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

January 2002

*We thank Ian Ayres, Judy Chevalier, Pinelopi Goldberg, 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. We gratefully acknowledge support from the Economics Program of the National Science Foundation, Grant #: SES-0111885.

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 dealers' difficulty in identifying consumer characteristics on the Internet and consumers' ease in finding information affects equilibrium prices in car retailing. Using a large dataset of transaction prices for new automobiles, the first part of the paper analyzes the relationship between car prices and demographics. We find that offline African-American and Hispanic consumers pay approximately 2% more than other consumers, however, we can explain 65% of this price premium with differences in income, education, and search costs; we find no evidence of statistical race discrimination. The second part of the paper turns to the role of the Internet. Online minority buyers who use the Internet Referral Service we study, Autobytel.com, pay nearly the same prices as do whites, irrespective of their income, education, and search costs. 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). In this paper, we analyze whether the increased difficulty in accurately assessing a consumer's willingness to pay on the Internet and consumers' ease in finding information affects race and gender discrimination in car retailing—a large industry in which prices are negotiated.

We show first, that disadvantaged minorities pay 2.0 - 2.3% more for their cars than white consumers. Most of this minority premium can be explained with differences in non-racial demographics and search costs between minority groups and whites. Second, we show that the Internet eliminates most of the offline minority premium in car buying.

Combined, these results clarify the nature of race discrimination in car buying. Academics and policy makers have been concerned with whether price discrimination in car buying has a "disparate impact" on minorities and whites or are evidence of a "disparate treatment" of minorities and whites. "Disparate impact" refers to the rational response by dealers to differences among individual consumers—if racial minorities pay different prices than do white males, it is because they have different education, income and perhaps bargaining ability. Alternatively, dealers may be treating customers differently ("disparate treatment") because they make statistical inferences based on group averages. This is a form of "racial profiling."

The combination of our results suggest that the race premium results from disparate impact, not disparate treatment. We can come to this conclusion because the Internet referral service we study, Autobytel.com, passes on the names and addresses of potential customers to its contract dealers. Hence, while the Internet removes important cues that salespeople can use to determine a consumer's willingness to pay (clothing, jewelry, body language, etc.), a dealer can easily infer the racial or ethnic background of an Internet consumer (thought a combination of name, address, and a phone interaction). Hence, the Internet eliminates the offline race premium despite the fact that the dealer is likely to know the minority status of Internet consumers. This suggests that dealers are not conditioning car prices on race, i.e. that there is no disparate treatment between minorities and whites. Instead, the revelation of personal characteristics, and the heavy premium placed on information during the bargaining process may have a disparate impact on minority consumers, who are less well educated on average.

Corresponding to our two primary findings, the paper is divided into two parts. To test whether the Internet's obfuscation of consumer characteristics affects equilibrium prices, we must first establish that *offline* negotiations result in differing car prices depending on individual consumer characteristics. We spend the first half of the paper just analyzing the relationship between car prices and demographics. We are particularly interested in whether characteristics exhibited by consumers of different races and genders can explain variation in new car prices. Since the two major papers in this literature come to different conclusions, we analyze this question in detail before investigating the effects of the Internet. 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 addition, Ayres (2002) finds $400 and $500 African-American premia in a smaller study of transaction prices at one dealer. 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

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

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 2.1% more (almost $500 for the average car) than do white buyers for identical cars. Since the average difference between price and invoice on a vehicle in our sample is $1700, this represents an almost 30% higher markup. 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 0.8% for blacks and 0.6% for Hispanics. Thus, about 65% of the minority price premium can be attributed to observable individual differences in income, education, and search costs. We find a small price premium for women, .43% ($100).

The second part of the paper turns to the role of the Internet. We test for the effect of the Internet on price discrimination with data on whether consumers used the Internet referral service Autobytel.com in purchasing their car. We have found in previous research that Autobytel.com users pay lower than average prices for new cars (Zettelmeyer, Scott Morton, and Silva-Risso 2001). Autobytel.com allows consumers to request a price quote from an affiliated dealer without engaging in personal interaction. Consequently, dealers are exposed to fewer cues that signal a consumer's willingness to pay. They can (imperfectly) identify, however, members of certain ethnic and racial groups by their names and addresses. 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 non-racial demographics and search costs. This has not been previously established, to our knowledge. Secondly, our large dataset allows us to estimate a race premium with good precision; in particular, we establish that disadvantaged minorities pay 2.0 - 2.3% more for their cars than white consumers. Finally, we conclude that the Internet eliminates most variation in new car prices that is due to race and ethnicity. Our results suggest that a car market where prices are bargained (as opposed to posted and fixed) has a disproportionately negative impact on minority buyers, but that the negative impact is not necessarily due to different treatment of different races by dealers. Instead, our evidence points to the role of information and search costs as determinants of prices; we find that any group that is less educated or less able to search pays higher offline prices. Members of these groups are also those who disproportionately benefit from using the Internet. 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 benefit most 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. Section 4 contains the first set of results, establishing that offline car prices depend on individual consumer characteristics. Section 5 contains the second set of results, showing that the Internet reduces most of the difference in pricing between racial 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.[2] Communication may occur by email or telephone. In this 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'

[2] 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.

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.[3] 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, body language, vehicle, or accent as signals of her reservation value or bargaining ability. The last two - vehicle and accent - may be revealed to the saleperson during the course of the negotiation if it takes place over the phone and includes discussion of a trade-in. However, the dealer clearly has less information about the buyer than he would have if the buyer were in the dealer's showroom.

The preceding arguments suggest that dealers should be less likely to price discriminate for online than offline 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. Note however, that this could be done by "offline" dealers also, provided they have a few minutes away from the customer.

[3]According to J.D. Power and Associates (2000a) some dealers follow these behavioral recommendations closely, while others do not.

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 these 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). This number may be positive or negative depending on whether the dealer is over- or under-paying for the trade-in. We adjust the price of the new car for this amount to account for the possibility, for example, that a dealer may offer a consumer a high price for the new car so he can artificially subsidize the trade-in. (This pattern is the most common in our data.)

