Demographic Preferences and Price Discrimination in New Vehicle Sales*

Ashley Langer
UC Berkeley

Fall 2009

Abstract

In this paper I expand the literature on discrimination in automobile pricing by exploring the extent to which differences in the prices paid by different demographic groups can be explained by firms' profit maximizing price discrimination based on demographic groups' preferences. I estimate seperate random coefficient discrete choice models for each demographic group and use these estimates to calculate optimal oligopoly price markups for each vehicle to each demographic group. I find that, for gender and marital status groups, a one dollar increase in the difference between demographic groups in a vehicle's optimal markup leads to a 30 to 45 cent increase in the observed average price difference for that vehicle. Thus firms are engaging in third-degree price discrimination based on consumer gender and marital status. I find that, controlling for third-degree price discrimination, women and single consumers pay more for a vehicle than their male and married counterparts.

*I would like to thank David Card, Patrick Kline, Enrico Moretti, Kenneth Train, Clifford Winston, and Catherine Wolfram for their guidance throughout this project. Zachary Anderson at JD Power and Andrew Lumsden at Autodata were extraordinarily helpful with procuring the data.

1 Introduction

The market for new automobiles is a well-known and important setting for economists' investigations about whether a consumer will be charged a higher price because he or she belongs to a particular demographic group. For example, Ayers and Siegelman (1995) found that African American men negotiated substantially higher prices than white men for a particular car in an experiment. More recently, Morton, Zettelmeyer, and Silva-Risso (2003) found that both white women and African AMerican of both genders pay more than white men for similar new vehicles. Goldberg (1996) does not find statistically significant differences in the percent discount of off manufacturer's suggested retail price (MSRP) for different groups, but Harless and Hoffer (2002) do find some differences in dealer profits by gender and age.

These earlier papers have almost always looked at how the average price or profit varies with demographics, controlling for vehicle characteristics.[1] While this has allowed the authors to understand whether demographic groups experience different market outcomes, it has not been particularly useful at decomposing *why* these different outcomes may occur. The literature recognizes that there are many reasons why differences in average outcomes may occur. If firms are able to identify a consumer's vehicle preferences from her demographics, then they may be able to maximize profits by charging, for instance, men more for vehicles that they strongly prefer and women more for vehicles that they strongly prefer. Alternatively, it could be that dealer animus against certain consumers leads to the taste-based discrimination of Becker (1957). Other differences in consumer behavior or dealer competition could also lead groups to pay different amounts.

This paper is the first to explicitly begin to disentangle consumer preferences and third-degree price discrimination from other forces that may lead consumers to pay different prices for the same vehicle. If firms do use consumer demographics to engage in third-degree price discrimination, then the observed differences in prices paid by demographic group cannot be interpreted as evidence of taste-based discrimination because it may reflect economically rational price discrimination. Similarly, the lack of an average price differential between demographic groups does not rule out taste-based discrimination. These distinctions are obviously important for public anti-discrimination policy and are relevant to many industries besides automobiles.

To understand whether firms price discriminate based on demographic group preferences and whether price differences between demographic groups for the same vehicle remain once I control for third-degree price discrimination, I turn to the literature on discrete choice

[1] Goldberg (1996) does look at the variances in prices over demographic groups, but is contrained by her data in the depth at which she can explore this issue.

demand estimation. This literature has advanced substantially from McFadden's 1974 paper on the utility foundations of logit demand, and can now incorporate consumer heterogeneity along both observed and unobserved consumer dimensions to improve predicted substitution patterns and optimal product markups. I will estimate consumer demand separately for different demographic groups while allowing for extensive heterogeneity in preferences within each group. I assume that each demographic group's demand is structured similarly to Berry, Levinsohn, and Pakes (200X, henceforth MicroBLP). Thus I allow the mean utility of each demographic group to vary with the unobservable quality of the vehicle. I estimate demand for each demographic group using maximum likelihood as in Train and Winston (2007) but with the Berry (1994) contraction to reduce the dimensionality of the estimated coefficient vector. Using the demand paramters for each demogrpahic group, I can then calculate the profit-maximizing price that differentiated product oligopolies would charge each group for each vehicle.

For tractability, I focus on the demand of four distance demographic groups: married women, married men, single women, and single men.[2] I find substantial differences in vehicle demand between these groups. On average, I find that women are more price sensitive than men and single people are more price sensitive than married people. Given that a large amount of the heterogeneity in preferences for price within demographic groups is driven by income differences, this difference in average price sensitivity likely reflects the fact that married people and men have higher household incomes than their single and/or female counterparts.

I find that all demographic groups substitute substantially between vehicles of the same type (car, truck, SUV, or van), but that married women on average strongly prefer SUVs to cars while single women have the opposite average preference. Men, both single and married, prefer vehicles with high curb weight, although the high heterogeneity in the preference probably indicates the preference for both large, heavy vehicles, and lighter, sportier cars. Women, on the other hand, are fairly indifferent to curb weight after controlling for other vehicle characteristics, and do not have much heterogeneity in their taste for curb weight. I discuss the extent of preference differences between groups in section 4.

Using this variation in preferences, I find optimal product markups for each demographic group that are consistent with earlier results in BLP and MicroBLP. I find that married men have the highest optimal markups that average approximately 40% of transaction prices, while single women have the lowest optimal markups that average approximately 20% of

[2] I do not have enough observations to consider consumer race in my analysis, although future work may be able to look at differences in all preferences by consumer age, education, and income. For the current analysis I use income as an observed driver of preference heterogeneity within each demographic group, while age and education differences will contribute to unobserved heterogeneity.

transaction prices. Given that this is a market with high fixed costs and these are the markups over marginal cost for the entire supply chain (manufacturer and dealer) the magnitude of these markups is quite reasonable.

When I compare the differences in predicted optimal markups across demographic group pairs to the differences in observed average prices, I find that firms do engage in third-degree price discrimination. I find that a \$1 increase in the difference in optimal markups between two groups leads to a statistically significant 30 to 45 cent increase in the difference in observed average prices. The fact that firms do not appear to be able to completely capitalize differences in predicted markups into differences in observed prices could be the result of unmodelled competition between dealers of the same brand or concerns that full third-degree price discrimination might induce government intervention in the market to assure equal treatment between demographic groups. Additionally, once preference differences between demographic groups are considered, women and single buyers appear to pay more for new vehicles on average than their male and/or married counterparts. While this may be attributable to other characteristics of consumers such as preferences for search or negotiation, it also leaves open the possibility that taste-based discrimination is occurring in this market.

To understand the impact on consumer welfare of this price discrimination, I use the estimated demand functions to ask how the welfare of different groups would change in the absense of third-degree price discrimination. I find that... IN PROGRESS.

Finally, in section 6 I present the results of various specifications of the consumer demand model in order to investigate the robustness of my results to changes in functional form. I find that although different functional forms can lead to extreme outliers in terms of optimal markups, the main result that firms engage in third-degree price discrimination is quite robust to changes in the functional form of both the mean demographic group preference and the heterogeneity in preferences within each demographic group.

The remainder of the paper is organized as follows: in the following section I describe my empirical specification. The data used is explained in section 3. I then present results of the demand estimation and comparison of optimal markups to observed prices in section 4 and the results of the welfare calculations in section 5. Section 6 tests the robustness of my results to functional form and section 7 concludes.

2 Empirical Specification

The empirical model estimates the optimal price for each firm to charge each demographic group for each vehicle. This requires estimating the vehicle demand functions of each demographic group for each vehicle and pairing these estimates with a model of manufacturer and dealer pricing behavior. By comparing the optimal price-discriminating prices to the observed rpices, I can test for third-degree price discrimination in this market.

2.1 Demand Functions

The demand function follows directly from MicroBLP but is estimated using maximum likelihood as in Train and Winston (2007). I use the Berry (1994) inversion to reduce the dimensionality of the coefficient space.

