Biases from Overlooking Price Discrimination and a Solution

Benjamin Reed Shiller Shuran Zhang

Brandeis University

April 14, 2015

Abstract

Researchers often ignore price discrimination, instead assuming that all consumers face the sales-weighted average price. We show this introduces measurement error, biasing upwards the quantity demanded and the absolute slope of demand. These biases persist even with instrumental variable methods. We then propose a novel solution, and estimate the bias magnitude in the luxury car market. Despite only observing modest price discrimination - price rarely deviates from the average by more than 4% - we find evidence that demand is biased upwards by 5% to 14%, and its slope magnitude by 4% to 10%.

This paper first proves that because price discrimination is widespread but often ignored by researchers, estimates of both the quantity demanded at a given price and the absolute slope of demand are biased upwards. Moreover, these biases persist even when standard instrumental variable methods are used. Second, this paper introduces and implements a novel demand estimation method which solves this problem. In the process, a useful tangential result is found - biased demand estimates can still be used to yield unbiased estimates of firms' marginal costs.

The biases arise from aggregation. Researchers using aggregate data often assume that the price of a product equals revenues divided by quantity sold. This assumption is valid when the firm charges all consumers the same price. But when firms price discriminate, then dividing revenues by quantity sold yields the sales-weighted average price. Some consumers face prices above it, and others below it. Assuming all consumers are offered the average price introduces measurement error.

Problems ensue. Consumers offered prices below the average price would likely buy less if charged the average price. However, by assuming all consumers face the average price, we are implicitly assuming these consumers would buy the same amount at the average price. Thus their demand is weakly overestimated. However, by the same logic, assuming consumers who were charged higher than average prices instead faced the average price weakly underestimates their demand.

Thus, at first glance, the existence of a bias is obvious, but the direction is not. Others (Aguirre et al. [2010], Robinson [1933], Schmalensee [1981]) looking at the related question of whether discriminatory pricing yields higher or lower sales than optimally-set uniform prices find ambiguous results, seemingly confirming the ambiguity.

We show otherwise - when appropriately comparing sales under price discrimination to sales when uniformly charging the average price, rather than to sales at the *optimally-set* uniform price, this ambiguity disappears. The bias is always positive - demand is overestimated. It follows that the demand function's slope magnitude is biased upwards too. Note, measurement error bias is typically in the opposite direction, i.e. it usually lowers coefficient magnitudes.

Standard instrumental variable methods do not address these biases, due to the nature of the measurement error. One can interpret the measurement error as error in the quantity bought - it would be less if all consumers faced the average price. Instrumenting for price is therefore not a valid solution.[1]

Worryingly, these biases may go unnoticed, because holdout-sample validation tests typically use data which exhibit the same measurement error problems and thus may appear reassuring. Yet misleading conclusions, biased welfare estimates, and incorrect mapping of underlying mechanisms persist.

Moreover, the size of the bias may be growing, since internet sales enable more widespread and intense price discrimination. Shiller [2014], for example, demonstrates a large benefit to employing web-browsing data for price discrimination.

The dominant strategy for demand estimation essentially traces movement along the demand curve as supply shifts, but demand does not.[2] We propose a novel solution that instead uses the variation in the firm's perceived profit-maximizing prices across groups located at different points along the demand curve, in a given snapshot in time too short for new demand or supply shocks to arise.[3] The approach shares similarities with Guerre et al. [2000], who recover the distribution of underlying valuations from auction bids.

The method's intuition is as follows. Consumers in disaggregated data are grouped by the price they were offered. Once marginal costs are estimated, as explained below, one can find the slope of a group's demand function which implies the observed price and quantity bought by the group maximizes profits.[4] Hence we observe one point on each group's demand, the observed

[1]Denote: $\hat{Q} = Q + u$, where u is the measurement error, and $\hat{Q}$ denotes Q measured with error. If $Q = \alpha + \beta P + \epsilon$, then $\hat{Q} = \alpha + \beta P + (\epsilon + u)$. We show the mean of u is greater than zero and increasing in Q, problems which persist after instrumenting for price.

[2]Often instrumental variables are used to isolate movement along the demand curve.

[3]Another potential method for addressing the biases identified in this paper involves augmenting standard IV techniques. Essentially, one could instrument for each percentile of price offered to consumers. This method, however, has some major drawback - it requires a lot of data, and introduces selection issues.

[4]In discrete choice applications, the slope of demand is defined for expected demand.

price and quantity, and can recover an estimate of its slope locally. However, extrapolation requires an assumption on the curvature of demand, which is not known ex-ante. So demand curvature parameters are estimated by finding values which minimize the difference between (i) biased demand simulated from these group demand functions and (ii) biased demand estimated directly from aggregate data using standard instrumental variables techniques. After estimating curvature parameters, market demand is calculated by aggregating demand across groups.

In a second proof, we show that marginal costs can be inferred as usual from biased demand function estimated using aggregate data. This result implies that the bias from using overestimated demand and the bias from using its overestimated slope magnitude exactly offset when estimating marginal costs. The fact that these biases do not extend to cost estimates may be useful beyond their application in this paper.

Bias magnitudes are estimated using disaggregated Volvo S40 sale and price data from a Volvo dealership in China. We find demand is overestimated by around $5\% - 13\%$, and its slope magnitude by $4\% - 10\%$. We view this bias as large, especially since we also find only modest use of price discrimination, suggesting the bias may be a lot larger in other contexts.

A couple of papers tackle related questions. D'Haultfoeuille et al. [2014] address a more obvious bias by replacing manufacturer's list price with predicted transaction price from bargaining.[5] Manchanda et al. [2004] estimate physician demand for pharmaceuticals, modeling heterogeneity in physician preference and salesmens' ability to extract surplus.

This paper is organized as follows. Section 1 provides intuition for the biases and offers a formal proof. Section 2 presents the model designed to alleviate these biases. In Section 3, we describe the data and industry background. Sections 4 and 5 describe estimation and results. A brief conclusion follows.