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 one indicates a female customer. The problem with this variable is that many cars are officially bought by two people and the dataset only records the first name on the registration. If the owners are "Mary and John Doe," our dataset records Mary as having purchased the car and the *Female* variable is one. However, John may have been the one who actually bargained for the car. While we cannot fix this problem, we will later compare subsamples of the data to better measure the true impact of gender. 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. 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.

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.

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 forwarded slightly over over 2 million referrals 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 (*Autobytel*) 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), and (2) an indicator for Autobytel.com franchise dealer (*AutobytelFranchise*) 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.

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 biases our test against our hypotheses since we will be comparing a group that used Autobytel.com to a group that may include users of competing services.

3.4 Controls

We use car fixed effects to control very precisely for the cost of the car. 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 an observation's invoice price (its *VehicleCost*) from

[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%).

the average *VehicleCost* of that type of car in the dataset.[5] We call this variable *DVehCost*. For example, if the car has a sunroof and we don't observe it, the car's invoice price will be higher than average. Our *DVehCost* variable will be positive in this case because the focal car is more expensive. In the regression this variable will have a positive coefficient close to one because it is measuring the part of cost we cannot control for with our fixed effects.

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.[6]

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:

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

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.

[5] 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.

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

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 (see column 1 of Table 2): higher income lowers car prices until the average block income reaches $80,000, at which point increases in income increase price. Coming from a block with a higher percentage of people who have gone to college and higher house values lowers prices. Home ownership, a proxy for creditworthiness, also lowers prices. The probability of being a blue-collar worker or an executive are insignificant. Coming from a block with a higher percentage of people who are professionals increases car prices. So does a higher probability of not finishing high school.

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. 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% 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 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% (not reported). 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.[7]

We are concerned with the interpretation of the *Female* coefficient because in cases of joint car ownership by a couple the variable may not accurately measure whether a woman has bargained over the price of the car. To estimate the effect of joint ownerships on the *Female* coefficient we compare the female premium for cars that are predominantly purchased by couples ("Minivans") with the premium for cars that are likely to be purchased by women alone ("Compact Entry" and "Compact Sporty"). [8] Our conjecture is that a male is more likely to have participated in the purchase of a "Minivan" than a "Compact Entry" or "Compact Sporty" car that is listed in our data as a female purchase. Comparing the *Female* coefficient in the two columns of Table 3, we see the conjectured result: while gender plays no role in the price of minivans, women pay 0.43% more for small cars, or $98 for the average car. The likely smaller measurement error in the small car segments leads us to prefer this estimate of the gender premium to the sample-wide one. The estimate of 0.43% is still likely to be conservative since women are frequently advised to bring a man along to negotiations with car dealers.

Our estimate of a minority premium between $350 and $500 is much smaller than those of Ayres and Siegelman (1995), whose testers find unexplained minority premia of $410 (female) and $1100 (male). They are closer to, but still smaller than, the Ayres (2002) transaction results. We investigate whether our data show the same relationship between minority female

[7] We do not create a new Hispanic indicator because the maximum *PctHispanic* is only 55% and thus too low to create an equivalent variable.

[8] XXX JORGE source of demographics

and minority male prices. Column 4 of Table 2 shows that the interaction of *PctBlack* with *Female* is positive, insignificant, and additionally, only 0.13%, or about $30 on the average car. The coefficient on *PctBlack* remains fairly steady at 1.3%. Hispanic women appear to pay -.15% less than Hispanic men.[9] The estimates of the female interaction coefficients continue to be close to zero and insignificant if we use only the sample of small cars (unreported).

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 reported in the last two columns of Table 2.

Neither the black, Hispanic, nor gender coefficients change when the sample is restricted to buyers from census blocks with average incomes above the mean of $57,000. We find very similar results when we restrict the sample to buyers that reside in blocks where 32% or more of residents have a college education: only the *PctHispanic* coefficient declines. These results indicate that our basic finding is not driven by one end of the income or education distribution.

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 4). We find that a buyer who has a probability one of being black versus a buyer who 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. These results are consistent with the findings of Goldberg (1996) that the variance in minority prices is greater than that of white prices. However, in contrast to her, we find that blacks and Hispanics pay on average more than do whites, even at the low end of the price distribution. Our results on gender are not consistent with the findings of Goldberg (1996). Females pay .21% more in the 10% regression and .28% more in the 90% quantile regression. This indicates that the variance of female's reservation price distributions is not smaller than that for males.

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

[9]Note that we cannot test separately for "redlining" since our race data are neighborhood data and we are thus already measuring price differences based on where people live.

buyers pay a 1.1% premium. Hispanic women pay a little less than the Hispanic average. We also estimate that women pay about .5% more than do 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

We have focused on whether dealers are rationally responding to differences among individual consumers. Alternatively, they may be treating customers differently because they make statistical inferences based on group averages. The first is an 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.[10] However, if dealers treat customers differently because they make statistical inferences based on group averages, this is a form of "racial profiling."[11] Ayres and Siegelman (1995) attribute the causes of their results to such statistical discrimination. Our data does not allow us to distinguish between discriminatory behavior by the dealer and different behavior by customer groups because, unlike Ayres and Siegelman (1995) we do not see the dealer's offer.[12] However, we can observe the effect of race on price as we vary the specification and, in this way, indirectly test for some explanations for the race premium.

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 increase 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 5). This contrasts with 1.5% and 1.1%, respectively, when income, education, occupation, and wealth are controlled for.[13] 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.

[10] In legal terminology this is "disparate impact" as opposed to "disparate treatment."

[11] "Disparate treatment."

[12] Thus at no point in the paper will we be able to conclude that we have found disparate treatment.

[13] These results illustrate 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).

Minorities may not be able to finalize the transaction: Dealers may be less willing to engage in a lengthy bargaining process with minority buyers if they are 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.[14] Since dealers typically ask consumers early in the sales process whether they require financing, 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 buyers 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 5). 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.

Minorities may shop where dealers expect customers to be uninformed: We also 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 that 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. We cannot pursue this avenue of inquiry any further without individual, rather than neighborhood-level, data.

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

[14] We do not use information on financing elsewhere in the paper for two reasons. First, it is a large topic that deserves throrough treatment in a separate paper. Secondly, preliminary correlations suggest that financing profits are not cross-subsidizing car prices. Rather the two tend to move together. Thus we feel it is reasonable to omit an analysis of the price of financing from this paper.