Consumers are each assumed to belong to a single demographic group, $d = 1, ..., D$. Within these demographic groups, consumers are heterogeneous along both observable and unobservable individual characteristics. Consumer i's utility for vehicle $j = 0, 1, ..., J$ is assumed to be:

$$U_{idj} = p_{jd}\tilde{\alpha}_{id} + \sum_k x_{jk}\tilde{\beta}_{idk} + \xi_{dj} + \epsilon_{idj} \tag{1}$$

where p_{jd} is the price charged to i's demographic group d; $x_{j1}, ..., x_{jK}$ are the non-price attributes of vehicle j; ξ_{dj} is the preference of demographic group d for the unobservable attributes of vehicle j; and ϵ_{idj} is an EV1 residual preference parameter. The $\tilde{\alpha}_{id}$ and $\tilde{\beta}_{idk}$ are the individual's preference for vehicle attributes p_{jd} and x_k respectively, which are assumed to have the form:

$$\begin{aligned} \tilde{\alpha}_{id} &= \bar{\alpha}_d + \sum_r z_{idr}\alpha^o_{dr} + \nu_{idp}\alpha^u_d \\ \tilde{\beta}_{idk} &= \bar{\beta}_{dk} + \sum_r z_{idr}\beta^o_{dkr} + \nu_{idk}\beta^u_{dk} \end{aligned} \tag{2}$$

Thus the individual's preference for vehicle attribute x_k is decomposed into a component ($\bar{\beta}_{dk}$) that is constant within that individual's demographic group, a component (β^o_{dkr}) that varies with consumer characteristics z_{idr} that are observed by the econometrician, and a component (β^u_{dk}) that varies with consumer characteristics ν_{idk} that are unobserved by the econometrician, but are assumed to have a known distribution.[3] These unobserved consumer characteristics capture the fact that there is heterogeneity in preferences for different vehicle attributes in every demographic group, although we may not have variables that allow us to identify those consumers who get particular utility from horsepower, for instance, rather than side air bags.

[3] I will generally assume that unobserved consumer characteristics have normal distributions.

Combining equations (1) and (2) leads to the consumer's choice model:

$$U_{idj} = \delta_{dj} + \sum_r p_{jd} z_{idr} \alpha^o_{dr} + \sum_{k,r} x_{jk} z_{idr} \beta^o_{dkr} + p_{jd} \nu_{idp} \alpha^u_d + \sum_k x_{jk} \nu_{idk} \beta^u_{dk} + \epsilon_{idj} \tag{3}$$

$$\text{where} \qquad \delta_{dj} = p_{jd} \bar{\alpha}_d + \sum_k x_{jk} \bar{\beta}_{dk} + \xi_{dj} \quad \text{for each} \quad j = 1, ..., J \tag{4}$$

The consumer chooses the vehicle $j = 1, ..., J$ or the outside option ($j = 0$, not purchasing a new vehicle) that maximizes this utility function.[4] As this notation makes clear, there is a component (δ_{dj}) to each individual's utility for each vehicle that is common across all members of his or her demographic group d. Additionally, the term $\sum_r p_{jd} z_{idr} \alpha^o_{dr} + \sum_{kr} x_{jk} z_{idr} \beta^o_{dkr}$ allows consumers with different observable characteristics to have different tastes for certain vehicle attributes (and thus specifies the extent to which vehicle substitution varies with observable consumer demographics). Finally, there is a component of consumer preference ($p_{jd} \nu_{idp} \alpha^u_{dk} + \sum_k x_{jk} \nu_{idk} \beta^u_{dk}$) that is unobserved by the econometrician but helps explain why certain consumers have stronger preferences for some vehicle attributes rather than others, and helps to explain why individuals may substitute more strongly between certain vehicles. The β^u_{dk} and α^u_d coefficients can be thought of as representing the standard deviation in the unobserved preference within demographic group d for the vehicle attribute. For notational ease, I define the vector of distributional coefficients $\theta_d \equiv [\alpha^o_{dr},\ \alpha^u_{dr},\ \beta^o_{dkr},\ \beta^u_{dkr}]'$.

I estimate the θ_d and δ_d coefficients via maximum-likelihood. The extreme-value error term guarantees that the probability of vehicle j maximizing consumer i's utility conditional on the observable attributes of the vehicle (p_{jd}, x_{jk}) and the consumer's observable (z_{idr}) and unobservable ($\nu_{id} = [\nu_{idp}, \nu_{idk}]'$) characteristics is:

$$Pr_{idj}(p_{jd}, x_{jk}, z_{idr}, \nu_{id}; \theta_d, \delta_d) = \frac{\exp(V_{idj}(p_{jd}, x_{jk}, z_{idr}, \nu_{id}; \theta_d, \delta_d))}{\sum_{l=0}^{J} \exp(V_{il}(p_{ld}, x_{lk}, z_{idr}, \nu_{id}; \theta_d, \delta_d))} \tag{5}$$

where $V_{idj}(p_{jd}, x_{jk}, z_{idr}, \nu_{id}; \theta_d, \delta_d)$ is the non-stochastic component of consumer i's utility for vehicle j from equation (3). To condense notation, I will write $V_{idj}(\nu_{id}; \theta_d, \delta_d)$ and $Pr_{idj}(\nu_{id}; \theta_d, \delta_d)$ with the understanding that the non-stochastic utility and probability are also a function of the observable data. Because ν_{id} is unobserved to the econometrician, the expected value of the probability unconditional on ν_{id} is:

[4]The outside option of not purchasing a new vehicle is assumed to have utility equal to $U_{i0} = \beta^0_{d0} \frac{1}{I_i} + \beta^u_{d0} \nu_{i0}$, where $I_i$is the income of consumer i, and $\nu_i$is a draw from a standard normal distribution.

$$Pr_{idj}(\theta, \delta_d) = \int \frac{\exp(V_{idj}(\nu_{id}; \theta_d, \delta_d))}{\sum_{l=0}^{J} \exp(V_{idl}(\nu_{id}; \theta_d, \delta_d))} f(\nu) d\nu \tag{6}$$

where again $Pr_{idj}(\theta_d, \delta_d)$ is understood to also be a function of the observable data.

Because the θ_d coefficients determine how consumers substitute between vehicles as attributes change, it is important to include information on consumers' first and second choice vehicles in the likelihood function. Thus, the joint probability that consumer i chooses vehicle $j = 1$ out of the full choice set, and $j = 2$ out of the choice set with $j = 1$ and the outside good removed is:[5]

$$Pr_{i1}(\theta_d, \delta_d) Pr_{i2}(\theta_d, \delta_d | 1) = \int \frac{\exp(V_{id1}(\nu_{id}; \theta_d, \delta_d))}{\sum_{l=0}^{J} \exp(V_{idl}(\nu_{id}; \theta_d, \delta_d))} \left(\frac{\exp(V_{id2}(\nu_{id}; \theta_d, \delta_d))}{\sum_{l=2}^{J} \exp(V_{idl}(\nu_{id}; \theta_d, \delta_d))} \right) f(\nu) d\nu$$

Since the probability of observing a particular first and second choice for an individual is conditional upon the individual's ν vector, the integration over the distribution of ν must be for the joint probability. I approximate this integral using simulation, and then sum the log of this probability over consumers i in demographic group d to calculate the log-likelihood function.[6]

The log-likelihood function is maximized over θ_d. For each value of θ_d, I choose δ_d to set the predicted market shares for each demographic group equal to the observed market shares for that group as in Berry (1994):

$$\begin{aligned} S_{dj} &= \int_{z_{idr}} \int_{\nu} Pr_{idj}(\theta_d, \delta(\theta_d)) f(\nu) f(z_{idr}) d\nu dz_{idr} \\ &= Pr_{dj}(\theta_d, \delta(\theta_d)) \end{aligned} \tag{7}$$

where $f(z_{idr})$ is the pdf of the consumer characteristics z_{idr}in the demographic group d. Thus the maximum-likelihood procedure solves for the value of θ_d that maximizes the likelihood function subject to a market share constraint that is a function of both θ_d and $\delta(\theta_d)$.

This model differs from previous random-coefficient demand models in an important way: the preferences of each demographic groups d are assumed to be completely independent

[5]I remove the outside good from the second-choice choice set because the second choice information is based on the vehicle the consumer said she considered but did not purchase. It is not clear whether she would have purchased the second choice vehicle if the first choice were not available (she may not have purchased any vehicle), but it is her preferred alternative out of the set of vehicles once her first choice is removed.

[6]Simulation uses 100 scrambled Halton draws to approximate the integral for each consumer.

of the preferences of every other demographic group.[7] While this means that demographic groups may value the observable (to the econometrician) attributes of the vehicles differently, it is particularly important that the unobservable (to the econometrician) characteristics of a vehicle (ξ_{dj}) are allowed to be valued differently by members of different demographic groups. A prime example of such varying preference for vehicle unobservables would be the vehicles that are commonly referred to as "chick cars" or "guy cars"[8] such that the opposite gender might be interested in the vehicle for its physical attributes, but dissuaded from buying the car because of its social connotation. Additionally, options packages that appeal to one group rather than another (for instance spoilers or wheel rims) would potentially change the unobservable quality of the car for different groups differently.