[5] Chen et al. [2008] offer an approach similar to D'Haultfoeuille et al. [2014]

1 Biases in Standard Demand Estimation

In many economic fields, researchers infer the price consumers pay for a product by dividing revenues by quantity sold, thus implicitly assuming that all consumers are offered the sales-weighted average of prices charged under discriminatory pricing. In this section, we show that this simple fix introduces problems, yielding upwards biased estimates of quantity demanded and the demand curve's slope magnitude.

To prove this, we show that whenever the firm chooses to set separate prices to two potentially unequally sized submarkets, sales increase. This implies the more general finding for the case of many different prices - any final submarket classification can be reached from the overall market by repeatedly splitting submarkets in two.

Before presenting the main theorem, which requires no controversial assumptions, we develop some intuition using a toy model with the restrictive assumptions that (1) each group has linear demand, and (2) both groups buy the same quantity in equilibrium. We label the group charged a higher (lower) price as the strong (weak) market, with subscript "s" ("w").

In this over-simplified model, the weak market, which is charged a lower price, will have an inverse demand curve with flatter slope.[6] The flatter slope implies that as the firm moves from discriminatory prices to uniform pricing at the average price, the sales loss in the weak market exceeds the mitigating increase in the strong market. See Figure 1. In Appendix A, we show this holds generally for linear demand functions, even if groups purchase different quantities.

Now let's relax restrictive assumptions and suppose curved demand. A concave inverse demand curve, relative to a linear demand curve with tangency at the optimal discriminatory price, implies a greater reduction in sales at higher prices, and smaller increase in sales at lower prices. Hence, concave inverse demand increases the magnitude of the bias. See Figure 2. In

[6] Rearranging the elasticity formula yields $\frac{dP}{dQ} = \frac{P}{\epsilon Q}$. In this intuitive but over-simplified example, quantities are the same in both markets by assumption. The market with the lower price has larger in magnitude elasticity (according to the inverse elasticity rule), implying the right hand side of this equation is smaller in magnitude for the weak market than it is for the strong market. Hence, so too must be the left hand side $\left(\frac{dP}{dQ}\right)$.

the extreme, with perpendicular demand, the weak market buys none at $\bar{P}$, and the strong market buys the same as it would at higher price P_s^*. This reasoning bounds the magnitude of the bias at the sum of all sales occurring at prices below the sales-weighted average price.

Instead consider convex inverse demand curves. Relative to a linear inverse demand curve that is tangent at the optimal discriminatory price, a convex inverse demand curve implies lower reduction in sales when price increases and a bigger increase in sales when price decreases. See Figure 3. Hence, if one or both of the strong and weak markets exhibit inverse demand curves with convex shape, this might imply higher sales at $\bar{P}$ than under discriminatory pricing, changing the direction of the bias. However, it would also make charging $\bar{P}$ relatively more attractive in either market.

Using very mild assumptions, the next proof shows that if inverse demand curves are sufficiently convex for sales to be higher at $\bar{P}$ than at discriminatory prices, then those discriminatory prices cannot be profit-maximizing. Hence, if assuming rational firms, the existence of price discrimination precludes such shapes of inverse demand curves, proving the direction of the bias.

Theorem 1. *Assume continuous and downward sloping aggregate demand. Then, if total sales are greater at $\bar{P}$ than under discriminatory pricing, either P_w^*, or P_s^* (or both) must not be profit-maximizing.*

Proof. First note that the sales-weighted average price, $\bar{P}$, equals $\frac{P_w^* Q_w(P_w^*) + P_s^* Q_s(P_s^*)}{(Q_w(P_w^*) + Q_s(P_s^*))} \equiv \frac{R_{PD}}{Q_{PD}}$, where P_k^* and $Q_k(P_k^*)$ are the optimally set price and corresponding quantity purchased by group k, and R_{PD} and Q_{PD} denote revenue and quantity sold under discriminatory pricing. If the quantity sold when uniformly pricing at $\bar{P}$ exceeds the quantity sold under discriminatory pricing, then there exists an even higher uniform price $\widehat{P}$ which yields the same total quantity sold as under price discrimination. But then, profits from uniformly charging $\widehat{P}$ must exceed profits from discriminatory pricing: $\pi_{uniform} \equiv \widehat{P} Q_{PD} - C(Q_{PD}) > \bar{P} Q_{PD} - C(Q_{PD}) \equiv \frac{R_{PD}}{Q_{PD}} Q_{PD} - C(Q_{PD}) = R_{PD} - C(Q_{PD}) \equiv \pi_{PD}$.

If instead the quantities are the same in both cases, i.e. $Q_{PD} = Q(\bar{P})$, then uniform pricing can yield revenues at least as high as price discrimination. If we additionally assume implementing price discrimination implies epsilon > 0 marginal costs, uniform pricing would again yield strictly higher profits. Price discrimination can therefore raise profits only if $Q_{PD} > Q(\bar{P})$.

□

Theorem 1 tells us that if we observe discriminatory prices, and are willing to assume that the firm's decision to price discriminate raises profits, then falsely assuming all consumers pay the average price biases upwards estimated sales at that price.

It follows that the slope estimate is biased as well. First note that both the biased and true aggregate demand curves intersect at the demand choke price - when the firm sells one unit to the consumer with highest valuation, there is no bias created by aggregation. Because the biased inverse demand curve must also connect with the upward-biased demand at the sales-weighted average price, the biased inverse demand curve must be flatter than the true aggregate inverse demand curve, at least on average. See Figure 4. A flatter inverse demand function signifies a steeper demand function, implying the estimated absolute slope of demand is biased upwards.

2 Model

In this section, we first demonstrate one can estimate marginal costs even when demand is biased from overlooking price discrimination. Then, in subsection 2.2, we explain a way to infer overall demand in a given snapshot in time from the set of prices the firm charges to different market segments, the quantities sold to each segment, and estimated marginal cost. Average demand over time can then be computed by averaging demands from the different snapshots in time.

2.1 Estimating Marginal Costs

Theorem 2. *Assume $\frac{d\bar{P}}{dQ}$, i.e. the slope of the biased inverse demand curve which does not account for price discrimination, equals the ratio of changes in sales-weighted average prices to changes in total quantity sold arising from changes in costs. Then, the marginal cost in period s can be inferred from the sales-weighted average price paid in period s ($\bar{P}$), the total quantity sold in period s (Q), and $\frac{d\bar{P}}{dQ}$. Specifically, marginal cost C' equals: $C' = \frac{d\bar{P}}{dQ} * Q + \bar{P}$.*

Proof. The slope of the biased inverse demand equals: $\frac{d\bar{P}}{dQ} = \frac{\frac{d\bar{P}}{dC'}}{\frac{dQ}{dC'}}$. We find the numerator and denominator of this complex fraction separately.