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 nearly identical race and gender coefficients in this specification (compare column 2 in Table 5 with column 2 in Table 2). Hence, the minority premium is not due to purchases at higher cost dealerships.[15]

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 are 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 our minority race variables and market structure is positive (see column 3 of Table 5), which is counter to theory. It seems we are picking up the high prices paid by blacks and Hispanics in central urban neighborhoods where there are many dealerships within 10 miles of a buyer.

The next specification in the table checks this conjecture by including a population density measure and its interaction with race. We find that the main race coefficients decline slightly, but the interaction coefficients are positive and significant: a minority consumer living in a more urban area pays a higher price. Someone in a block with 15% black (Hispanic) residents would pay an additional 1.1% (2.8%) if population density rose by one standard deviation. In summary, we find no evidence that minorities or women pay more because they have an aversion to bargaining.

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 and Winston 1991). Collecting basic information about features, prices, and availability for a 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* or *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 5 of Table 5 show that

[15] 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.

this is indeed the case. Notice that we include an indicator variable for transactions with a trade-in in all specifications in the paper. Consumers of all races who trade in a vehicle pay a small premium for that convenience.[16] 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 .8% for black and 0.6% for Hispanic consumers who have traded in a vehicle.[17] This suggests that higher search costs when buying a vehicle may be responsible for a large part of the price premium paid by minorities. Our finding provides interesting evidence about the bargaining strategies being used by dealers: searching is required to get a good price. Those who cannot search pay a high price, those who can pay a lower price. Our specification assumes the majority of white consumers can search and uses the trade-in to measure which minority consumers can search. A system that requires price searching of this type could be said to have a disparate impact if the minority consumers are disproportionately the ones who cannot search.

In conclusion, we find that the minority premium of 2.0% or 2.3% (when no demographics are in the regression) declines to .8% or .6% when we control for differences between groups of consumers. 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%.[18] 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 statistic for African-American and Asian blocks is 1.7% and 4.1%, respectively. Census blocks where the sum of black and Hispanic residents exceeds 75% of the population have only a 1% use of Autobytel.com.

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

[17] 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 (unreported).

[18] The overall Autobytel.com use is closer to 6% before we drop consumers who buy a different car.

5.1 Result

We begin with a specification that includes an indicator variable *Autobytel* that is one if the car buyer submitted a purchase request using Autobytel.com. We also include an indicator variable *AutobytelFranchise* for Autobytel.com network dealers. The specification is a follows:

$$\ln\left(Price_i\right) = \alpha_1 Autobytel_i + \alpha_2 AutobytelFranchise_i + \gamma D_i + \beta X_i + \epsilon_i$$

Column 1 of Table 6 shows that Autobytel.com users pay about 1.2% less than do other customers. The first effect is the main Autobytel.com discount of 0.9%. Users are also sent to an Autobytel.com dealer for an additional .5% discount (which they would get by chance with 1/3 probability since Autobytel.com dealers sell 1/3 of all cars), resulting in additional savings of .32%.

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, obfuscation of consumer characteristics, 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 6 shows that the coefficient on *PctBlack***Autobytel*, is -1.2% and significant. This substantially offsets the *PctBlack* coefficient of +1.5%. The *Autobytel* 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 do 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 $Autobytel * PctHispanic$ has a coefficient of -2.0%, which more than makes up for the premium of 0.75% we estimate for *PctHispanic*. The interaction coefficient is large because the variable *PctHispanic* is correlated with education, income, and home ownership variables that also have "Autobytel.com" effects. While these are included separately in the specification, their Autobytel.com effects seem to be partially picked up by the $Autobytel{*}PctHispanic$ measure. This can be seen by repeating the interaction specification with no demographics other than race and gender. Column 3 shows that Hispanics who use Autobytel.com exactly eliminate the offline Hispanic premium.[19]

Finally, we estimate the interaction of Autobytel.com with minorities and women while controlling for all demographics as well as franchise fixed effects (see column 4 in Table 6). We see again that Autobytel.com use offsets essentially all minority premia and half the gender premium.

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. 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 be classified as high probability white in our data and cannot therefore be driving the minority results. In fact we have the reverse problem; consumers 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 (not block group) level from 1990. A census block contains only about 100 people on average. We examine heavily minority blocks (in regions in our

[19] Buyers who are "disadvantaged" according to other metrics also pay lower prices when using Autobytel.com. We interact *%CollegeGraduates*, *%<Highschool*, and *MedianHHIncome* with *Autobytel*. As expected, *%CollegeGraduates* has a positive and High school dropouts and low income buyers have negative *Autobytel* interaction coefficients.

main dataset), to see if white residents of those tracts have higher educational levels than do 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 not more likely to be 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."[20]

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 *Hispanic*, *Autobytel*Hispanic* in columns 2 through 4, and *Black*, *Autobytel*Black* in column 5, Table 6). We conclude therefore that it is unlikely that we misidentify minority consumers.

5.2.2 Selection Effect by Race

Notice that regardless of whether the coefficient on *Autobytel* 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. A race-based selection effect might occur, for example, if 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 now turn to a formal IV model to handle this potential problem.

[20]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

5.2.3 Instrumental variables model

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

$$\begin{aligned} Autobytel_i &= \gamma Z_i + \alpha B_i + \mu_i = \gamma Z_i + \epsilon_{1\,i} & (1) \\ \ln(Price_i) &= \phi Autobytel_i + \beta X_i + \delta B_i + \nu_i = \phi Autobytel_i + \beta X_i + \epsilon_{2\,i} & (2) \end{aligned}$$

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 δ. Since B is unobserved, *Autobytel* will be correlated with equation 2's error term. In this scenario the estimated coefficient on *Autobytel* will be negatively biased relative to 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. Our second instrument is family size, which may be correlated with personal computers in the home and/or wanting to shop for a car outside of normal business hours. We measure *FamilySize* at the census block group level.

Our third 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 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.