Once I have estimated θ_d and calculated $\delta(\theta_d)$, I can use the δ_ds to extract information about the $\bar{\alpha}_d$ and $\bar{\beta}_{dk}$ coefficients rather than just the θ_d coefficients. Recall that:

$$\delta_{dj} = p_{jd}\bar{\alpha}_d + \sum_k x_{jk}\bar{\beta}_{dk} + \xi_{dj}$$

The problem is that unobservable vehicle quality, ξ_{dj} may include vehicle attributes that increase the price of the vehicle (for instance, a sunroof or leather seats). Therefore, an OLS regression of the δ_{dj} vector on vehicle price and attributes will estimate that consumers are less price sensitive than they actually are. In order to correct for this bias, I run a weighted IV regression of δ_{dj} on the vehicle price and attributes.[9] I use the standard Bresnahan (YEAR??) instruments:

$$\sum_{l \in F_j, l \neq j} x_{lk} \qquad \text{and} \qquad \sum_{l \notin F_j} x_{lk} \tag{8}$$

which are the sum of each competing vehicle attribute for vehicles produced by the same firm as vehicle j, F_j, and the sum of each competing vehicle attribute for vehicles produced by other firms. These instruments are intended to capture the extent of price competition faced by vehicle j in the market. For instance, if a vehicle is competing with a set of vehicles that have particularly high horsepower, then competitive pressure will keep the vehicle's price fairly low conditional on its attributes. If the observed price is actually high conditional on

[7]This also means that the preferences in the population (as estimated in the previous discrete choice literature) are a mixture of the preferences in each demographic group.

[8]See, for instance: http://www.cartalk.com/content/features/Guy-Chick-Cars/index.html

[9]The weights are equal to the number of consumers of demographic group d who chose vehicle j, which is a measure of the inverse of the variance of the estimate of $\delta(\theta_d)$ from the maximum-likelihood estimation.

attributes, it must be that the vehicle has a lot of unobservable attributes that are increasing its demand.

Because demographic groups face different prices and value vehicle attributes differently, the competitive pressure on price created by competing vehicles' attributes should vary over demographic groups. Therefore the instrumental variables regression is also run separately for each demographic group. Each vehicle observation is weighted by the total number of purchasers of the vehicle in the data to approximate the inverse sampling variance of the δ_{dj}.

The estimated demand coefficients allow me to calculate demand elasticities. Because the predicted demand of demographic group d for vehicle j is just the number of people in group d times the predicted market share of group d for vehicle j, the own-price elasticity of demand is just:

$$\frac{\partial Pr_{dj}(\hat{\theta}_d)}{\partial p_{dj}} \left(\frac{p_{dj}}{Pr_{dj}(\hat{\theta}_d)} \right) \tag{9}$$

which is straightforward to calculate given $\hat{\theta}_d$. This is the key component of demand for firms that are choosing prices to maximize profits.

2.2 Supply

Paired with this demand specification is a stylized model of vehicle supply. The standard supply model in the discrete choice literature assumes that manufacturers are able to perfectly contract with their dealers to charge the joint profit maximizing price for each vehicle. The supply model extends this idea to allow for third-degree price discrimination across demographic groups. "Firms" will consist of a manufacturer and a network of dealers who have optimal contracts to charge the profit maximizing price for each vehicle to each demographic group. Vehicles are assumed to have constant marginal costs, and supply is perfectly elastic so that the demand of any demographic group does not affect the demand of any other demographic group. As in the standard model, the product line is taken as exogenous and fixed.

Thus firm F sets prices to maximize profits over the vehicles it sells:

$$\pi_F = \sum_{d=1}^{D} \sum_{j \in F} Q_{dj}(p_d)(p_{dj} - c_j - D_d) \tag{10}$$

where the demand function of demographic group d for vehicle j is a function of the vector of the demographic group's prices for all vehicles, p_d, allowing for the possibility of taste-based discrimination (or other differences in the transaction for different demographic groups), D_d. The maximization of this set of profit functions for all firms leads to the vector of optimal prices given the vector of marginal costs, c:

$$\begin{aligned} P_d^* &= c - \Omega_d^{-1} Q_{dj} + D_d && (11) \\ &\equiv c + M_d + D_d && (12) \end{aligned}$$

where P_d^* is the optimal price vector for group d, M_d is the vector of optimal markups, and Ω_d is the matrix of own and cross-price derivatives of demand:

$$[\Omega_{djk}] = \begin{cases} \frac{\partial \mathrm{Q}_{dk}(\theta_d, p_{dj})}{\partial p_{dj}} & \text{if } j \text{ and } k \in F \\ 0 & \text{otherwise} \end{cases}$$

From the demand estimation, I have estimates of $\hat{\theta}_d$ and $\bar{\alpha}_d$, and I can therefore construct estimates of the demographic group's demand and price derivative matrix, $Q_{dj}(\hat{\theta}_d)$ and $\Omega_d(\hat{\theta}_d, \bar{\alpha}_d)$. Thus I have enough information to construct estimates of the optimal markups for each vehicle j sold to demographic group d, $\hat{M}_{dj}$. While I do not have information on the costs of vehicle j,[10] I do assume that the marginal vehicle costs are the same for all demographic groups, and therefore that the difference in the optimal price between demographic groups is equal to the difference in the optimal markup between groups plus any difference in the treatment of groups: $P_A^* - P_B^* = M_A - M_B + D_A - D_B$.

This relationship provides the basis for my estimation of the extent of third-degree price discrimination between different demographic groups in the market. While there may be many other considerations in price setting, the extent to which observed price differences track differences in the predicted optimal markup (which is based entirely on differences in preferences across demographic groups) should be a measure of the extent to which firms engage in third degree price discrimination. Therefore, in order to understand the extent to which observed price differences between demographic groups follow differences in predicted

[10]In BLP and MicroBLP, the authors use a moment similar to equation 11 to estimate their model, allowing the cost of each vehicle to be a linear combination of the vehicle's observed attributes. I do not exploit this moment, and therefore do not assume that the observed prices, p_{dj},are optimal. This leaves open the possibility of taste-based discrimination or another deviation from this pricing model in the observed data.

markups between groups, I run the regression:

$$\bar{p}_{Aj} - \bar{p}_{Bj} = \gamma_0 + \gamma_1(\hat{M}_{Aj} - \hat{M}_{Bj}) + \epsilon_j \tag{13}$$

where $\bar{p}_{Aj}$ is the average transaction price for vehicle j for demographic group A, and ϵ_j is the measurement error in the predicted price differences. In this regression, the coefficient $\gamma_0 = D_A - D_B$ is a measure of the difference in the experience of different demographic groups in the market, and γ_1 is a measure of the extent of third-degree price discrimination. If the above model were completely correct and firms were perfectly able to price discriminate based on the exact demographic groups used in estimation, then $\hat{\gamma}_1$ should be indistinguishable from one. If, on the other hand, firms do not take differences in preferences across demographic groups into consideration at all in thier pricing (or if firms are completely unable to observe the demographics of their consumers), then $\hat{\gamma}_1$ should be indistinguishable from zero. I will therefore use weighted OLS estimates[11] of $\hat{\gamma}_1$as my measure of the extent of third-degree price discrimination between any two demographic groups in my analysis, and plots of $p_{Aj} - p_{Bj}$ versus $\hat{M}_{Aj} - \hat{M}_{Bj}$ as a graphical represenation of how price differences vary with predicted markup differences. While the estimate of γ_0 does include any differences in taste-based discrimination between demographic groups A and B, it also would include any differences in preferences for negotiation or search behavior that is not modeled here.

3 Data

The primary data for this analysis is a survey of new vehicle buyers conducted by JD Power. This data is augmented with data from the Current Population Survey, the Automotive News Market Data book, and Autodata Solutions.

The survey of new vehicle buyers includes 25,875 respondents who purchased new vehicles in the second quarter of 2005. The survey includes information on the model of vehicle purchased and the other models considered, but does not include information on the trim level or the options packages of the vehicle.[12] The survey asks respondents a series of questions about their purchase, including the price they paid for the vehicle, and whether they paid cash for the vehicle, leased it, or secured a loan for the vehicle. Additionally, respondents indicated their age, gender, marital status, education, household income, and race on the

[11]The weights are equal to $\frac{1}{N_{Aj+N_{Bj}}}$ where $N_{Aj}$is the total number of consumers in demographic group A who purchase vehicle j. This weighting scheme takes into account the fact that the variance in the mean price difference across demographic groups for each car decreases as the number of people purchasing the car increases.