We start with the numerator. First, recall that $\bar{P} = \frac{\sum_k R_k}{\sum_k Q_k}$, where R_k and Q_k denote the revenues and total sales from segment k, respectively. By the quotient rule, the derivative with this substitution becomes:

$$\frac{d\bar{P}}{dC'} = \frac{\left(\frac{d}{dC'}\sum_k R_k\right)\left(\sum_k Q_k\right) - \left(\frac{d}{dC'}\sum_k Q_k\right)\left(\sum_k R_k\right)}{\left(\sum_k Q_k\right)^2} \tag{1}$$

By the chain rule, $\frac{dR_k}{dC'} = \frac{dR_k}{dQ_k}\frac{dQ_k}{dC'}$. $\frac{dQ_k}{dC'}$ gives the change in profit-maximizing sales volume for segment k as costs change. $\frac{dR_k}{dQ_k}$, i.e. the marginal revenue for segment k, should, assuming profit maximization, equal C'. Plugging these equivalences into equation 1 and pulling out C' yields:

$$\frac{d\bar{P}}{dC'} = \frac{C'\left(\sum_k \frac{dQ_k}{dC'}\right)\left(\sum_k Q_k\right) - \left(\sum_k \frac{dQ_k}{dC'}\right)\left(\sum_k R_k\right)}{\left(\sum_k Q_k\right)^2} \tag{2}$$

Using $Q = \sum_k Q_k$, $\bar{P} = \frac{\sum_k R_k}{Q}$, and simplifying further yields:

$$\frac{d\bar{P}}{dC'} = \left(\frac{C'}{Q} - \frac{\bar{P}}{Q}\right) * \left(\frac{dQ}{dC'}\right) \tag{3}$$

That completes solving for the numerator of the complex fraction $\frac{\frac{d\bar{P}}{dC'}}{\frac{dQ}{dC'}} = \frac{d\bar{P}}{dQ}$. To get $\frac{d\bar{P}}{dQ}$, we divide the numerator, equal to equation 3, by the denominator $\frac{dQ}{dC'}$, yielding:

$$\frac{d\bar{P}}{dQ} = \left(\frac{C'}{Q} - \frac{\bar{P}}{Q}\right) \tag{4}$$

Lastly, rearrange to solve for C':

$$C' = \frac{d\bar{P}}{dQ}Q + \bar{P} \tag{5}$$

□

Note that one can rearrange Equation 5 to yield the inverse elasticity rule with price replaced with $\bar{P}$. This is intuitively surprising. The inverse elasticity rule is derived from profit maximization of a single price, but the price in Equation 5, $\bar{P}$, is not the price the firm is choosing to charge each consumer.[7] Rather it is the sales-weighted average of the optimal prices for each segment. Moreover, $\frac{d\bar{P}}{dQ}$ is not the slope of demand, but a biased estimate of the slope of demand.

Marginal costs are estimated as follows. First, one estimates $\frac{dQ}{dP}$ using two-stage least squares with cost-based instruments and finds its reciprocal. This yields the biased inverse demand slope estimated by improperly assuming all consumers are offered the average price. Equation 5 can then be used to estimate marginal cost.

The next subsection explains how to use this information to estimate aggregate demand.

2.2 Unbiased Demand Estimation with Known Marginal Cost

We begin by showing that one can estimate the slope of demand for group k in period t from the price $P_{k,t}$, quantity $Q_{k,t}$, and marginal cost C'_t, where k includes all consumers charged

[7] This does not imply that the elasticity of demand is unbiased, unless by chance the profit-maximizing uniform price happens to equal the weighted average price under price discrimination.

price $P_{k,t}$ in period t. We later discuss combining group demands.

The next few paragraphs show the method is valid even when group k is comprised of smaller subgroups i for which the firm separately optimizes prices, prices which just happened to be the same. Therefore, the researcher only needs to observe the price charged, not the granularity of subgroups in the firm's pricing algorithm.

First note that the slope of demand for group k, who are all charged price P_k, is:

$$\frac{dQ_k}{dP_k} = \sum_{i \in k} \frac{dQ_i}{dP_k} \tag{6}$$

where Q_i is the quantity bought by subgroup i. i could be an individual, or a subgroup with easily observable characteristics, e.g. men. Different subgroups may be of different sizes.

Since all subgroups i in group k have the same profit maximizing price, they must have the same elasticity. Substituting in the elasticity E_k, we can rewrite Equation 6 as:

$$\frac{dQ_k}{dP_k} = \sum_{i \in k} E_k \frac{Q_i}{P_k} \tag{7}$$

Because the elasticity and price are the same for all subgroups i in group k, we can pull them out of the sum, yielding:

$$\frac{dQ_k}{dP_k} = \frac{E_k}{P_k} \sum_{i \in k} Q_i = \frac{E_k Q_k}{P_k} \tag{8}$$

Lastly, simplifying using the inverse elasticity rule yields:

$$\frac{dQ_k}{dP_k} = \frac{Q_k}{C' - P_k} \tag{9}$$

Thus, we can solve for the slope of segment k's demand, since all parts on the right hand side are known - P_k and Q_k are observed directly, and the marginal cost can be estimated as shown in Section 2.1. Hence, we can recover one point on group ks demand curve and its slope at that point.

Complications arise when combining group demands. For market segment k charged price P_k, we know how much that group purchased at that price. But we do not know how much another market segment j charged different price P_j would have purchased at the price offered to segment k. At no price can we directly observe all segment's purchase decisions.

Therefore, when price discrimination exists, the only way to reach an estimate of aggregate demand at any given price of interest is by extrapolating demand for each segment. Usually at least for some segments, and often for most, this will entail extrapolating demand outside of the range of prices offered to the segment. It therefore necessitates functional form assumptions.