The intuition behind the identification strategy is that different consumers will have different exogenous tendencies to use the Internet based on the car they are looking for, the mix of industries in their city, word of mouth in their zipcode, and the prevalence of kids in their neighborhood. These measures of the costs and benefits of using the Internet for a particular consumer are assumed (and tested) to not be correlated with price. Two minority consumers with the same demographics who live in the same city but differ in the car they are searching for or in their census block group will have different predicted probabilities of using Autobytel.com. This is then related to the transaction price to find an effect of the Internet.[21]

We estimate in a probit specification the use of Autobytel.com on our instruments and all demographics used in the price equation. The pseudo 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 estimate an additional probit predicting minority use of Autobytel.com with the same explanatory variables. We use the predicted values from the probits as additional instruments in a two stage least squares regression of price, with *Autobytel* and race∗*Autobytel* as the endogenous variables. We also include the variable of interest, such as *Hispanic*, times the predicted probability of using Autobytel.com as an instrument.

We begin with interaction coefficients on the indicator variables for the two disadvantaged minority groups in our sample, Hispanics and blacks. The results are reported in Table 7 (the instruments used in each specification are reported at the bottom of the table). The coefficient on *Autobytel* increases in magnitude in all the specifications. This direction of movement is consistent with results we have obtained in prior work.[22] The first column contains the indicator variable *Hispanic* interacted with Autobytel.com. The Autobytel.com coefficient is -1.2% (p=.15), while the coefficient of the interaction term *Autobytel*∗*Hispanic* is -2.9% and significant at the 5% level. The interaction coefficient is probably larger than the main Hispanic effect for the same reason we saw earlier, namely that *Hispanic* is correlated with other demographics.

[21] We cannot estimate Internet use variation within a block group for people buying the same car. However, our cars are so specific and we control for costs so carefully that we are not concerned that our estimate relies on across-car variation.

[22] See Zettelmeyer, Scott Morton, and Silva-Risso (2001)

Next we estimate the interaction of *Autobytel* and *Black* in the restricted sample with more than 75% or less than 2% blacks. The interaction coefficient is -33%, which is unreasonably large. The large standard error is due to the small number (98) of Autobytel.com users from heavily black neighborhoods. With a restricted sample and weak instruments, we cannot get reasonable coefficient estimates for the black interaction.

When we run an instrumental variables procedure on an interaction between Autobytel.com and a census block percentage such as *PctBlack* or *PctHispanic*, we find coefficients that are also large and unstable, so we do not report these results. Our instruments do not do a good job of pinning down results with demographic averages.

We also try approaching the problem by limiting the sample to consumers with a high value of some characteristic and estimating a single *Autobytel* coefficient, rather than an interaction, for that subsample. We focus on two characteristics that affect car pricing but have more heterogeneity across census blocks: income and education. In columns 5 and 6 in Table 7, we restrict the sample to the bottom half of the income and education distribution, respectively. Our theory suggests that the Autobytel.com coefficient should increase in magnitude relative to the OLS estimates for two reasons; the IV controls for the unobserved reasons to use Autobytel.com and the sample choice increases the observable reasons to use Autobytel.com. The estimated coefficients are negative and significant, although larger than we would expect, at -5% and -7%. We use the estimated coefficients on only the census block measures from the OLS regression in column 1 of Table 6 to calculate a measure of the effect of demographic variables on car prices. We find the total impact of demographics ranges from negative two percent to positive one percent in our sample. We average the demographic effect within the top and bottom income quartiles and find that it changes by one percent between the two groups. If a demographic effect of this magnitude is added to the estimated IV ABT coefficient of approximately negative two, it would be reasonable to find ABT coefficient estimates for low income and low education samples of about negative three percent. This suggests that the true values of the ABT coefficients in the disadvantaged sample are probably contained in the 95% confidence intervals of our point estimates, but our instruments are not allowing for precise estimation.

While our instruments introduce substantial variance, the estimated Autobytel.com effect for the disadvantaged group in question is consistently negative and larger than the OLS estimate, indicating that the "aggressive bargainers use Autobytel.com" story is not driving our results. The instruments pass an exogeneity test described in Hausman (1983) in all specifications. 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.

5.3 Supply Side Pricing

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.[23] We do find less dispersion for Autobytel.com sales, which is likely contributing to less variation in prices by race.

We calculate the standard deviation of the dollar margin and the percentage margin for each dealer-model-quarter that has greater than 5 sales per period in each channel. We compare the standard deviation between "street" and Internet channels for the same dealer-model-quarter and find that the this difference has a negative mean; Internet sales have less dispersion. We examine the largest selling model-dealer combinations who have both Autobytel.com and "street" sales, and plot the errors for each separately in Figure 2. 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. The standard deviation of dollar margin for the 30 largest model-dealer-quarter combinations has less variation for Autobytel.com sales in 22 out of 30 cases. This is also true for 9 out of the largest 10 model-dealer combinations.

6 Concluding remarks

We have shown that pricing of new cars to offline 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 .6% to .8% 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: online buyers who use the Internet Referral Service we study, Autobytel.com, pay the same prices as do whites, irrespective of their income, education, and search costs. Our findings suggest that disadvantaged minorities have more to gain from using an online buying service than whites do. Consistent with this finding, a consumer survey in J.D. Power and Associates (2000b) shows that minority buyers who use an online buying service submit on average more purchase requests than white buyers

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

(1.42 versus 1.35, difference significant at 5% level).

Our results suggest that dealerships condition to a lesser extent on individual characteristics for online than offline 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 online: dealerships may have access to as much or more information about online than offline 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 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 online 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.*

Zettelmeyer, F., F. Scott Morton, and J. Silva-Risso (2001): "Does the Internet Lower New Car Prices?," mimeo, UC Berkeley, Berkeley, CA.