[12]I will follow the standard practice of assuming that the other models considered are listed in the order in which they were considered in order to identify the consumer's second choice vehicle.

survey.[13] In a particularly relevant question, the survey asks whether the respondent is both the "principle buyer and driver" of the vehicle. 21,085 respondents indicated that he or she was both the principle buyer and driver, and I will limit my analysis to these respondents in order to assure that the demographic information matches the driver of the vehicle and the person who physically purchased the vehicle.[14]

I remove from consideration any consumers who purchased a vehicle with an average sales price of over 75 thousand dollars in order to limit the analysis to commonly purchased vehicles. My primary analysis will focus on four demographic groups: married women, married men, single women, and single men. In order to calculate prices for each demographic group for every vehicle, I only include vehicles which at least one survey respondent of each demographic group purchased. When combined with the restriction that all of the relevant questions were answered, these restrictions bring my dataset down to 10,735 consumers. 58% of my sample is male and 64% is married. New car buyers tend to be wealthier than the average American, with 38% of respondents coming from households making over $100k and only 25% coming from households making less than $50k. 53% of respondents in my sample have a college degree.

In order to account for consumers who decided not to purchase a new vehicle in the second quarter of 2005, I append observations to my sample with consumers from each demographic group who purchased the "outside good", e.g. decided not to purchase a new vehicle in this quarter. This requires me to weight observations so that the outside good "purchasers" are accurately represented in the data. To do this, I first use information from the Automotive News Market Data Book on the total number of vehicles of each model sold in the second quarter of 2005 to weight the data for consumers who purchased a new vehicle.[15] I assume that the rates at which each demographic group purchases each vehicle model and the prices they pay in my data are representative of all purchasers of that model. This assures that the total vehicle weight on new vehicle purchasers will be equal to the total number of new vehicles purchased in the quarter. Then I use information from a survey by GfK Automotive Research that indicates that approximately 20% of Americans seriously consider purchasing a new car in a given year. I assume that this means that 10% seriously consider purchasing a new car in the second quarter of the year, meaning that the total weight on purchasers and non-purchaser in the data should be equal to 10% of the US population. I use data from the

[13]Consumers were asked to indicate the range in which their education and household income fell, rather than the exact amount.

[14]Of course, many people may take a friend or family member with them to purchase a vehicle, in which case the dealer may not be completely sure who the primary driver of the vehicle is.

[15]Note that in the second quarter almost every vehicle has the model year equal to the calendar year, which alleviates the issue of the mix of model years of vehicles sold.

Current Population Survey on the number of Americans over age 15 in each demographic group and give weights for the outside good observations equal to one tenth of the total US population in the demographic group minus the total number of people in the group who purchased a new car, as defined by the weights on the purchasers in the sample.[16]

This data on consumer purchases is paired with data from AutoData Solutions on the attributes of model year 2005 vehicles. This data includes a plethora of information on the vehicle, including (but not limited to) the manufacturer's suggested retail price (MSRP), horsepower, curb weight, wheel base, fuel economy, turning radius, and whether the vehicle has stability control, traction control, or side airbags. This data is at the vehicle trim level, which allows it to differ for the same vehicle model based on differences such as engine type (e.g. V6 vs V8) or body style (e.g. hatchback vs sedan). Since my consumer choice data only specifies a consumer's purchase decision at the model level, I use the vehicle attributes of the trim with the lowest MSRP as the model attributes and consider any deviations from this as "unobserved quality". This reinforces the idea that consumers of different demographic groups might have different valuations of unobserved quality, since not only the vehicle's styling may be valued differently but also the average trim level chosen may vary by demographic group. To the extent that many options such as leather seats, rear spoilers, or sunroofs may be fairly inexpensive to produce but command a large markup, these options packages may be a way for firms to encourage consumers to self-select into options packages that are priced to further price discriminate.[17]

4 Results

I divide consumers into four demographic groups: married women, married men, single women, and single men.[18] The results are presented in three steps: the demand coefficients are presented first and include the coefficients governing observed and unobserved preference

[16]This weighting scheme assumes that consumers who consider buying a new car but don't have the average demographics of the population as a whole, as is assumed in BLP. Another approach would be to assume that these marginal consumers have the demographic characteristics of the vehicle purchasers, which would make them richer, older, whiter, and more heavily male than the average American.

[17]Although this is technically second-degree price discrimination, I will estimate it as a part of what I call third-degree price discrimination.

[18]These groups have the advantage of being fairly observable to dealers, but gender and marital status are clearly only a subset of the demographics that a dealer may observe or infer. In this analysis, differences in average income, age, education, and race across these four demographic groups will enter into the mean preference coefficients, $\bar{\alpha}$ and $\bar{\beta}$. I use household income as an observed determinant of consumer heterogeneity within demographic groups, but assume that prices are set to the demographic group as a whole rather than to different income classes within the demographic group. Age, education, and race differences within demographic groups will contribute to unobserved consumer heterogeneity while differences across demographic groups will enter into the mean preference coefficients.

heterogeneity within demographic groups and the mean preference coefficients. I then present the elasticities and optimal markups that are calculated from these demand coefficients, and finally I compare the predicted markup differences between pairs of demographic groups to the observed average price differences for those groups.

4.1 Demand Estimation Results

Since the mean demand coefficients are estimated using the δ vector from the estimation of the demand heterogeneity coefficients, I will discuss the heterogeneity in demographic group demand first, followed by the estimates of the mean preference parameters. The demand specification is the same for each demographic group.

The specification of consumer demand heterogeneity is comprised of three sets of coefficients. I specify price as having a normally distributed unobservable heterogeneity component as well as a component that varies with the demeaned inverse of consumer's household income to capture the fact that price sensitivity may depend upon the relative vehicle price. I then include normally distributed unobservable heterogeneity terms for the vehicle's type (SUV, pickup truck, van, car) and the outside good. This set of coefficients allows consumers to substitute more intensely within vehicle types, even conditional on vehicle attributes, and makes the model a generalization of a nested-logit framework. Finally, I allow for normally distributed unobservable heterogeneity in each demographic group's demand for vehicle attributes including fuel use (measured in gallons per hundred miles), curb weight, horsepower, and whether the vehicle has side air bags. The coefficients on all of these normally distributed unobservable heterogeneity terms can be interpreted as the standard deviation in the demographic group's preference for the vehicle attribute.

Table 2 presents the coefficient estimates for these consumer heterogeneity terms for all four demographic groups. I find that all groups except married females have small but statistically significant heterogeneity in their price preference, even after controlling for income differences. All groups except for single women exhibit substantial heterogeneity in price preferences based on income, and the signs are what is expected: wealthier consumers are less price sensitive than averae and poor consumers are more price sensitive than average. Consumers of all demographic groups display high variation int heir preference for different types of vehicles, which indicates that consumers of all groups substitute substantially within vehicles of the same type. Of particular interest is the fact that married women show relatively little heterogeneity in their demand for SUVs and single men show relatively little heterogeneity in their demand for pickup trucks. Since we might expect both groups to have fairly high mean preferences for these vehicle types (which is indeed supported in the mean preference regressions), the lack of heterogeneity in these preferences indicates that the demographic group is surprisingly united in their taste for these types of vehicles. The final component of the vehicle nests is the heterogeneity in preference for the outside good. Only married men and single women appear to have substantial heterogeneity in their demand for the outside good relative to a new vehicle, which may be a results of these groups being at the two extremes of the income distribution, on average.

Finally, I estimate varying amounts of heterogeneity in consumer demand for different vehicle attributes. Consumers of all demographic groups exhibit heterogeneity in their demand for fuel use, which might not be surprising in a country where Toyota Priuses and Chevrolet Suburbans increasingly share the road. Men, both married and single, display heterogeneity int heir prefernce for curb weight, a likely result of heterogeneity in their preferences for large, heavy trucks and SUVs and smaller, sportier performance cars. Perhaps similarly, married women exhibit large differences in their demand for horsepower, perhaps displaying differences in preference for power and ease of driving. There is not much heterogeneity within any demographic group in the demand for side air bags.

