the employer, so it is plausibly exogenous. For example, secretaries may have easy access to the Internet and their bargaining skill may be low as compared to a construction worker in the same city. (We also use other variables from this survey as instruments in some specifications: percentage who have a computer, use the Internet, and have ever used the Internet.) Our second instrument is family size, which is correlated with having a personal computer in the house and/or wanting to shop for a car outside of normal business hours.We measure *FamilySize* at the census block group level.

Our next instrument varies at the zip code level. It is a count of all the Autobytel.com referrals to that zip code that are not in our dataset (because they do not match a purchase transaction) divided by the zip code's population. We expect there to be some idiosyncratic variation in who uses Autobytel.com and that it might spread by word of mouth to neighbors within the zip code. We are concerned that use of Autobytel.com in a zip code is positively related to low prices at the local Autobytel.com dealer. However, the correlation between usage and Autobytel.com residuals from the price equation is zero. We also recognize that consumers in the same zip code share many demographics that predict Autobytel.com use; we control for these in the selection equation. Our two remaining instruments are based on the number of observations in the data that belong to the same "car" as the focal observation. This measures the popularity of a combination of attributes for which the consumer has been searching. We include *NumberOfCars*, linearly, squared, and cubed in our selection equation. (We find that people are less likely to use Autobytel.com when they are seeking very rare or very common bundles of attributes.) This instrument varies by consumer, rather than on a geographic basis. We do not include *SameDealer* in the IV specifications as it is endogenous also and we have not yet found a good instrument for it. Since it has a very small effect on price, for now we assume its omission will not bias the other coefficients.

We estimate in a probit specification the use of Autobytel.com on our instruments and all demographics used in the price equation. The R^2 of about .06 is quite low, partly because we do not know which consumers used Autobytel.com's two largest competitors (Carpoint and Autoweb), so our dependent variable is undercounted. We use the predicted values from the probit as an additional instrument in a two stage least squares regression of price, with ABT and race*ABT as the endogenous variables. We also include the variable of interest, such as *PctBlack*, times the predicted probability of using Autobytel.com as an instrument. The instruments used in each specification are reported at the bottom of the table.

The results are reported in Table 6. The coefficient on ABT increases in magnitude to -1.4% from under -1%. This direction of movement is consistent with results we have obtained in prior work (though smaller, as one would expect with the interaction term included). The estimated coefficient on the interaction term with *PctBlk* is -.12, which seems unreasonably large. Since we believe that the correct interpretation of the marginal effect of race is to move from 0 to 100% minority in a census block group, the price saving to a black buyer from using Autobytel.com would be 12%. We find this same pattern - the coefficient is significant but too large - for the *PctHispanic* interaction. For some reason, the instruments are not doing a good job of pinning down results with demographic averages. We therefore turn to the indicator variable *Hispanic* and interact it with Autobytel.com. The results are reported in column two of the same table. The Autobytel.com coefficient is -1.3% (significant at p=.15), while the interaction is -3.5% and significant at the 5% level. The magnitude of these estimates is somewhat closer to what we expect to see, although still large given a Hispanic premium of .6%. The instruments are clearly introducing more variance into the estimates. In all specifications the instruments pass an exogeneity test described in Hausman (1983). The test statistic is $N * R^2$ from a regression of the IV errors on all the exogenous variables in the system. It is distributed χ^2 with K-1 degrees of freedom, where K is the number of instruments.