Appendix

Table 1: Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
Autobytel	671,468	0.03	0.17	0	1
AutobytelFranchise	671,468	0.24	0.43	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
MedianHouseValue	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
DVehCost	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

Table 2: Regressions for results section†

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Variable ln(price)	Full Sample	Full Sample	> 75% or < 2% Black	> 75% or < 2% Black	Full Sample	> $57,000 Income	> 32% College
%Black	0.01457 (0.00054)**	0.01469 (0.00054)**			0.01467 (0.00054)**	0.01478 (0.00130)**	0.01664 (0.00135)**
%Hispanic	0.01105 (0.00102)**	0.00671 (0.00104)**	0.00292 (0.00136)*	0.00291 (0.00136)*	0.00670 (0.00104)**	0.00770 (0.00202)**	0.00065 (0.00209)
%Asian	-0.00390 (0.00096)**	-0.00096 (0.00098)	-0.00031 (0.00138)	-0.00031 (0.00138)	-0.00096 (0.00098)	0.00175 (0.00131)	0.00162 (0.00127)
Hispanic		0.50539 (0.02761)**	0.53896 (0.03767)**	0.53936 (0.03767)**	0.55760 (0.03339)**	0.49703 (0.04578)**	0.51710 (0.04813)**
Asian		-0.96564 (0.04341)**	-0.85375 (0.05799)**	-0.85368 (0.05799)**	-0.96576 (0.04341)**	-0.83471 (0.05446)**	-0.77833 (0.05388)**
%Black> 75			1.36567 (0.06193)**	1.29715 (0.08689)**			
Female	0.20610 (0.01391)**	0.20896 (0.01390)**	0.19364 (0.01789)**	0.18963 (0.01804)**	0.22107 (0.01431)**	0.18852 (0.01929)**	0.16345 (0.01961)**
Female ∗ %Black> 75				0.12694 (0.11611)			
Female ∗ Hispanic					-0.15039 (0.05431)**		
Customer Age	0.00449 (0.00063)**	0.00474 (0.00063)**	0.00303 (0.00081)**	0.00303 (0.00081)**	0.00475 (0.00063)**	0.00291 (0.00089)**	0.00594 (0.00092)**
Age> 64	-0.16841 (0.02953)**	-0.16796 (0.02952)**	-0.14502 (0.03665)**	-0.14465 (0.03665)**	-0.16797 (0.02952)**	-0.08025 (0.04194)	-0.14526 (0.04326)**
MedianHH Income	-0.00002 (1.39e-06)**	-0.00002 (1.39e-06)**	-0.00002 (1.71e-06)**	-0.00002 (1.71e-06)**	-0.00002 (1.39e-06)**	-0.00002 (3.38e-06)**	-0.00002 (2.18e-06)**
(Median HHInc.)^2	1.26e-10 (7.58e-12)**	1.25e-10 (7.57e-12)**	1.23e-10 (9.11e-12)**	1.23e-10 (9.11e-12)**	1.25e-10 (7.57e-12)**	1.21e-10 (1.60e-11)**	9.97e-11 (1.08e-11)**
%College Grad	-0.00305 (0.00095)**	-0.00325 (0.00095)**	-0.00109 (0.00119)	-0.00108 (0.00119)	-0.00324 (0.00095)**	-0.00038 (0.00138)	-0.00045 (0.00143)
%<High School	0.00394 (0.00128)**	0.00310 (0.00128)*	0.00329 (0.00167)*	0.00329 (0.00167)*	0.00310 (0.00128)*	0.00950 (0.00297)**	-0.00309 (0.00325)
%HouseOwn.	-0.00274 (0.00045)**	-0.00271 (0.00045)**	-0.00241 (0.00062)**	-0.00240 (0.00062)**	-0.00271 (0.00045)**	-0.00427 (0.00079)**	-0.00276 (0.00070)**
%Professional	0.00459 (0.00139)**	0.00472 (0.00139)**	0.00157 (0.00174)	0.00157 (0.00174)	0.00471 (0.00139)**	0.00574 (0.00187)**	0.00475 (0.00180)**
%Executives	-0.00013 (0.00147)	0.00008 (0.00146)	-0.00130 (0.00179)	-0.00130 (0.00179)	0.00007 (0.00146)	0.00133 (0.00200)	0.00120 (0.00198)
%BlueCollar	0.00018 (0.00102)	0.00024 (0.00102)	0.00082 (0.00125)	0.00083 (0.00125)	0.00023 (0.00102)	0.00179 (0.00187)	0.00178 (0.00199)
%Technicians	0.00460 (0.00347)	0.00421 (0.00347)	-0.00120 (0.00440)	-0.00120 (0.00440)	0.00420 (0.00347)	-0.01053 (0.00500)*	-0.00711 (0.00501)
MedianHouse Value	-2.73e-06 (1.28e-07)**	-2.58e-06 (1.28e-07)**	-2.38e-06 (1.60e-07)**	-2.38e-06 (1.60e-07)**	-2.58e-06 (1.28e-07)**	-2.06e-06 (1.72e-07)**	-1.40e-06 (1.59e-07)**
EndOfMonth	-0.34539 (0.01538)**	-0.34545 (0.01537)**	-0.35572 (0.01955)**	-0.35577 (0.01955)**	-0.34538 (0.01537)**	-0.32204 (0.02104)**	-0.33576 (0.02166)**
Weekend	0.11224 (0.01579)**	0.11079 (0.01577)**	0.10154 (0.02031)**	0.10160 (0.02031)**	0.11076 (0.01577)**	0.07238 (0.02181)**	0.04264 (0.02220)
DVeh-Cost	88.18283 (0.13374)**	88.15525 (0.13375)**	88.11184 (0.17123)**	88.11188 (0.17123)**	88.15550 (0.13375)**	88.00853 (0.19257)**	87.40996 (0.19463)**
Competition	-0.02174 (0.00351)**	-0.02222 (0.00351)**	-0.02060 (0.00434)**	-0.02059 (0.00434)**	-0.02221 (0.00351)**	-0.03943 (0.00472)**	-0.03751 (0.00508)**
AnyTrade	0.30869 (0.01377)**	0.30861 (0.01377)**	0.33353 (0.01767)**	0.33353 (0.01767)**	0.30853 (0.01377)**	0.43029 (0.01973)**	0.48460 (0.02013)**
Constant	1,001.7 (0.13088)**	1,001.7 (0.13081)**	1,003.9 (0.17713)**	1,003.9 (0.17713)**	1,001.7 (0.13082)**	1,009.40189 (0.27175)**	1,009.20522 (0.23718)**
Observations	650850	650850	386155	386155	650850	285231	276632
R^2	0.97	0.97	0.98	0.98	0.97	0.98	0.98

[*] 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: Female**%Black>75* 6134; column 5: Female*Hispanic: 18491.