The maximum-likelihood estimation of the consumer hetergeneity terms also generates estimates of the δ vector of mean preference parameters. These mean preference coefficients order vehicles by preference for the average consumer of each demographic group and therefore can serve as a useful reality check of the first stage. For both single men and married men, the five highest mean preference vehicles are all pickup trucks. Women, however, vary much more substantially by marital status. The top five vehicles for married women are all SUVs, while three of the five *lowest* preference vehicles for single women are SUVs (and the other two are vans). Single women prefer sedans, with the Toyota Camry Sedan topping the list. At the bottom of all demographic groups' lists are luxury cars and SUVs that most likely appear to a wealthy minority. For instance, married owmen have both the Cadillace Escalade Sport Utility Truck (basically an SUV with a pickup truck bed) and the Hummer H2 Sport Utility Truck in the bottom five. Three of the four groups (all except for single women) have the ultra-expensive Audi A8 in the bottom five. Generally, these results correspond with our expectations about the types of vehicles that different demographic groups prefer.

Regressing these deltas on vehicle attributes using weighted instrumental variables generates the mean preference coefficients for each demographic group. In these regressions, I again include price (instrumented with the BLP instruments as discussed earlier)[19] and the

[19] I do not use the instruments constructed from every mean preference variable. The van instrument has very little variation such that it is primarily picking up whether the vehicle is produced by a major manufacturer. I exclude the curb weight instrument because the combination of curb weight, horsepower, and fuel use are nearly colinear, and the deviations from colinearity are likely picking up some of the

vehicle types with cars as the excluded group. Because the δ vector is scaled such that the δ for the outside good is zero and not included in the regression, the constant term captures the preference for cars relative to the outside good. Again, I include vehicle attributes such as fuel use, curb weight, horsepower, and turning radius (which can proxy for the inverse of vehicle performance) in the mean preference specification.

The results of the mean preference regression are reported in Table 3. Married women are more price sensitive than married men, single women are more price sensitive than single men, and single people of either gender are more price sensitive than their married counterparts. This most likely reflects the fact that single people generally have lower household incomes than married people and women have lower household incomes than men.

While all groups other than single women prefer SUVs to car, married women have a particularly large preference for SUVs. When paired with the lack of heterogeneity in married womens' preferences for SUVs, this result is particularly striking. Similarly, men have a strong preference for pickups over cars while women are more indifferent, and single men's lack of heterogeneity in their pickup demand means that this strong preference is fairly robust within the group. Although no group significantly prefers vans to cars, married women do have a fairly high van coefficient relative to other groups. All of the groups except single women have negative, significant constant terms, reflectin gthe low numbers of new car buyers conditional on income and choice set attributes in any given quarter.

In terms of vehicle attributes, all groups dislike high fuel use vehicles.[20] Men like high curb weight vehicles and women are indifferent, although again, men had high heterogeneity in their preference for curb weight. Single women have a surprisingly strong preference for horsepower, and married men hae a significant preference for low turning radius vehicles, although other groups have similar estimates of their preferences that are not statistically significant.

4.2 Elasticities and Optimal Markups

When converting these coefficient estimates into estimated markups, one useful statistic with some economic intuition is the aggregate own-price demand elasticity for each demographic group for each vehicle. The first three rows of Table 4 give some descriptive statistics on the distribution of own-price elasticities aross vehicles. Generally, these elasticities average from just under 3 (in absolute value) for married men to almost 6 for single women. These appear to be of the same magnitude as the elasticites reported in MicroBLP.

unobserved quality of the vehicle. This explaination needs work.

[20]Single females have a estimated mean preference for fuel use that's of the same magnitude as the other groups, but the high standard error means that the estimate is not statistically significant.

The second and third panels of Table 4 provide descriptive statistics for the vehicle markups for each demographic group in terms of both dollars and as a percent of the transaction price. As we would expect given the elasticities, single women have the lowest average elasticities and married men have the highest. The markups average between 20 and 40 percent of transaction prices. There are a few reasons to think that markups of this magnitude would be reasonable. First, this is an industry with high fixed costs for each model produced, and these are markups over marginal costs. Therefore, we would expect a firm to only bring a vehicle to the market if the firm believed that the vehicle would have high marginal profits. Additionally, BLP find markups in the range of 15 to 40 percent of MSRP. Since transaction prices are generally lower than MSRP we might expect slightly higher markups here.

I also calculated standard errors for the predicted optimal markups.[21] All of the markups are highly significantly different from zero. Yet it is more interesting to use these standard errors to test whether the markups for different demographic groups are statistically different from one another. IN PROGRESS...

4.3 Comparing Actual and Predicted Price Differences

The goal in calculating these optimal consumer markups is to understand whether firms are engaging in third-degree price discrimination between demographic groups. I test this by comparing the observed difference in average prices for a pair of demographic groups to the predicted price difference. Recall from Section 2that, under third-degree price discrimination,

$$\begin{aligned} P_{jA} - P_{jB} &= C_j + M_{jA} + D_A - (C_j + M_{jB} + D_B) \\ &= (D_A - D_B) + M_{jA} - M_{jB} \end{aligned}$$

where P_{jA} is the price charged for vehicle j to demographic group A. M_{jA} is the optimal markup for vehicle j to demographic group A, and D_A is any taste-based discrimination or consumer behavior that changes prices of all vehicles. When estimated with average prices and predicted optimal markups, the regression function is

$$\bar{P}_{jA} - \bar{P}_{jB} = \alpha + \beta(\hat{M}_{jA} - \hat{M}_{jB}) + \epsilon_j$$

where the ϵ_j captures the measurement error in the average price difference and therefore has a variance that is a function of the number of observations o n purchasers of vehicle j from demographic groups A and B. The α term measures the differences between demographic

[21] Using the delta method with a numerical derivative of the markup function with respect to the θcoefficients.

groups in the prices they pay controlling for preferences. The β allows for the fact that firms may not be able to extract the full third-degree price discriminating markup difference from consumers. This could be because of unmodelled competition between dealers of the same brand or because firms avoid having large price differences between demographic groups in order to avoid the appearance of taste-based discrimination.[22] It may also be that consumers would react to extreme price differences by sending a friend or family member with different demographics to purchase the vehicle, thus undermining the profitability of the practice. Regardless of the reasons, if I can reject the null hypothesis that $\beta = 0$, I will say that dealers engage in third-degree price discrimination, and if I reject the secondary null hypothesis that $\beta \geq 1$, I will say that dealers are unable to perfectly third-degree price discriminate.

Before diving into the results, a graphical representation of this regression is useful. Figure BB has the observed difference in average prices for two demographic groups on the y-axis and the predicted markup difference on the x-axis. If firms perfectly price discriminate based on differences in preferences but do not charge demographic gropus different prices for any other reason, then $\alpha = 0$ and β =1 and the regression line is the 45 degree line. If firms are unable to perfectly third-degree price discriminate, but still do not charge groups different prices for any other reason, then α =0 and $0 < \beta < 1$ and the regression line goes through the origin but with a positive slope less than one, as shown by line i. Finally, if firms engage in taste-based discrimination against group B or if group B has a lower taste for negotiation than group A (or many other reasons) then at $\hat{M}_A - \hat{M}_B = 0$, where the two groups' preferences lead to identical optimal markups, group B will still pay more than group A and α will be less than zero as in Figure BB's line ii. Thus the regressions can give us substantial information on both the intensity of third-degree price discrimination in the market and the difference in the average prices paid once I have controlled for third-degree price discrimination.

Table 5 gives the estimates of α and β for four demographic group pairs. The first set of results are for married people, comparing the prices and markups of married men minus married women. Figure 3 shows the plot of the weighted data and the regression line with the dotted 45 degree line included for reference. The slope of the regression line is 0.450, meaning that for every dollar increase in the difference between married men and married women's optimal markups, firms are able to increase the difference between the groups' average price paid by 45 cents. This coefficient is statistically different from both zero and one, so firms are engaging in third-degree price discrimination but are not able to extract the full third-degree price discriminating markup differences. Additionally, the estimate of the intercept is

[22]The practice of stating MSRPs for each vehicle may serve to enforce this type of behavior across dealers in that it limits the total difference in prices that dealers can charge.

negative and statistically different from 0 meaning that, once I control for optimal markups, married women pay on average more than married men. Since dollar values are measured in tens of thousands of dollars, the coefficient of -0.0594 means that married women are paying $594 more on average, controlling for optimal markups, than married men.

The remainder of Table 5 gives the regression coefficient for single men minus single women, married women minus single women, and married men minus single men. Figures 4,5, and 6 represent these results graphically. While the first two sets of results obviously represent discrimination based on gender conditional on marital status, the second two represent discrimination based on marital status conditional on gender.[23] The recurring result is that firms do engage in third-degree price discrimination, although not to the full extent predicted by the model. None of the slope coefficients are statistically different from any of the others, but all of them are statistically different from 0 and 1. Thus, third-degree price discrimination based on consumer demogrpahics does contribute to the differences between demographic groups in the rpices paid for new vehicles.