We can, however, use additional information to guide the choice of functional form assumptions for extrapolations. Specifically, we can first estimate an unbiased demand curve for each group by the above method, under a specific functional form assumption. We can then simulate biased demand by the steps in the following paragraph. Simulated biased demand curve should coincide with the biased demand curve estimated by ignoring latent price discrimination in aggregate data. We can therefore choose functional form assumptions for group demand's which make them so.

We simulate biased aggregate demand from groups' unbiased demand estimates by the following steps. First, for a given marginal cost, we compute the optimal price to charge each group and the expected quantity purchased. Then we combine across groups to calculate the total quantity and sale's weighted average price. This gives one point along the simulated biased demand curve. Repeating for a vector of marginal costs yields the entire simulated biased demand curve. Further details on implementation are provided in Section 4.

3 Data and Background

3.1 "4S dealerships" in China

In China, there is only one Volvo general distributor, but 116 certified dealerships, known as "4S dealers." A "4S dealer" (or "certified brand dealer") provides full services to consumers including **S**ale, **S**pare-parts, **S**ervice and **S**urvey feedback (i.e. 4**S**). The arrangement is mutually exclusive in the sense that a certified dealer of a particular brand can only sell cars and parts of its brand, and the car manufacturer only provides cars and parts to the certified dealers.[8]

The distributor, on authority of the manufacturer, coerced its dealers not to compete with each other. Each dealer was assigned a "responsible area," which was typically one city or a part of a metropolitan area. The "conduct regulation of responsible area and relevant fines" section of the "dealer commercial policy" (a binding contract) stated that "Dealers should NOT be seeking sales and service in other dealers' 'responsible areas' OR (sales and service) within 5 kilometers of the other dealer's facilities/stores (if shares the same responsible area)."[9] The contract also stipulated punishments for violating this rule. If a dealer's inter-provincial sales ratio was greater than 10% of its total sales then it would lose 0.2% rebate of its total sales value.[10] Moreover, the general distributor had the right to revoke the certification of the dealer if the dealer violated this rule. The contract also encouraged cooperation: if dealers in one or multiple provinces offered "Sales Events" together, all participating dealers would get 0.5% rebate. Additional market power arose from consumers' reluctance to purchase from a distant dealer over fears of whether their warranty would be honored by local dealers.

The dealers not only had market power for new car sales, but also for service and parts. Only the certified dealers could sell parts produced by the original manufacture. Unable to

[8]Since 2005, the "4S dealership framework" has been officially enforced by "Administration of Automobile Brand Sales Implementing Procedures," a federal act. On Oct 1st, 2014, they stopped authorizing new certified dealers, and are believed to in the near future allow dealers to sell multiple car brands.

[9]Translated from Chinese-language contract.

[10]The car's plate region is used to identify whether a sale is local or inter-provincial. A dealer's inter-provincial sales ratio is calculated as the accumulative inter-provincial sales over last quarter/ total sales last quarter.

buy official Volvo parts, un-certified garages could only use "deputy-plant parts" produced by un-authorized producers.[11] If a 4S dealer found any "deputy-plant parts" during a car's warranty period, the car's warranty automatically expired.

3.2 Data

Our data are provided by a well-established certified Volvo dealer in the Jiangsu Province of China. There are 13 other similar dealers in the Province, owned by 7 different companies, but none nearby. See Figure 5.

The data contain a host of information about each car the dealer sold between July 2006 and July 2010, including the Vehicle Identification Number (VIN), model, trim, date it arrived on the lot, and cost to the dealer. Importantly, the data also include each car's final sale price and date sold. Inventory levels are constructed from the arrival dates.

We restrict our attention to the Volvo S40, the best-selling Volvo model. There were five separate trims: S40 2.4, S40 T5, S40 2.5T, S40 2.5, and S40 2.0. The major differences between these trims are their engine specifications. 737 Volvo S40s were sold between July 2006 to July 2010.[12,13] Models S40 2.4, S40 T5 and S40 2.5T account for most of the sales, and the dealer gets on average 2.31%, 6.13%, and −1.05% profit rates on each sale.[14] See Table 1. The profit rate may seem low, but the dealer also received from the manufacturer a rebate worth $5-11\%$ of the invoice value.[15] All but one of the cars in the data were assembled in China, but 60% of

[11]Un-authorized parts producers can market parts with phrasing like "this part suits S40", but the parts cannot have a Volvo Logo on it without breaking the law.

[12]We dropped four observations, three which implied 25.2%, 25.2% and 29.6% losses. We were informed cars sold for large losses are likely either politically motivated or for cars used for test-driving. The fourth dropped observation had a final sale's price of 363 thousand Yuan, while several others of the exact same model sold for 263 thousand Yuan, suggesting there was a recording error.

[13]Two months had missing sales, May and June 2010.

[14]The profit rate is calculated as final sale price minus the cost of the car to the dealer.

[15]The distributor retains most of the dealer's profits for each quarter and month, and later pays the dealers in the form of incentives-based rebates. Mature dealers (dealerships established for more than 12 months) received rebates worth between $5-11\%$ of the total invoice value of new car sales. New dealers received higher rebates, worth $6-11\%$ of the total invoice value.

the parts by value were imported, mainly from Sweden.[16,17]

The data confirm that the dealer's aquisition costs in Chinese Yuan were somewhat sticky, suggesting the dealer paid the same price for a given quality car over long periods. See Figure 6, which shows the cost vs. date car arrived on the lot for the most popular the trim, the S40 2.4.

3.3 Evidence of Price Discrimination

Price discrimination is defined by two characteristics. First, consumers pay different prices for similar items. Second, price differences are not merely due to cost differences, but instead reflect differences in consumers' reservation prices.

Figure 7 provides illustrative evidence suggesting the dealer employed price discrimination when selling Volvo S40s. The points in the figure denote residuals from a kernel regression of the price on (1) the cost of the car to the dealer, and (2) date.[18] These explanatory variables are designed to flexibly capture price differences due to differences in costs across cars with different features and across time. Residual differences in prices presumably reflect price discrimination.[19]

Figure 7 shows that price differences due to price discrimination are rarely greater than 4% of the average price. Hence, any biases in demand estimates demonstrated in later sections apply to a market with relatively meager use of price discrimination. Biases thus might be much larger in other contexts.