We also try approaching this problem by running instrumental variables with only one endogenous variable, ABT, on a subsample of the data such as *PctBlk*$>$ 75 or $Asian$=1. We find that there are too few observations to produce stable and significant coefficients. We also try the approach with households in low income and education neighborhoods. If the general principle is that people with poor education, poor access to information, and high search costs pay more for cars, we should be able to see the benefits of using the Internet for disadvantaged

Table 6: Selection results

	(1)	(2)	(3)	(4)
Dep. Variable ln(price)	IV	IV	IV, Income < \$28000	IV, % Less HighSch. > 25
ABT	-0.01387 (0.00917)	-0.01345 (0.00930)	-0.08847 (0.06972)	-0.17233 (0.05660)**
ABTFranchise	-0.00446 (0.00025)**	-0.00414 (0.00026)**	-0.00282 (0.00114)*	-0.00371 (0.00079)**
ABT*%Black	-0.12316 (0.03402)**			
ABT*Hispanic		-0.03548 (0.01816)		
%Black	0.01604 (0.00067)**	0.01415 (0.00050)**		
%Hispanic	0.00747 (0.00102)**			
Hispanic	0.00490 (0.00028)**	0.00580 (0.00039)**		
%Asian	0.00045 (0.00097)	0.00049 (0.00099)	-0.00835 (0.00400)*	-0.00117 (0.00321)
Asian	-0.00964 (0.00049)**	-0.00975 (0.00049)**	-0.01303 (0.00261)**	-0.01899 (0.00194)**
Female	0.00200 (0.00015)**	0.00198 (0.00016)**	0.00295 (0.00057)**	0.00195 (0.00050)**
CustomerAge	0.00004 (0.00001)**	0.00004 (0.00001)**	0.00004 (0.00003)	0.00004 (0.00002)
Age> 64	-0.00167 (0.00030)**	-0.00179 (0.00032)**	-0.00124 (0.00103)	-0.00306 (0.00104)**
AnyTrade	0.00300 (0.00016)**	0.00311 (0.00016)**	-0.00163 (0.00063)**	-0.00308 (0.00052)**
Competition	-0.00030 (0.00003)**	-0.00025 (0.00004)**	0.00011 (0.00014)	0.00052 (0.00012)**
Constant	10.01703 (0.00135)**	10.01498 (0.00141)**	9.92493 (0.01077)**	9.93322 (0.00493)**
Observations	625341	576076	60418	70378

* significant at 5%; ** significant at 1%
Standard errors in parentheses

† Unreported are *EndOfMonth, WeekEnd, VehicleCost, MedianHHIncome, (MedianHHInc.)^2, %Executives, %BlueCollar, MedianHouseVal., %House-Own., %CollegeGrad, %<HighSchool, %Professional, %Technicians,* car, month, region, and model recency fixed effects.

[1] Instruments for column one: references in the zip code, family size, number of cars linear and squared. Column two: references in the zip code, family size, number of cars linear, squared, and cubed, percent with Internet access at work. Column 3: references in the zip code in spline form, family size, number of cars linear, squared, and cubed, percent with Internet access at work, a computer, who use the Internet, who have ever used the Internet. Column 4: references in the zip code, family size, number of cars linear and squared.

groups other than racial minorities. We run our instrumental variables specification on two subsamples of the data: buyers from blocks with greater than 25% of residents have not finished high school, and buyers from blocks with average household income below \$28,000. These samples have more than 50,000 observations each. The IV coefficient on the interaction of *ABT* and *LowEd* (*LowInc*) is -17% (-9%). The income coefficient is not significant, but in both cases the instruments pass the exogeneity test described above. The results are similar to the *PctBlk* results; whatever is driving the population averages seems to also be in effect here.

One might think that uniform pricing by Autobytel.com salespeople was driving our results, since uniform pricing would, by definition, eliminate discrimination. However, we do not find uniform pricing for Autobytel.com sales.[13] We examine the largest selling model-dealer combinations who have both Autobytel.com and "street" sales, and plot the errors for each separately in Figures 4 and 5. Keep in mind that options on the cars vary as does the time of year, which may be creating some base level of dispersion. The first franchise shows approximately similar dispersion between the two channels, while the other three show noticeably less dispersion for Autobytel.com sales. We also calculate the standard deviation of dollar margin for the 30 largest model-dealer-quarter combinations. We find less variation for Autobytel.com sales in 22 out of 30 cases. This also holds for 9 out of the largest 10 model-dealer combinations.