Table 3: Female coefficients by segment†

	(1)	(2)
Dep. Variable ln(price)	Segments: Compact Entry and Sporty	Segment: Minivan
%Black	0.01246 (0.00651)	0.01685 (0.00205)**
%Hispanic	-0.02245 (0.01287)	0.00505 (0.00372)
%Asian	-0.01857 (0.01639)	0.00460 (0.00343)
Hispanic	0.80267 (0.32345)*	0.44631 (0.09679)**
Asian	1.46397 (1.04951)	-0.72971 (0.13057)**
Female	0.42559 (0.19073)*	0.02765 (0.04804)
CustomerAge	-0.00156 (0.00799)	0.00993 (0.00248)**
Age> 64	0.83626 (0.47895)	-0.29870 (0.10730)**
MedianHHIncome	-0.00004 (0.00002)	-0.00003 (0.00001)**
(MedianHHInc.)^2	2.81e-10 (1.50e-10)	1.96e-10 (2.83e-11)**
%CollegeGrad	-0.02172 (0.01448)	-0.00487 (0.00330)
%<HighSchool	0.02760 (0.01594)	0.00373 (0.00446)
%HouseOwn.	-0.00949 (0.00592)	0.00026 (0.00164)
%Professional	0.01669 (0.02116)	0.01211 (0.00488)*
%Executives	-0.01211 (0.02261)	-0.00639 (0.00510)
%BlueCollar	-0.02221 (0.01362)	-0.00119 (0.00350)
%Technicians	0.02038 (0.04528)	0.00258 (0.01191)
MedianHouseVal.	-6.10e-07 (2.22e-06)	-1.77e-06 (4.87e-07)**
EndOfMonth	-0.32131 (0.22640)	-0.30078 (0.05132)**
Weekend	0.36512 (0.23626)	0.01997 (0.05248)
DVehCost	89.31072 (1.54931)**	89.90300 (0.45892)**
Competition	-0.57803 (0.07208)**	0.00231 (0.01027)
AnyTrade	-0.00599 (0.20295)	0.28481 (0.04577)**
Constant	937.6 (1.38177)**	1,008.7 (0.47868)**
Observations	5335	57541
R^2	0.91	0.88

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

† Unreported are car, month, region, and model recency fixed effects

Table 4: Quantile regressions[†]

	(1)	(2)	(3)
Dep. Variable ln(price)	.1 Quantile	.9 Quantile	Median
%Black	0.007 (0.001)**	0.025 (0.001)**	0.012 (4.57e-04)**
%Hispanic	1.53e-04 (0.002)	0.010 (0.002)**	0.009 (0.001)**
Hispanic	0.259 (0.043)**	0.861 (0.054)**	0.441 (0.024)**
%Asian	-0.010 (0.002)**	0.014 (0.002)**	-0.002 (0.001)**
Asian	-0.591 (0.071)**	-0.01432 (0.093)**	-0.867 (0.046)**
Female	0.215 (0.023)**	0.277 (0.029)**	0.144 (0.013)**
CustomerAge	0.005 (0.001)**	0.006 (0.001)**	0.004 (0.001)**
Age> 64	-0.184 (0.049)**	-0.129 (0.061)*	-0.177 (0.028)**
MedianHHIncome	-1.53e-05 (2.29e-06)**	-2.04e-05 (2.91e-06)**	-1.45e-05 (1.34e-06)**
$(\text{MedianHHInc.})^2$	1.24e-10 (1.26e-11)**	1.39e-10 (1.62e-11)**	1.09e-10 (7.75e-12)**
%CollegeGrad	-0.001 (0.002)	-0.008 (0.002)**	0.001 (0.001)
%<HighSchool	0.005 (0.002)*	0.005 (0.003)	0.002 (0.001)*
%HouseOwn.	-0.002 (0.001)*	-0.006 (0.001)**	-0.001 (4.18e-04)**
%Professional	0.004 (0.002)	0.006 (0.003)*	0.005 (0.001)**
%Executives	4.04e-04 (0.002)	-0.001 (0.003)	1.39e-04 (0.001)
%BlueCollar	9.19e-05 (0.002)	-0.002 (0.002)	0.003 (0.001)**
%Technicians	-0.008 (0.006)	0.016 (0.007)*	-0.001 (0.003)
AnyTrade	-0.007 (0.023)	0.586 (0.028)**	0.313 (0.013)**
Competition	-0.086 (0.006)**	0.063 (0.007)**	-0.034 (0.003)**
Constant	995.8 (0.224)**	1007.8 (0.282)**	1001.3 (0.127)**
Observations	650850	650850	650850

* significant at 5%; ** significant at 1%
Standard errors in parentheses
† Unreported are *EndOfMonth, WeekEnd, DVehCost*, car, month, region, and model recency fixed effects