Additionally, the slope coefficients are consistent in their estimates of which groups pay more on average for new vehicles controlling for third-degree price discrimination. Women, both married and single, pay more than their male counterparts once I control for third-degree price discrimination. Single people, both female and male, pay more than their married counterparts. These results should be interpreted with caution, however, since they are highly dependent upon the relative price sensitivity of each demographic group, which stems largely from the price coefficient in the IV regression of mean preferences on vehicle attributes. If the instruments are not perfect and allow different amounts of bias into the different demographic groups' estimates, then these estimates will be biased. Future work could focus on more robust estimates of these average price differences controlling for optimal product markups.

5 Welfare Implications of Price Discrimination

To understand which demographic groups are most impacted by third-degree price discrimination, I calculate the welfare change for each demographic group from a switch to a regime without third-degree price discrimination where firms charge all consumers the same optimal markup.[24] This thought experiment is of increasing interest as firms move to selling cars

[23]Remember that other consumer demographics may, on average, be entering here as well. For instance, if married people tend to be older than single people, then discrimination based on marital status may also be capturing discrimination based on age.

[24]In order to assure that I am only capturing the effect of third-degree price discrimination and not including other reasons why demographic gropus may pay different prices on average, I allow D_d to remain

more through online intermediaries where the firm may not even see the consumer's name and address until after a price is decided on.

I calculate the welfare changes under both a movement from perfect third-degree price discrimination and form a model like the one estiamted where firms can only extract 40% of the optimal markup differences. I find that...

IN PROGRESS: Results will be presented in seminar 10/28

6 Robustness

IN PROGRESS: Results will be presented in seminar 10/28

7 Conclusion

This paper explores the extent to which differences in demographic groups' preferences may lead to third-degree price discrimination. I find that firms do engage in third-degre price discrimination, but that differences between demographic groups in the price paid for a new vehicle remain even after controlling for this profit-maximizing behavior. Removing the ability to engage in third-degree price discrimination would benefit married men and hurt single women, but increase consumer surplus overall. figure 3

constant so that the thought experiment is a move from price P_0 to price P_1 where $P_1 - P_0 = C + M_1 + D - (C + M_0 + D) = M_1 - M_0$

References

[1] Ayres, Ian. 1991 "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations" 104 *Harvard Law Review* 817.

[2] Ayres, Ian and Peter Siegelman 1995 "Race and Gender Discrimination in Bargaining for a New Car" *American Economic Review* 85(3): 304-21

[3] Azar, Omar 2003 "Can Price Discrimination be Bad for Firms and Good for All Consumers? A Theoretical Analysis of Cross-Market Price Constraints with Entry and Price Discrimination" *Topics in Economic Analysis & Policy*, 3(1)

[4] Berry, Steve, Jim Levinsohn, and Ariel Pakes. 2004. "Differentiated Product Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy*, 112(1): 68-105.

[5] Berry, Steve, Jim Levinsohn, and Ariel Pakes. 1995 "Automobile Prices in Market Equilibrium" *Econometrica*, 63(4): 841-890

[6] Bresnahan, Timothy F., and Peter C. Reiss. "Dealer and Manufacturer Margins" *The RAND Journal of Economics*, 16(2): 253-268

[7] Borenstein, Severin 1985 "Price discrimination in free-entry markets" *Rand Journal of Economics* 16(3): 380-397

[8] Goldberg, Pinelopi. 1998. "The effects of the corporate average fuel economy standards in the automobile industry." *Journal of Industrial Economics*, 46: 1-33.

[9] Goldberg, Pinelopi. 1996 "Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey" *Journal of Political Economy*: 622-54

[10] McFadden, Daniel. 1974 "Conditional Logit Analysis of Qualitative Choice Behavior" in P. Zarembka (ed.) *Frontiers in Econometrics*: 105-142

[11] Revelt, D. and K. Train. 1998 "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level" *Review of Economics and Statistics* LXXX(4): 647-657

[12] Train, Kenneth and Cliff Winston. 2007. "Vehicle choice behavior and the declining market share of U.S. automakers." *International Economic Review*, 48: 1469-1496.

Table 1: Correlations Between Vehicle Price Paid and Consumer Demographics

	Dependent Variable: Price Paid						
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Male	135.98	150.48	165.58	9.47	110.25	-298.22	238.20
	(96.94)	(99.76)	(103.14)	(103.14)	(99.51)	(380.38)	
African-American		605.44**				676.42***	660.22***
		(247.54)				(250.43)	(250.97)
Household Income (000s)				10.35***		11.79***	10.73***
				(1.48)		(2.10)	(2.12)
Household Income (000s) if Male						4.65*	6.40**
						(2.54)	(2.65)
Married			-145.55			-502.66***	-496.43***
			107.91			(119.96)	(120.06)
Education					-39.51	-151.74***	-71.80
					30.96	(33.33)	(49.79)
Education if Male							-124.45*
							(64.63)
R^2	0.8472	0.8473	0.8473	0.8492	0.8492	0.8505	0.8505
# of observations	18,171	18,171	18,171	15,920	15,920	15,853	15,853
# of vehicle fixed effects	256	256	256	256	256	256	256

Eicker-White standard errors in parentheses. Astriks identify coefficients that are significant at the * 10%, ** 5%, or *** 1% level. The number of observations change because of missing observations in the survey. All regressions are weighted using sampling weights.

Table 2: Maximum Likelihood Results: Preference Heterogeneity Coefficients by Gender and Marital Status

		Gender and Marital Status			
Variable	Variable Type	Married Females	Married Males	Single Females	Single Males
Price	Standard Deviation	0.06	0.13***	0.13**	0.37***
(tens of thousands of dollars)		(0.12)	(0.03)	(0.06)	(0.07)
	Divided by Income	-6.20***	-5.79***	0.15	-5.28***
		(0.88)	(0.35)	(0.44)	(0.61)
SUV	Standard Deviation	0.60***	2.29***	3.76***	2.24***
		(0.21)	(0.23)	(0.67)	(0.59)
Pickup	Standard Deviation	3.59*	2.80***	1.84***	0.36
		(2.13)	(0.27)	(0.47)	(0.22)
Van	Standard Deviation	3.81***	3.85***	2.81***	2.39**
		(0.63)	(0.51)	(0.85)	(0.99)
Car	Standard Deviation	4.07***	4.22***	1.14***	4.26***
		(0.48)	(0.43)	(0.16)	(0.73)
Outside Good	Standard Deviation	0.38	0.55***	0.64**	0.43
		(0.27)	(0.17)	(0.32)	(0.26)
Fuel Use	Standard Deviation	0.24**	0.33***	0.23***	0.36***
(gallons per hundred miles)		(0.10)	(0.04)	(0.03)	(0.09)
Curb Weight	Standard Deviation	0.21	0.21***	0.01	0.63***
(thousands of pounds)		(0.17)	(0.05)	(0.10)	(0.13)
Horsepower	Standard Deviation	1.18***	0.02	0.01	0.37**
(hundreds)		(0.22)	(0.03)	(0.07)	(0.19)
Side Air Bag Dummy	Standard Deviation	0.10	0.09	0.35	0.26
		(0.25)	(0.15)	(0.31)	(0.29)

Standard errors in parentheses.

Significance level indicated by: *=10%, **=5%, ***=1%.

Table 3: IV Results: Mean Preference Coefficients by Gender and Marital Status

	Gender and Marital Status			
Variable	Married Females	Married Males	Single Females	Single Males
Price	-1.55***	-1.34***	-1.90***	-1.75***
(tens of thousands of dollars)	(0.43)	(0.29)	(0.58)	(0.33)
SUV	4.06***	1.90***	-3.58***	1.12**
	(0.54)	(0.41)	(0.58)	(0.46)
Pickup	0.05	3.30***	-0.38	4.09***
	(0.72)	(0.48)	(1.33)	(0.77)
Van	1.33	-0.32	-0.15	0.40
	(1.22)	(1.00)	(3.95)	(2.16)
Constant	-4.82***	-6.87***	0.94	-6.29**
	(1.41)	(1.31)	(3.96)	(3.20)
Fuel Use	-65.09**	-48.85***	-46.43	-64.18***
(gallons per 100 miles)	(30.37)	(14.31)	(51.17)	(16.35)
Curb Weight	0.34	1.33**	-0.05	1.11*
(thousands of pounds)	(1.02)	(0.59)	(1.12)	(0.63)
Horsepower	-0.09	0.47	2.26***	0.42
(hundreds)	(0.53)	(0.38)	(0.87)	(0.56)
Turning Radius	-0.15	-0.24**	-0.28	-0.21
(feet)	(0.11)	(0.10)	(0.22)	(0.20)

Standard errors in parentheses.