6 Explanations based on statistical discrimination

We have focused on whether dealers are rationally responding to differences between individual consumers. Alternatively, they may be treating customers differently because they make statistical inferences based on group averages. The first is considered a legitimate artifact of the bargaining process—if women and racial minorities pay different prices than do white males, it is because they have different education, income and perhaps bargaining ability, not because dealers are discriminating on the basis of race and gender. However, if dealers treat customers differently because they make statistical inferences based on group averages, this is a form of "racial profiling." Ayres and Siegelman (1995) attribute the causes of their results to such statistical discrimination.

Without considering the effect of the Internet we have been able to explain a substantial portion of the minority price premium with differences between individual consumers. This raises the question whether the remaining price differential is due to statistical discrimination. We present three tests:

First, statistical discrimination may occur because dealers are less willing to engage in a

[13] We find one dealer selling Dodge Durangos who appears to be selling at a uniform price.

lengthy bargaining process with minority buyers. This is because dealers may be afraid that such shoppers will not be able to purchase the car due to poor credit. If so, dealers effectively "bargain harder" with minority buyers since they expect no gains from trade. The sale may in many cases be lost by the dealer. However, since we only observe transactions, not offers, those minorities that purchase a car should pay higher prices under this conjecture. To exclude consumers that may be affected by this argument, we restrict our sample to buyers who did not obtain financing from their dealer. Dealers typically ask consumers early in the sales process whether they require financing. So minority consumers that do not, should not cause the dealer to exert low effort due to perceived credit risk. While many buyers that turn down dealer financing undoubtedly take out a loan elsewhere, some pay cash. In either case, such a buyer should have greater than average financial savvy. The estimated race and gender coefficients are only slightly smaller in this restricted sample (compare column 2 in Table 2 with column 5 in Table 4). We thus find no evidence that minorities pay a higher price because dealers may be less willing to engage them in a bargaining process due to credit risk.

Second, we use our previous results for black and Hispanic buyers. Recall that the dealership that received a purchase request from Autobytel.com knows only the name, address, and contact information of the potential customer. To identify whether dealers practice statistical discrimination, we capitalize on an important difference between the two largest minority groups in the United States. African-Americans can normally *not* be identified on the basis of name, but Hispanics often can be. Consequently, if dealers practice statistical discrimination, and since the Internet grants African-Americans more anonymity than Hispanic buyers, then African-Americans should benefit more than Hispanics from using Autobytel.com. Instead, we find that, if anything, Hispanics benefit more from using Autobytel.com.

Finally, we examine how the price a consumer pays for her car varies with a dealer's assumptions about his clientele. A dealer's assessment of how informed a consumer is, may be generated from the average educational level of consumers who patronize that dealership. We take buyers from census blocks with high levels of highschool dropouts. We measure how much those consumers pay at dealerships that serve many consumers from high dropout areas versus dealerships that usually serve highschool graduates. In previous results we found that buyers from less educated census blocks paid more. However, these buyers pay relatively higher prices at dealerships who usually serve uneducated people and relatively lower prices at dealerships that serve educated people. This is consistent with dealers who have a preconception about how well informed the average consumer is, and charge accordingly. Note, however, that what may draw a person from a high-dropout census block to a dealership that sells to educated people is — education.

Although we cannot prove that there is no statistical race discrimination when an individual buys a new car, our limited tests show no evidence of it. We do find weak evidence of statistical discrimination with respect to education.[14] When looking at the effect of education on prices, we find suggestive evidence that dealers may assume customers have the average educational level of the dealer's clientele. This type of behavior may be why less educated people can benefit from using Internet referral services such as Autobytel.com.

7 Concluding remarks

We have shown that pricing of new cars to off-line consumers strongly depends on individual characteristics of car buyers, in particular income, education, and search costs. Using data from J.D. Power and Associates on more than 700,000 new car purchases in 1999, we find a minority race premium of 2.0% to 2.3% when we do not control for any demographics, 1.1% to 1.5% when we control for neighborhood characteristics, and .7% when we (imperfectly) control for search costs. Our results are different from those in the previous literature, which finds either no role or conflicting results on the effect of demographics.