Table 5: Regressions for explanations section†

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable ln(price)	Full Sample	Franchise Fixed Effects	Full Sample	Full Sample	Full Sample	No Financing
%Black	0.01966 (0.00051)**	0.01301 (0.00054)**	0.01312 (0.00089)**	0.01302 (0.00070)**	0.01861 (0.00065)**	0.012 (0.001)**
%Hispanic	0.02332 (0.00081)**	0.0102 (0.00106)**	0.00060 (0.00136)	0.00770 (0.00118)**	0.01356 (0.00113)**	0.007 (0.002)**
%Asian	-0.00963 (0.00093)**	0.00023 (0.00098)	-0.00174 (0.00159)	-0.00123 (0.00120)	-0.00391 (0.00096)**	-0.002 (0.002)
Hispanic		0.49058 (0.02636)**				0.323 (0.061)**
Asian		-0.75659 (0.04176)**				-0.692 (0.068)**
Female	0.20619 (0.01388)**	0.19335 (0.01322)**	0.22654 (0.02233)**	0.20607 (0.01391)**	0.20536 (0.01391)**	0.287 (0.025)**
CustomerAge		0.00444 (0.0006)**	0.00446 (0.00063)**	0.00448 (0.00063)**	0.00446 (0.00063)**	0.002 (0.001)*
Age> 64		-0.1342 (0.02822)**	-0.16903 (0.02954)**	-0.16803 (0.02953)**	-0.17039 (0.02954)**	0.294 (0.044)**
Competition	-0.03703 (0.00342)**		-0.04851 (0.00522)**	-0.02974 (0.00357)**	-0.02170 (0.00351)**	-0.017 (0.007)*
—— ∗ %Black			0.00042 (0.00020)*			
—— ∗ %Hispanic			0.00316 (0.00029)**			
—— ∗ %Asian			-0.00067 (0.00038)			
—— ∗ Female			-0.00705 (0.00618)			
PopDensity				0.02442 (0.00364)**		
—— ∗ %Black				0.00023 (0.00011)*		
—— ∗ %Hispanic				0.00059 (0.00019)**		
—— ∗ %Asian				-0.00089 (0.00017)**		
AnyTrade	0.33640 (0.01373)**	0.248 (0.013)**	0.30688 (0.01377)**	0.30933 (0.01377)**	0.43094 (0.01807)**	0.811 (0.024)**
—— ∗ %Black					-0.01061 (0.00103)**	
—— ∗ %Hispanic					-0.00719 (0.00143)**	
Constant	1,000.99 (0.10778)**	986.4 (1.83e+9)	1,001.7 (0.13132)**	1,001.6 (0.13122)**	1,001.6 (0.13082)**	1006.8 (0.253)**
Observations	650850	615349	650850	650850	650850	159819
R^2	0.97	0.98	0.97	0.97	0.97	0.98

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

† Unreported are *EndOfMonth, WeekEnd, DVehCost*, 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*.

Table 6: Regression for Autobytel.com results†

Dep. Variable ln(price)	(1) Full Sample	(2) Full Sample	(3) Full Sample	(4) Franchice Fixed Effects	(5) > 75% or < 2% Minority
Autobytel	-0.87654 (0.02826)**	-0.59146 (0.04535)**	-0.62544 (0.04530)**	-0.00614 (0.04386)**	-0.82428 (0.03649)**
AutobytelFranchise	-0.45880 (0.01511)**	-0.45926 (0.01510)**	-0.48615 (0.01510)**	0.16628 (0.06887)*	-0.38789 (0.01938)**
%Black	0.01463 (0.00053)**	0.01481 (0.00054)**	0.01955 (0.00051)**	0.01317 (0.00054)**	
%Hispanic	0.00708 (0.00102)**	0.00746 (0.00102)**	0.01858 (0.00084)**	0.01055 (0.00105)**	0.00342 (0.00134)*
%Asian	-0.00066 (0.00095)	-0.00033 (0.00097)	-0.00544 (0.00094)**	0.00003 -(0.00097)	-0.00010 (0.00134)
Hispanic	0.50563 (0.02722)**	0.51176 (0.02758)**	0.53103 (0.02752)**	0.49312 (0.02635)**	0.53824 (0.03715)**
Asian	-0.95493 (0.04190)**	-0.96174 (0.04330)**	-0.97540 (0.04321)**	-0.76124 (0.04164)**	-0.84341 (0.05583)**
Female	0.20768 (0.01359)**	0.21127 (0.01388)**	0.21156 (0.01384)**	0.19429 (0.01320)**	0.19209 (0.01748)**
%Black> 75					1.36784 (0.06182)**
Autobytel ∗ %Black		-0.01230 (0.00281)**	-0.01137 (0.00281)**	-0.01199 (0.00265)**	
— ∗ %Hispanic		-0.02026 (0.00383)**	-0.02052 (0.00382)**	-0.01200 (0.00373)**	
— ∗ %Asian		-0.00696 (0.00326)*	-0.00719 (0.00326)*	0.00075 -(0.00321)	
— ∗ Hispanic		-0.57086 (0.14903)**	-0.57085 (0.14919)**	-0.53253 (0.13843)**	
— ∗ Asian		0.14260 (0.16351)	0.14319 (0.16369)	0.08893 -(0.15661)	
— ∗ Female		-0.12125 (0.05836)*	-0.12180 (0.05836)*	-0.09965 -(0.05537)	
— ∗ %Black> 75					-0.86518 (0.41742)*
CustomerAge	0.00455 (0.00062)**	0.00453 (0.00062)**		0.00427 (0.00059)**	0.00278 (0.00079)**
Age> 64	-0.16518 (0.02906)**	-0.16411 (0.02906)**		-0.12776 (0.02779)**	-0.14263 (0.03609)**
AnyTrade	0.31169 (0.01350)**	0.31173 (0.01350)**	0.33687 (0.01346)**	0.25608 (0.01304)**	0.33617 (0.01731)**
Competition	-0.03006 (0.00346)**	-0.03015 (0.00346)**	-0.04411 (0.00337)**		-0.02589 (0.00427)**
Constant	1,001.9 (0.13868)**	1,001.9 (0.13866)**	1,001.3 (0.11706)**	1010.8 -(8.5E+12)	1,004.08348 (0.17538)**
Observations	671468	671468	671468	635050	398566
R^2	0.98	0.98	0.98	0.98	0.98