[16] We can determine the country where the car was manufactured using the Vehicle Identification Number's (VIN) World Manufacturing Identifier (WMI). All but one of the model S40 cars in the data had VINs beginning with "LVS," indicating the car was manufactured in Chang'an Automobile Co.Ltd's plant. The plant began producing S40s in 2006

[17] The domestic-value percentage ratio (value of domestic parts / total value of parts) was obtained from technicians working for the dealer.

[18] We used a uniform kernel over an ellipse, where the diameter in each direction, the bandwidths, where optimized using leave-one-out cross-validation.

[19] Since dynamic pricing or seasonal variation in prices could be from price discrimination, actual price discrimination might be more intense than suggested by Figure 7.

3.4 Accounting for Minor Quality Differences

The presence of quality differences among trims of Volvo S40s introduces complications. We use the simplifying assumption that consumers' incremental valuations for higher quality cars is exactly captured by differences in dealer costs. Specifically, we subtract from each trim's selling price the difference between it's cost to the dealer and the lowest of these costs that year. This seemed a reasonable method for controlling for minor quality differences.[20] For a robustness check, we rerun the main regressions using a second method for adjusting prices for quality differences.[21]

4 Estimation

In this section, we begin by reviewing in outline form the steps in the proposed model. We then provide implementation details. Note "aggregate data" refers to monthly data which include total sales and average prices, and "disaggregated data" refers to data which include monthly sales and prices by group, where group is determined by the purchase price.

4.1 Outline

In Aggregate Data

1. Estimate biased demand $Q = f(\bar{P})$ using cost-based instruments. Record its inverse $\bar{P} = g(Q)$

[20] Dealer costs seem to be quite sticky, as shown in Figure 6, suggesting cost differences reflect only differences in product attributes. Influences of any randomness in the lowest of these costs, if it exists, will be mitigated in later regressions by the inclusion of year fixed effects.

[21] We first regress the sale's price on the dealer's cost and year and month fixed effects. The corresponding coefficient on dealer's cost suggests that for each additional Yuan spent on increasing quality, willingness to pay on average increases by 1.127 Yuan. As a robustness check, we adjust prices for quality differences by subtracting $1.127 * (\textit{arrival price of car} - \textit{lowest arrival price that year})$. The main regression results, shown in Tables 3 and 6, show insubstantial differences between the main specification and the second one.

2. Infer marginal costs (C'_t) by plugging the reciprocal of $\frac{df(\bar{P})}{dP}$ obtained from step 1 into Eq. 5

<u>In Disaggregated Data</u>

3. Find $\frac{dQ_{k,t}}{dP_{k,t}}$ for each specific group k and time t by plugging C'_t from step 2 into Eq. 9

4. Choose initial guess at parameter γ determining curvature of group demand functions

5. Extrapolate demand for each group k & period t using $\frac{dQ_{k,t}}{dP_{k,t}}$ from step 3 and γ

6. Simulate biased inverse demand $\bar{P} = g^{SIM}(Q)$: find one point on the inverse demand curve by averaging predicted optimal prices ($P^*_{k,t}$) and aggregating corresponding sales ($Q_{k,t}(P^*_{k,t})$) for an assumed marginal cost (C'_t). Repeat for vector of C'_t to yield entire curve

7. Calculate the value of the objective function: $obj = \sum \left(g^{SIM}(Q) - g(Q)\right)^2$

8. Repeat steps 5-7 searching over γ until objective function minimized

4.2 Implementation Details

In step 1, we regress the log of quantity sold in a month on the log of average price in order to estimate the biased inverse demand function and by extension its slope at any given price.[22] Due to the usual endogeneity concerns, we instrument for log price using two cost based instruments: (1) lagged dealer inventory levels of Volvo S40s, and (2) the Swedish Krona/Chinese Yuan exchange rate at time of sale.[23] The intuition behind these instruments is explained below.

[22] We tried linear demand, but found it led to strange results. In particular, the linearity assumption results in a severe underestimation of demand at high prices. This led to substantial problem in later steps - there was no level of assumed curvature for group demands that could come close to rationalizing the simulated biased demand and the biased demand estimated via regression.

[23] It is necessary to use cost based instruments because the proof for Theorem 2 was based on price and quantity changes arising from cost changes. Theorem 2 may not apply for level-of-competition based instruments.

Inventory of the model impacts the opportunity cost of selling a car on the lot (Copeland et al. [2011]). We expected higher inventory levels would, through its impact on opportunity costs of selling a car, imply lower marginal costs, and thus inventory should be negatively related to price.[24] However, one might worry that transient demand shocks may also influence current inventory, and thus current inventory may not be a valid instrument. We therefore take the lagged inventory level, from the previous month, which influences the inventory state at the beginning of the month, but is not influenced by transient demand shocks.

The exchange rate is also a good instrument, for the following reason. In Section 3 we showed that for the S40, which is manufactured in China, the cost of the car to the dealer is quite sticky in the short run. However, the exchange rate between China and Sweden has an influence on the long-run cost of parts, since 60% of the parts were manufactured in Sweden.[25] Since the dealership had market power in the repair market, the exchange rate influences future expected costs of repair, which impacts its profits from servicing its previous customers. As the exchange rate, defined as $\frac{Swedish\ Krona}{Chinese\ Yuan}$, increases, the Yuan appreciates and costs of parts for repair therefore falls, increasing profit to the dealer from previously sold cars. Hence, the higher the exchange rate, the greater the incentive for the firm to sell cars, which it can do by lowering the initial purchase price.

We would thus expect a negative relationship between the exchange rate at time of sale and price markup over dealer's cost. Table 2 confirms this relationship. The table shows regressions of markups, i.e. selling price minus dealer's acquisition cost, on the exchange rates at different points in time: (1) when the car was pre-ordered by the dealer, (2) when the car arrived on the lot, and (3) when the car was sold. Consistent with our assertion, the exchange rate at time of sale has by far the strongest and most significant influence on the markup over the direct cost of the car.

Including these instruments yields the results shown in Table 3. When instrumenting

[24]Unlike Copeland et al. [2011], we found squared inventory was not a significant or meaningful predictor of the sale's price, implying a linear rather than hump-shaped relationship between inventory and marginal costs.