Our main finding is that the Internet eliminates most variation in new car prices that is due to individual characteristics associated with race and ethnicity: African-American on-line buyers who use the Internet Referral Service we study, Autobytel.com, pay almost the same prices as do whites; Hispanics pay less than whites. These finding suggests that dealerships are conditioning to a lesser extent on individual characteristics for on-line than off-line consumers. There are two possible explanations for this. First, dealerships may have less information about a consumer because mediating the interaction through the Internet removes important cues that salespeople can use to determine a consumer's willingness to pay. We would expect this if dealerships rely on salespeoples' experience in interacting with consumers to price discriminate. Since the Internet hides some of the information normally available to salespeoples, price discrimination is likely to be less pronounced.

However, there is an alternative explanation for why dealerships condition less on individual characteristics on-line: dealerships may have access to as much or more information about on-line than off-line consumers but choose not to use it. Recall that a dealership knows the name and address of a consumers before replying with a price offer. Hence, the dealer could look up the average demographics of the consumers' zip code or purchase individual-level data of the type normally used by direct marketers. The fact that dealers seem not to be conditioning

[14] One should keep in mind, however, that evidence of statistical discrimination can be very hard to observe. As Hylton and Rougeau (1996) write "If race is a relatively good proxy for the information the statistical discriminator does not collect, then the more information an empirical researcher collects in order to test for racial discrimination, the less evidence there will be of discrimination" (p.252).

on such information indicates that dealers may be reluctant to institute a formal "process" (referral arrives, sales person accesses database, etc.) by which they price discriminate. This could be for fear of litigation, negative publicity, or because it might not be customary for dealers to rely on more than the judgement of salespeople to set prices.

In addition, our finding could be partly due to the easier access that on-line consumers have to pricing and technical information. This may have resulted in more standardized price expectations and thus bargaining outcomes. Also, Autobytel.com's dealer training and suggested volume-based compensation may have contributed to less price discrimination.

We conclude that the Internet seems to benefit disproportionately those who lack information or who have personal characteristics that put them at a disadvantage in negotiating. These results suggest an additional aspect of the "Digital Divide": not only are disadvantaged minorities less likely to use a computer, but they are also the group that would most benefit from it.

References

Admati, A. R., and M. Perry (1987): "Strategic Delay in Bargaining," *Review of Economic Studies*, 54(3), 345–364.

Ayres, I. (2002): *Pervasive Prejudice?* University of Chicago Press, Chicago, IL.

Ayres, I., and P. Siegelman (1995): "Race and Gender Discrimination in Bargaining for a New Car," *American Economic Review*, 85(3), 304–321.

Goldberg, P. K. (1996): "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey," *Journal-of-Political-Economy*, 104(3), 622–654.

Hausman, J. (1983): "Specification and Estimation of Simultaneous Equation Models," in *Handbook of Econometrics*, ed. by Z. Griliches, and M. Intriligator, p. p. 433. North Holland, Amsterdam.

Heckman, J. J. (1979): "Sample selection bias as a specification error," *Econometrica*, 47(1), 153–161.

Hylton, K. N., and V. D. Rougeau (1996): "Lending Discrimination: Economic Theory, Econometric Evidence, and the Community Reinvestment Act," *The Georgetown Law Journal*, 85(2), 237–294.

J.D. Power and Associates (2000a): "Dealer Satisfaction with Online Buying Services," Management report, Agoura Hills, California.

——— (2000b): "New Autoshopper.com Study," Management report, Agoura Hills, California.

Mannering, F., and C. Winston (1991): "Brand Loyalty and the Decline of American Automobile Firms," *Brookings Papers on Economic Activity, Microeconomics*, pp. 67–103.

Salop, S., and J. Stiglitz (1977): "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion," *Review of Economic Studies*, 44(3), 493–507.