Significance level indicated by: *=10%, **=5%, ***=1%.

Table 4: Estimated Elasticities and Markups: Descriptive Statistics Over Vehicles

		Gender and Marital Status			
Variable	Statistic	Married Females	Married Males	Single Females	Single Males
Elasticity	Mean	-3.25	-2.81	-5.80	-3.87
	Min	-6.57	-4.71	-13.66	-5.42
	Max	-1.70	-1.66	-2.41	-1.99
$ Markup	Mean	10,114	11,689	5,537	8,470
	Min	7,324	7,308	5,219	5,727
	Max	15,292	18,413	6,238	16,676
% Markup	Mean	35.76	40.62	20.37	29.22
	Min	15.44	21.54	7.37	19.29
	Max	61.01	70.75	41.91	51.46

Table 5: Comparison of Observed Price Differences to Predicted Markup Differences

	Markup Difference	Intercept
Difference by Gender:		
Married Men minus Married Women	0.450***	-0.059***
	(0.106)	(0.018)
Single Men minus Single Women	0.315**	-0.066**
	(0.130)	(0.031)
Difference by Marital Status:		
Married Women minus Single Women	0.367***	-0.151***
	(0.127)	(0.053)
Married Men minus Single Men	0.417***	-0.145***
	(0.106)	(0.038)

Standard errors in parentheses

Significance level indicated by: **=5%, ***=1%.

Figure 1:

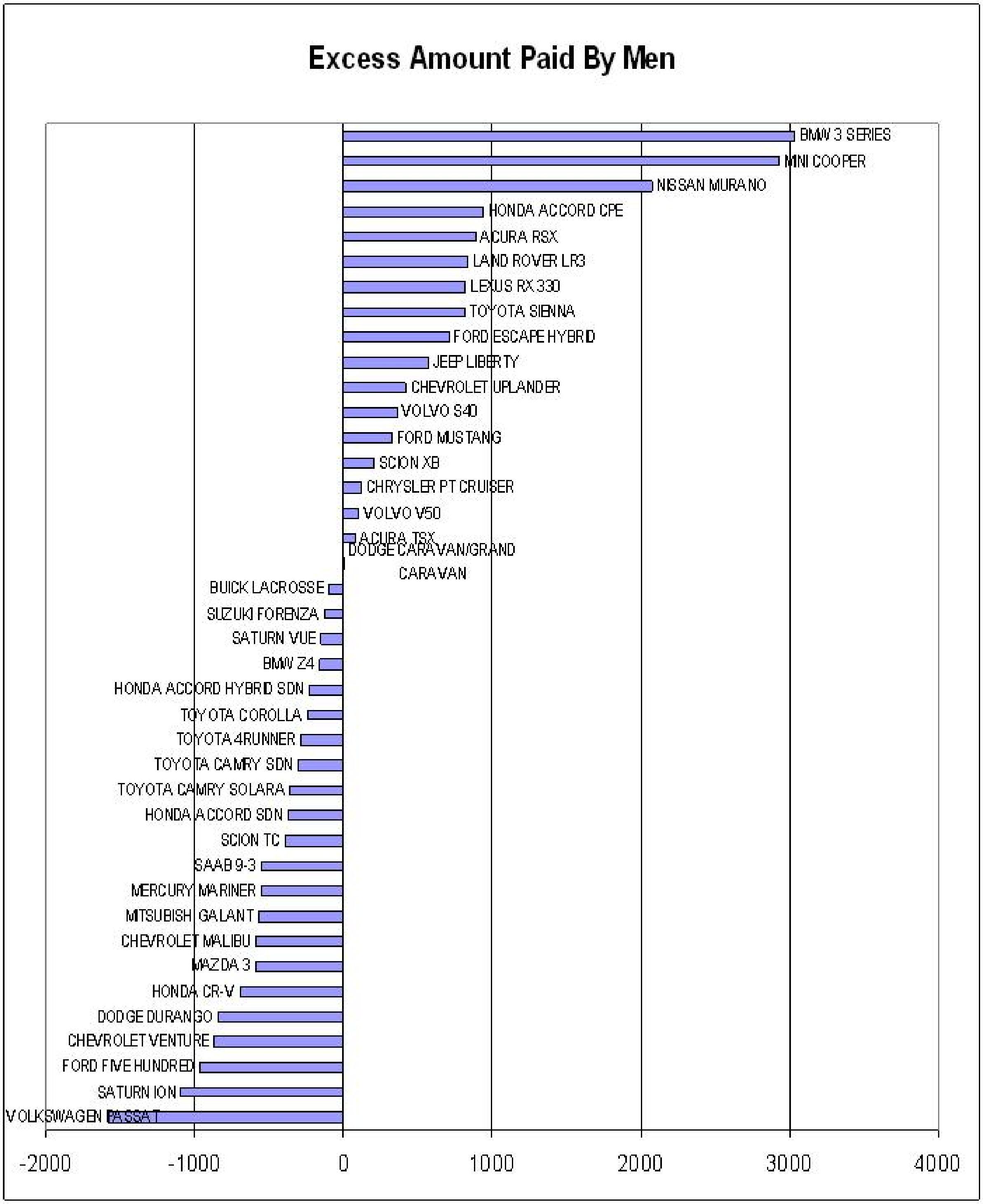

Excess Amount Paid By Men
BMW 3 SERIES
MINI COOPER
NISSAN MURANO
HONDA ACCORD CPE
ACURA RSX
LAND ROVER LR3
LEXUS RX 330
TOYOTA SIENNA
FORD ESCAPE HYBRID
JEEP LIBERTY
CHEVROLET UPLANDER
VOLVO S40
FORD MUSTANG
SCION XB
CHRYSLER PT CRUISER
VOLVO V50
ACURA TSX
DODGE CARAVAN/GRAND CARAVAN
BUICK LACROSSE
SUZUKI FORENZA
SATURN VUE
BMW Z4
HONDA ACCORD HYBRID SDN
TOYOTA COROLLA
TOYOTA 4RUNNER
TOYOTA CAMRY SDN
TOYOTA CAMRY SOLARA
HONDA ACCORD SDN
SCION TC
SAAB 9-3
MERCURY MARINER
MITSUBISHI GALANT
CHEVROLET MALIBU
MAZDA 3
HONDA CR-V
DODGE DURANGO
CHEVROLET VENTURE
FORD FIVE HUNDRED
SATURN ION
VOLKSWAGEN PASSAT
-2000
-1000
0
1000
2000
3000
4000

Figure 2: Example Comparison of Observed Price Difference to Predicted Markup Difference