[25]See Section 3.2.

for price and including month and year fixed effects, the coefficient on logged average price is significant at the 10% level. Its point estimate implies that a one percent increase in price reduces sales by 7.7%. However, recall that this is an upward biased estimate of the magnitude.

Next, we can find estimates of marginal costs by inverting the regression results to yield the inverse demand function, finding its derivative $\left(\frac{d\bar{P}_t}{dQ_t}\right)$, and plugging into Equation 5 to yield an estimate of the marginal cost in a given month. Specifically, if we denote the log-log regression from column vi of Table 3 as $\ln(Q_t) = \alpha_t + \beta \ln(\bar{P}_t)$, where α_t is the constant inclusive of month and year fixed effects, then:

$$\frac{d\bar{P}_t}{dQ_t} = \frac{\bar{P}_t}{\beta \exp(\alpha_t + \beta \ln(\bar{P}_t))} \tag{10}$$

Plugging the resulting values of $\frac{d\bar{P}_t}{dQ_t}$, the total monthly sales (Q_t), and average prices ($\bar{P}_t$) into Equation 5 yields the monthly marginal costs shown in Figure 8.

Figure 8 also shows the dealer's cost for a "base-level" car.[26,27] Note that the marginal cost estimates fall below the actual cost to the dealer of acquiring the car. This is because the dealer's direct cost for the car does not include opportunity costs nor incentives provided by the distributor. As a contact at the dealership explained, a large share of its profits come from servicing of previously sold cars. Factoring in this opportunity cost lowers the relevant marginal cost.

[26] The base-level car is assumed to be the lowest cost car in a given year.

[27] The observed noise in marginal costs in Figure 8 suggests there is an errors-in-variables problem, which was expected in a dataset with relatively few sales. The first-order profit maximization conditions, used to estimate marginal cost each period according to Equation 5 and later to estimate demand curves for specific groups, should apply only for the firm's ex-ante expectations of the quantity they will sell at chosen prices. But only ex-post realized sales are observed. This point is most apparent in a discrete choice context. Each individual that enters the dealership is expected to buy with some probability between 0 and 1, but is only observed buying 1. Hence, ex-post realized sales at a given price, which are observed, are random and may differ from expected sales in small samples. We tried attenuating these errors-in-variables problems by smoothing both the marginal costs and quantities sold at each price, implicitly imposing smoothness priors for both. Specifically, we first used local linear regressions to yield a smoothed marginal cost estimate for each month based on X neighboring months. We then completed an analogous procedure in the disaggregated data to smooth the quantities sold in each price range. Results using these smoothed values differed insubstantially from non-smoothed results, and were thus omitted from the paper.

Figure 8 also shows that marginal costs tended to decline over time. This is due to two factors. First, as the figure shows, the dealer's direct cost of acquiring the car falls over time. Second, the dealership, which opened in the first period of our data, likely learned over time how profitable post-sale servicing was, thus increasing perceived opportunity costs of not selling an additional car.

Next, using disaggregated data, we infer the slope of demand for a specific group $\left(\frac{dQ_{k,t}}{dP_{k,t}}\right)$ at the observed point on the demand curve, i.e the point corresponding to the price for each group $(P_{k,t})$ and total quantity bought by the group $(Q_{k,t})$. We form groups by each month sorting consumers into K price ranges of equal width centered at the average price. If assuming K price groups each month and T time periods, then we have $K * T$ total groups. For each, we compute the total sales $(Q_{k,t})$ and average price $(P_{k,t})$. We can then as desired estimate $\frac{dQ_{k,t}}{dP_{k,t}}$ for each group by plugging in $Q_{k,t}$, $P_{k,t}$, and the estimated marginal costs C'_t into Equation 9.

Having estimated the slopes at one point on each group's demand curve, we can extrapolate a given group's demand to arbitrary prices, based on an assumption of the demand function's curvature. We use the following functional form to allow for nonlinear inverse demand functions:

$$P_{k,t} = \delta_{k,t} + \lambda_{k,t} Q_{k,t}^{\gamma} \tag{11}$$

Where $\delta_{k,t}$, $\lambda_{k,t}$, and γ are parameters. With an assumed value of γ, we can infer the values of $\delta_{k,t}$ and $\lambda_{k,t}$, and hence the entire inverse demand curve for a specific group.[28] By summing group demand functions over k and averaging over t, we can yield an estimate of the unbiased aggregate demand curve in the average month.

We can then simulate the biased demand function from the individual group demands by reaggregation, i.e. simulating the process that causes bias in the first place. Based on our group demand estimates, we can calculate the profit-maximizing price for each group $P^*_{k,t}(C'_t)$

[28] $\frac{dP_{k,t}}{dQ_{k,t}} = \gamma \lambda_{k,t} Q_{k,t}^{\gamma-1}$. Hence, $\lambda_{k,t} = 1/\left(\gamma \frac{dQ_{k,t}}{dP_{k,t}} Q_{k,t}^{\gamma-1}\right)$. Then: $\delta_{k,t} = P_{k,t} - \lambda_{k,t} Q_{k,t}^{\gamma}$.

and the corresponding quantity sold $Q_{k,t}(C_t')$ for any arbitrary marginal cost.[29] We then aggregate. The total corresponding *per-period* demand is: $Q^{Sim\ Biased}(C_t') = \frac{\sum_k \sum_t Q_{k,t}(C_t')}{T}$, and the average price is: $\bar{P}(C_t') = \frac{\sum_k \sum_t P^*_{k,t}(C_t')Q_{k,t}(C_t')}{\sum_k \sum_t Q_{k,t}(C_t')}$. These valuees $(Q^{Sim\ Biased}(C_t'), \bar{P}(C_t'))$ comprise one point on the simulated biased inverse demand curve. Repeating these steps for a series of marginal costs allows one to construct the entire simulated biased inverse demand curve: $\bar{P} = g^{SIM}(Q)$.