Scott Morton, F., F. Zettelmeyer, and J. Silva-Risso (2001): "Internet Car Retailing," *Journal of Industrial Economics, forthcoming.*

Appendix

Table 7: Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
ABT	671,468	0.03	0.17	0	1
ABTFranchise	671,468	0.24	0.43	0	1
SameDealer	671,468	0.01	0.09	0	1
Price	671,468	23,367	8,103	5957	100190
%Black	671,468	5.95	14.49	0	100
%Hispanic	671,468	8.25	10.27	0	55.33
%Asian	671,468	4.93	7.94	0	100
Hispanic	671,468	0.08	0.27	0	1
Asian	671,468	0.02	0.14	0	1
Female	671,468	0.36	0.48	0	1
CustomerAge	671,468	43.90	14.13	16	100
Age> 64	671,468	0.09	0.29	0	1
MedianHHIncome	671,468	56,597	24,905	10403	150000
%CollegeGrad	671,468	30.95	17.71	0	100
%<HighSchool	671,468	12.47	10.54	0	100
%HouseOwn.	671,468	72.99	22.38	0.14	100
%Professional	671,468	16.42	8.42	0	100
%Executives	671,468	17.39	8.06	0	100
%BlueCollar	671,468	26.27	14.99	0	100
%Technicians	671,468	2.99	1.97	0	100
MedianHouseVal.	671,468	164,642	99,728	7500	500000
EndOfMonth	671,468	0.22	0.42	0	1
Weekend	671,468	0.23	0.42	0	1
VehicleCost	671,468	0.0004	0.06	-0.64	0.73
AnyTrade	671,468	0.40	0.49	0	1
Competition	671,468	2.98	2.28	0	23
ModelMonth5-13	671,468	0.73	0.44	0	1
ModelMonth14+	671,468	0.11	0.32	0	1
FamilySize	671,468	2.99	0.55	1.5	6
%InternetAtWork	615,899	0.15	0.05	0	0.41
#ofCarsSold	671,468	2,701	2,262	300	12063
%ReferralsInZip	625,722	1.22	8.13	0.004	1700

Figure 1: Distribution of percent African-American in Census block group for new car buyers

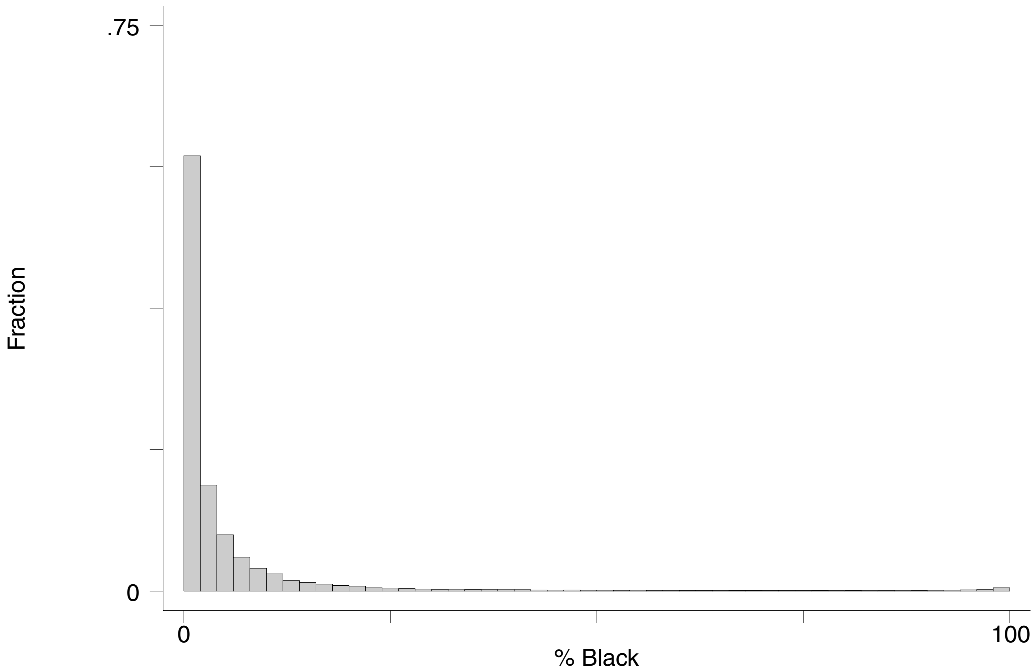

Figure 2: Distribution of percent Hispanic in Census block group for new car buyers