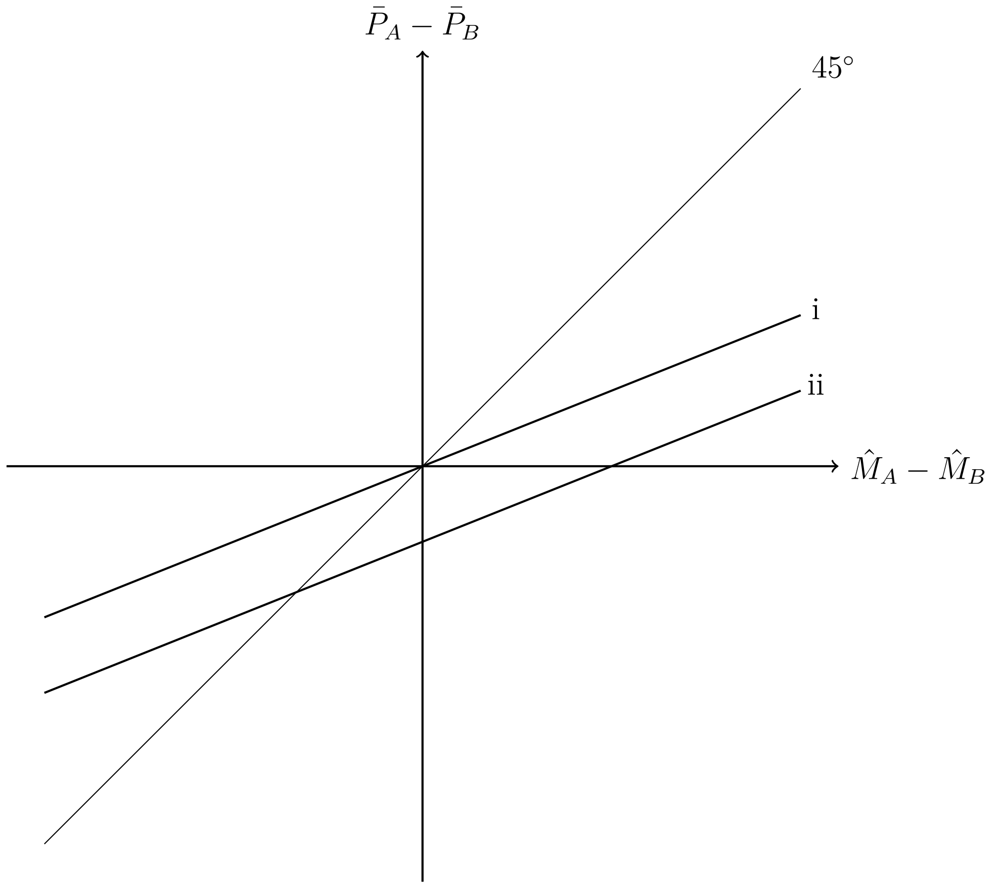
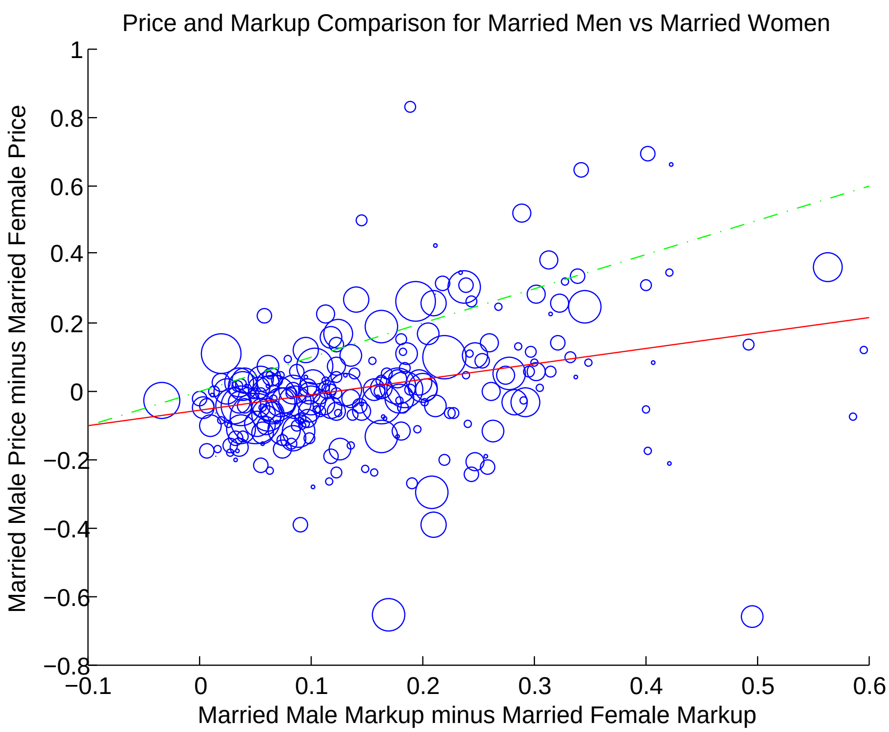


Figure 3:

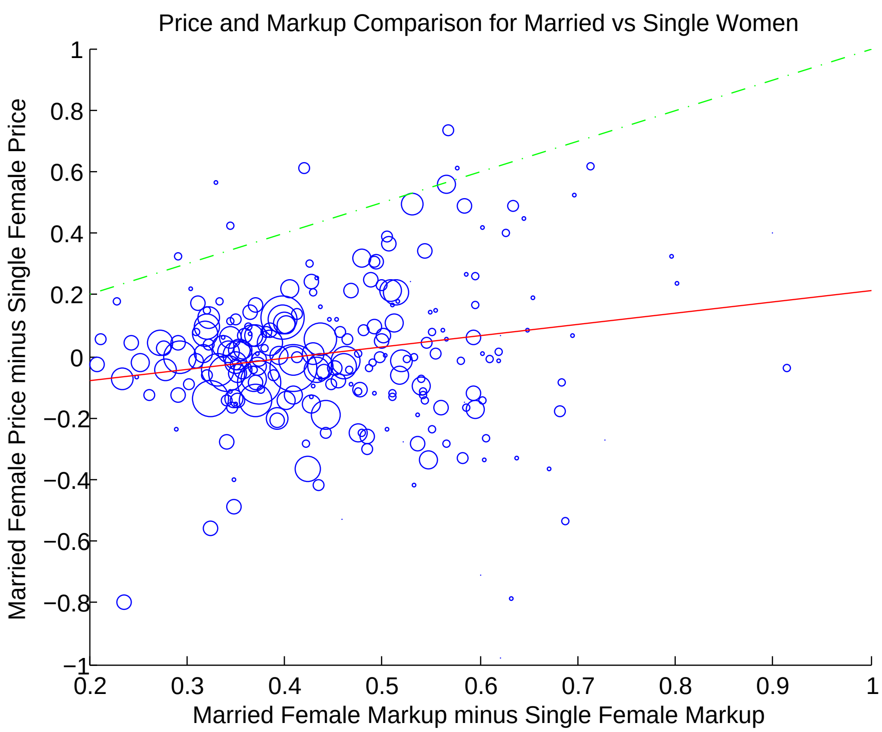


Price and Markup Comparison for Married Men vs Married Women
Married Male Price minus Married Female Price
Married Male Markup minus Married Female Markup
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.1
0
0.1
0.2
0.3
0.4
0.5
0.6

Figure 4:

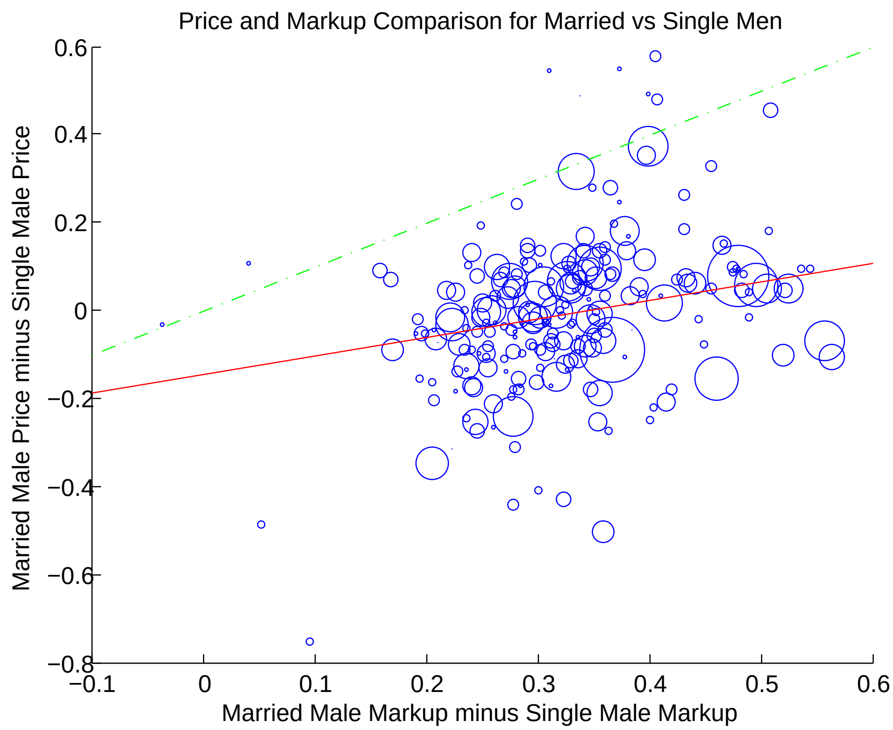

Price and Markup Comparison for Single Men vs Single Women
Single Male Price minus Single Female Price
1
0.5
0
−0.5
−1
−1.5
0
0.2
0.4
0.6
0.8
1
1.2
1.4
Single Male Markup minus Single Female Markup

Figure 5:



Price and Markup Comparison for Married vs Single Women
Married Female Price minus Single Female Price
Married Female Markup minus Single Female Markup
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Figure 6:


Price and Markup Comparison for Married vs Single Men
Married Male Price minus Single Male Price
Married Male Markup minus Single Male Markup
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−0.1
0
0.1
0.2
0.3
0.4
0.5
0.6