We begin by assuming linear demand for each group, i.e. $\gamma = 1$, and assuming $K = 5$ price groups per period. Figure 9 shows three inverse demand functions under this assumption: (1) simulated *un*biased, simulated biased, and *un*simulated biased, the last of which is from column vi of Table 3. If the model fits well, the simulated and unsimulated biased inverse demand curves should be similar. Clearly they are not. The simulated biased demand does have a convex shape, due to different groups of consumers having different demand intercepts, but it is clearly not as convex as the inverse demand function that was estimated directly from aggregate data and which does not rely on the assumed value of γ.

Thus, we need to relax the assumption of linear group demands, and search for the level of convexity which yields a biased simulated inverse demand curve which coincides with the unsimulated equivalent. Specifically, we search over γ for the value which minimizes the squared difference:

$$obj(\gamma) = \sum_{Q} \left(g^{SIM}(Q) - g(Q) \right)^2 \qquad (12)$$

Where $g^{SIM}(Q)$ and $g(Q)$ are the simulated and unsimulated biased inverse demand functions, and the squared differences are summed over increments of 0.1 between the minimum and maximum number of monthly car sales observed in the data. As Figure 10 shows, the optimized value of $\gamma = 0.158$ yields a simulated biased inverse demand that is quite close to the unsimulated equivalent.

[29] The profit-maximizing price to charge group k in period t is: $P^*_{k,t}(C_t') = \delta_{k,t} + \frac{C_t' - \delta_{k,t}}{\gamma+1}$.

5 Results

In this section, we present and discuss the biases apparent from the estimated inverse demand functions obtained in Section 4.

We could directly compare the unbiased demand estimates with the unsimulated biased demand estimates obtained from standard IV regression methods in aggregate data. However, this could be misleading because the difference includes small noise from functional form assumptions used for the group demands and for the initial IV regressions. This noise is apparent from the small differences between the unsimulated biased demand estimates yielded from IV methods in aggregate data and the biased demand estimates simulated from the groups' unbiased demand estimates, shown in Figure 10. We therefore instead focus on a more exact comparison, unbiased aggregate demand vs. simulated biased demand reached by simulating the aggregation process.

The bias in the amount demanded at a given price $\hat{P}$ is computed as follows. First, we calculate analytically the unbiased estimate of aggregate demand at price $\hat{P}$ in the average month. Specifically, we sum the amount demanded $\left(Q_{k,t}(\hat{P})\right)$ across groups k and time t and divide by number periods T to yield the average aggregate monthly demand. We then compute the analogous amount demanded at that same price $(\hat{P})$ according to the simulated biased demand function. To do so, we first numerically search for the marginal cost $(\hat{C}')$ that implies a simulated sales-weighted average price equal to the price of interest $\hat{P}$, assuming the firm is optimally price discriminating. We then use that marginal cost $(\hat{C}')$ to find the amount demanded at average price $\hat{P}$ - the aggregate amount demanded equals the sum across k and average across t of the group demands when prices are set to maximize profits separately for each group at marginal cost $\hat{C}'$. I.e. it equals: $\frac{\sum_k \sum_t Q_{k,t}(P^*_{k,t}(\hat{C}'))}{T}$. The bias in then calculated by subtracting the unbiased estimate of the amount demanded at price $\hat{P}$ from its biased equivalent.

The bias in the slope of demand is calculated similarly. We first compute the unbiased estimate of the slope of average aggregate monthly demand. Specifically, we compute the corresponding slope of demand at price $\hat{P}$ for each group k and time t, then sum across k and average across t. Then we compute the analogous slope according to the simulated biased demand. To compute it, we first compute its reciprocal $\frac{d\bar{P}}{dQ}$ by plugging the following variables into Equation 4: the price $\hat{P}$, the corresponding amount demanded at marginal cost $\hat{C}'$, and $\hat{C}'$. We then again take the reciprocal to yield the simulated biased value of $\frac{dQ}{d\bar{P}}$. The bias in then calculated by subtracting the unbiased estimate of the slope of aggregate demand at price $\hat{P}$ from its biased equivalent.

The bias was computed for various prices of interest. We chose five prices, which correspond to the 10th, 25th, 50th, 75th, and 90th percentiles of prices on the unsimulated biased aggregate demand curve, when restricted to the segment with quantities which fall between the minimum and maximum monthly sales observed in the data.

Table 4 shows that the percent bias in the amount demanded depends on the location along the demand curve.[30] Within the interdecile range of prices, point estimates of the percent bias range from 5.51% at the 10th price percentile to 13.84% at the 90th price percentile. Standard errors, computed via the delta method, are relatively large due to propagation of the large standard errors in the initial IV estimates from aggregate demand, shown in column vi of Table 3.

Table 5 shows the corresponding bias in the slope of estimated demand. Again, we see relatively large estimates of the bias, especially considering differences in prices due to price discrimination were small. Specifically, within the interdecile range of prices, point estimates of the percent bias range from 3.97% at the 10th price percentile to 10.64% at the 90th price percentile.

[30] We assume $K = 5$ price groups per month. We found that further increasing to ten price ranges yielded similar results.

6 Conclusion

In this paper, we first proved that overlooking price discrimination and assuming all consumers pay the average price yields upwards biased estimates of demand and its slope magnitude. We then found that, surprisingly, these biases offset each other when computing marginal costs from optimal pricing conditions. Hence we can still compute unbiased estimates of marginal costs even with the biased demand.

We then provided a method that uses estimated marginal costs and disaggregated data to yield an unbiased estimate of aggregate demand. Using this technique, we found that, in the context of Volvo S40 sales in China, standard instrumental variable techniques overestimate demand by 5% to 13%, and its slope magnitude by 4% to 10%.

References

Iaki Aguirre, Simon Cowan, and John Vickers. Monopoly price discrimination and demand curvature. *The American Economic Review*, 100(4):pp. 1601–1615, 2010. ISSN 00028282. URL `http://www.jstor.org/stable/27871267`.

Yuxin Chen, Sha Yang, and Ying Zhao. A simultaneous model of consumer brand choice and negotiated price. *Management science*, 54(3):538–549, 2008.

Adam Copeland, Wendy Dunn, and George Hall. Inventories and the automobile market. *The RAND Journal of Economics*, 42(1):121–149, 2011.

Xavier D'Haultfoeuille, Isis Durrmeyer, and Philippe Fèvrier. Automobile prices in market equilibrium with unobserved price discrimination. 2014.