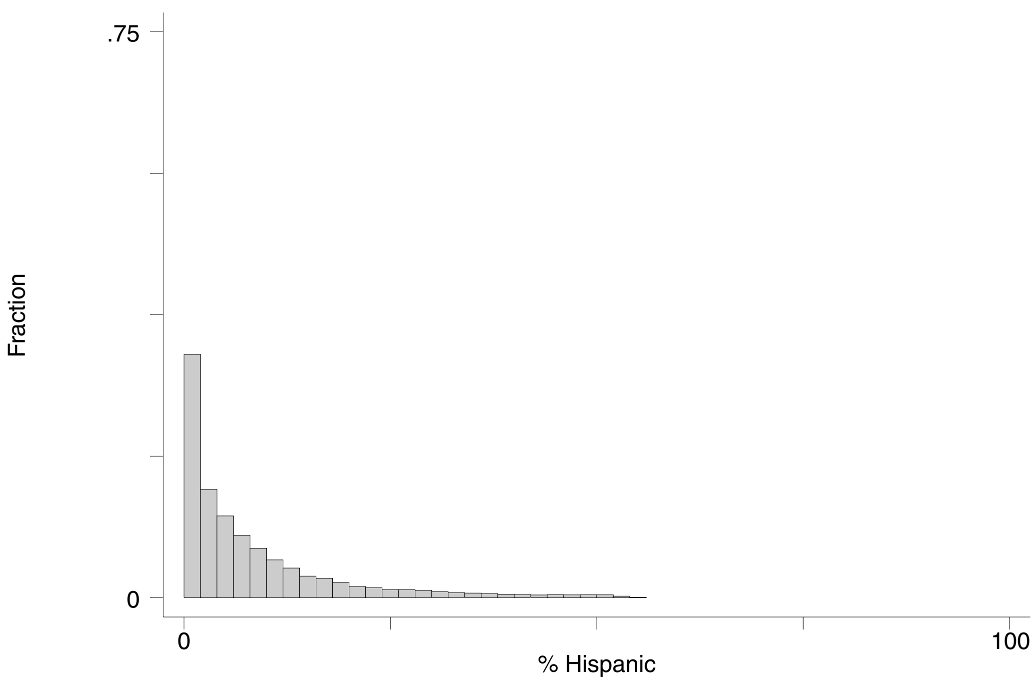

Figure 3: Distribution of percent Asian in Census block group for new car buyers

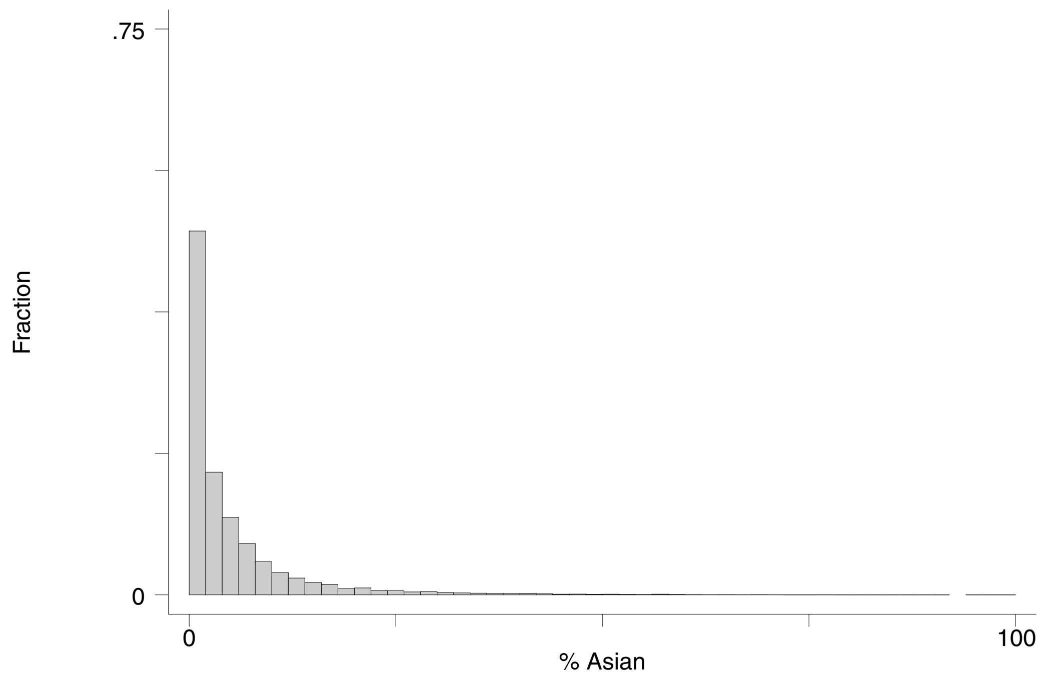