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

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

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

Table 7: Selection results

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Variable ln(price)	.1 Quantile	.9 Quantile	IV	IV, $> 75\%$ or $< 2\%$ Black	IV, Income $< \$53,000$	IV, % Less HighSch. > 10
Autobytel	-0.235 (0.080)**	-1.349 (0.100)**	-1.18388 (0.93250)	-2.46573 (1.21260)*	-7.42656 (1.63748)**	-5.14595 (1.65018)**
AutobytelFranchise	-0.408 (0.026)**	-0.491 (0.031)**	-0.41889 (0.02651)**	-0.33685 (0.03512)**	-0.34815 (0.03723)**	-0.37384 (0.03828)**
%Black	0.007 (0.001)**	0.025 (0.001)**	0.01461 (0.00051)**		0.01370 (0.00062)**	0.01305 (0.00060)**
%Hispanic	0.001 -0.002	0.011 (0.002)**	0.00610 (0.00107)**	0.00323 (0.00134)*	0.00402 (0.00136)**	0.00251 (0.00134)
%Asian	-0.01 (0.002)**	0.014 (0.002)**	0.00040 (0.00099)	0.00086 (0.00142)	-0.00130 (0.00164)	0.00163 (0.00144)
Hispanic	0.255 (0.044)**	0.842 (0.053)**	0.54132 (0.03972)**	0.50521 (0.03807)**	0.43245 (0.03857)**	0.45497 (0.03770)**
Asian	-0.624 (0.070)**	-1.450 (0.090)**	-0.97515 (0.04948)**	-0.86012 (0.06657)**	-1.15966 (0.08551)**	-1.17733 (0.08029)**
Female	0.219 (0.023)**	0.272 (0.028)**	0.19977 (0.01594)**	0.17344 (0.01997)**	0.17948 (0.02293)**	0.18907 (0.02261)**
%Black>75				1.62439 (0.09797)**		
Autobytel $*$ %Black	-0.011 (0.006)*	-0.022 (0.007)**				
— $*$ %Hispanic	-0.006 -0.007	-0.038 (0.009)**				
— $*$ %Asian	0.012 (0.006)*	-0.03 (0.007)**				
— $*$ Hispanic			-2.89706 (1.81231)			
— $*$ %Black>75				-33.32440 (10.09526)**		
CustomerAge	0.004 (0.001)**	0.006 (0.001)**	0.00437 (0.00074)**	0.00176 (0.00091)	0.00337 (0.00105)**	0.00351 (0.00103)**
Age>64	-0.184 (0.049)**	-0.133 (0.060)*	-0.17571 (0.03159)**	-0.14833 (0.03799)**	-0.24489 (0.04503)**	-0.22152 (0.04472)**
AnyTrade	-0.011 -0.023	0.596 (0.028)**	0.31496 (0.01628)**	0.28495 (0.01994)**	0.14703 (0.02382)**	0.12040 (0.02380)**
Competition	-0.097 (0.006)**	0.056 (0.007)**	-0.02738 (0.00358)**	-0.02750 (0.00430)**	-0.01294 (0.00555)*	-0.00521 (0.00511)
Constant	997.7 (0.242)**	1006.8 (0.293)**	1,001.5 (0.14884)**	1,004.0 (0.16893)**	995.3 (0.24738)**	996.2 (0.19013)**
Observations	671468	671468	576076	361870	312118	300558

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

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

[1] Instruments for column 3: references in the zip code, family size, number of cars linear, squared, and cubed, percent with Internet access at work, predicted probability of using Autobytel.com, the prediction times the black indicator, and predicted probability of *Black*Autobytel*. Column 4: references in the zip code, family size, number of cars linear and squared, predicted probability of using Autobytel.com, the prediction times the Hispanic indicator. Column 5: references in the zip code, number of cars linear, squared, and cubed, predicted probability of using Autobytel.com, predicted probability of using Autobytel.com if income<53K. Column 6: predicted probability of using Autobytel.com, predicted probability of using Autobytel.com if lesshs>10.

Figure 1: Distribution of percent minority in census block group for new car buyers

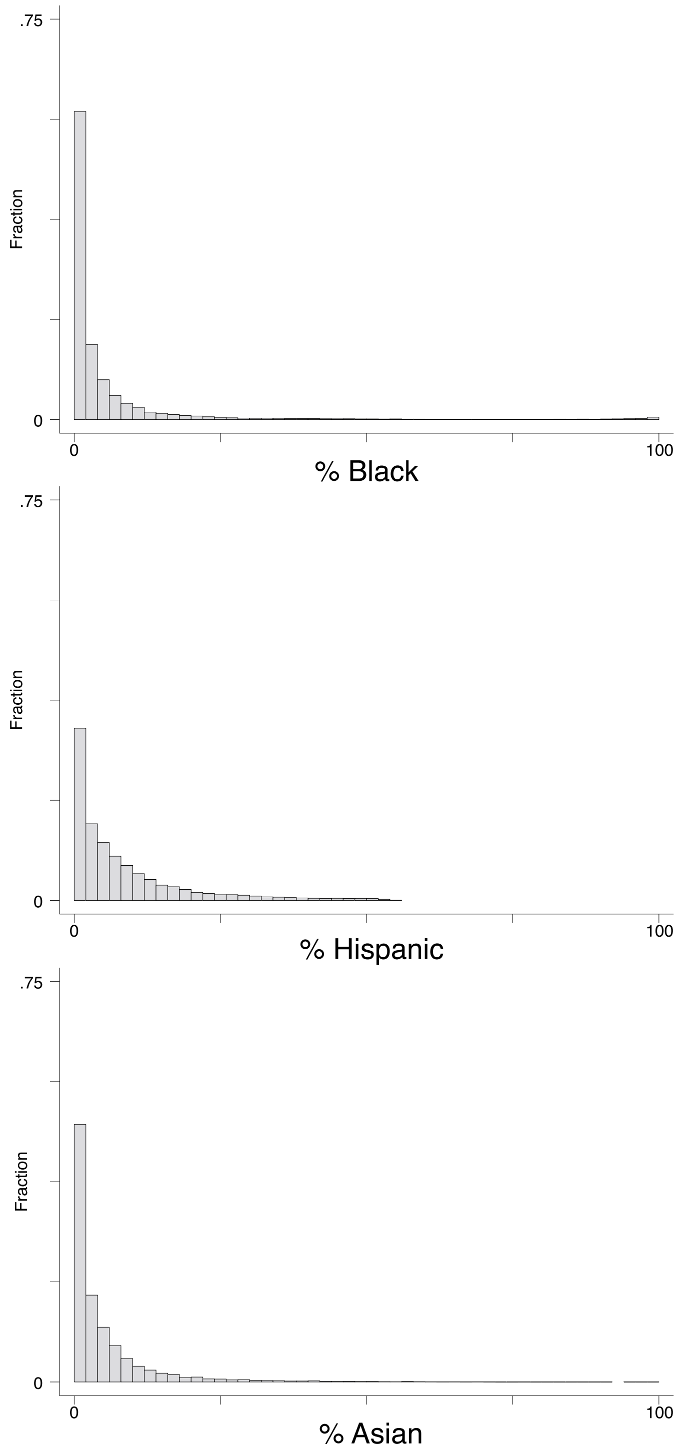

Figure 2: Distribution of percentage margins