Figure 4: Distribution of percentage margins

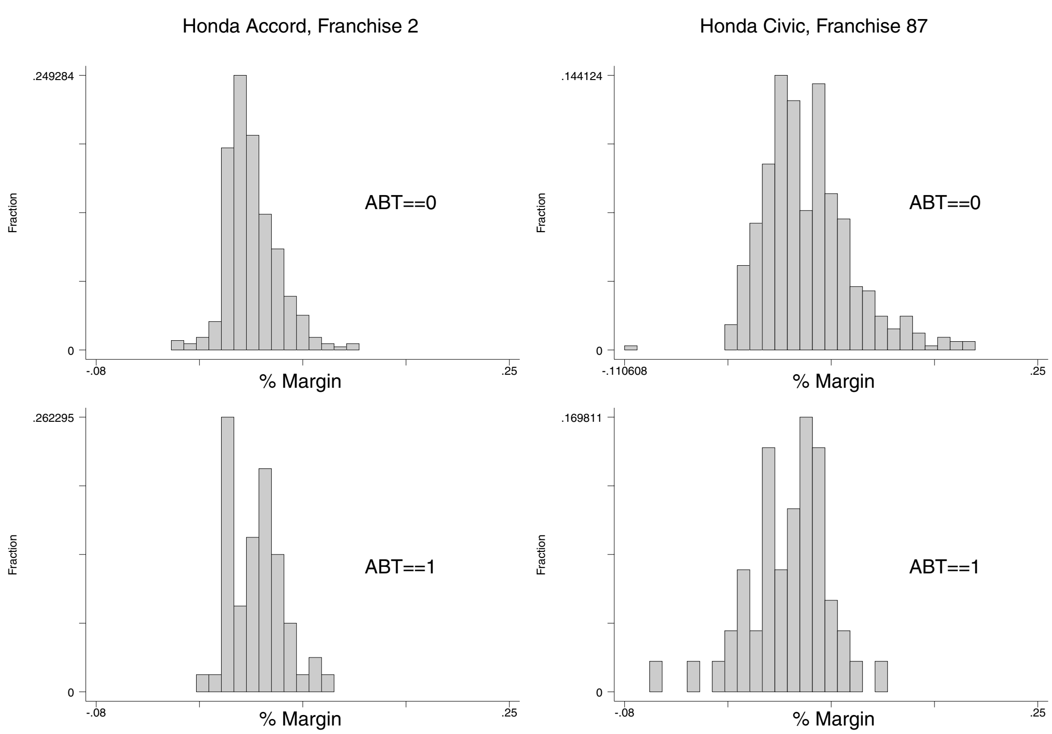

Figure 5: Distribution of percentage margins