Emmanuel Guerre, Isabelle Perrigne, and Quang Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000.

Puneet Manchanda, Peter E Rossi, and Pradeep K Chintagunta. Response modeling with nonrandom marketing-mix variables. *Journal of Marketing Research*, 41(4):467–478, 2004.

Joan Robinson. Economics of imperfect competition. 1933.

Richard Schmalensee. Output and welfare implications of monopolistic third-degree price discrimination. *The American Economic Review*, pages 242–247, 1981.

Benjamin Shiller. First-degree price discrimination using big data. *Working Paper*, 2014.

A Proof of Bias for Linear Demand

Proposition 1. *Suppose two markets with linear demand functions, $Q_s(P_s)$ and $Q_w(P_w)$. Denote P_w^* and P_s^* as the optimal discriminatory prices in each market, respectively, and $\bar{P} = \frac{P_w^* Q_w(P_w^*) + P_s^* Q_s(P_s^*)}{Q_w(P_w^*) + Q_s(P_s^*)}$, as the sales-weighted average price. Then, without loss of generality, if $P_s^* > P_w^*$, then $Q_w(P_w^*) + Q_s(P_s^*) > Q_w(\bar{P}) + Q_s(\bar{P})$*

Proof. Denote the change in quantity sold when moving from optimal discriminatory prices (P_s^*, P_w^*) to the sales-weighted average price $(\bar{P})$ as $\Delta Q_{PD \longrightarrow u}$. With linear demand, it exactly equals the change implied by the first order Taylor expansion:

$$\Delta Q_{PD \longrightarrow u} = Q'_w(P_w^*)(\bar{P} - P_w^*) + Q'_s(P_s^*)(\bar{P} - P_s^*).$$

Plugging in for $\bar{P}$ and simplifying yields:

$$\Delta Q_{PD \longrightarrow u} = (Q'_w(P_w^*) * Q_s(P_s^*) - Q'_s(P_s^*) * Q_w(P_w^*)) \left(\frac{P_s^* - P_w^*}{Q_w(P_w^*) + Q_s(P_s^*)}\right).$$

Multiplying both sides by $\frac{P_w^* P_s^*}{Q_w(P_w^*) Q_s(P_s^*)}$, a positive number, yields:

$$\Delta Q_{PD \longrightarrow u} * \frac{P_w^* P_s^*}{Q_w(P_w^*) Q_s(P_s^*)} = (\epsilon_w(P_w^*) * P_s^* - \epsilon_s(P_s^*) * P_w^*) \left(\frac{P_s^* - P_w^*}{Q_w(P_w^*) + Q_s(P_s^*)}\right).$$

where ϵ denotes the elasticity of demand. The fraction in latter parentheses, $\frac{P_s^* - P_w^*}{Q_w(P_w^*) + Q_s(P_s^*)}$ is greater than zero. So the proof depends on whether $\epsilon_w(P_w^*) * P_s^* - \epsilon_s(P_s^*) * P_w^*$ is negative. It is.

By the inverse elasticity rule, we know $|\epsilon_w(P_w^*)| > |\epsilon_s(P_s^*)|$, and hence $|\epsilon_w(P_w^*)|P_s^* > |\epsilon_s(P_s^*)|P_w^*$. Since demand elasticities are negative, so is the expression. $\square$

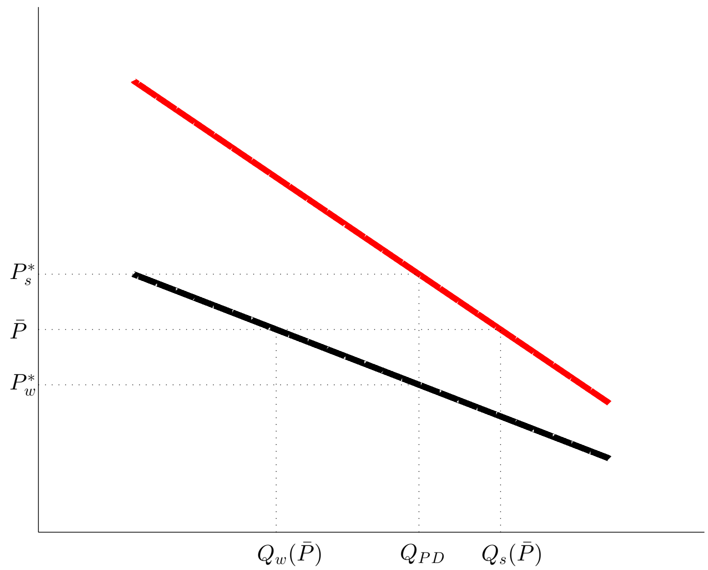

Figure 1: Illustration of Bias With Linear Inverse Demand Curves

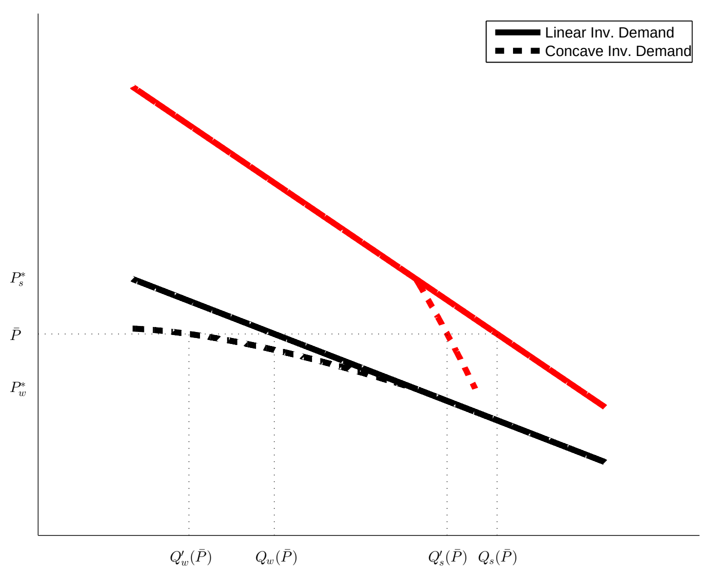

Figure 2: Illustration of Bias With Concave Inverse Demand Curves

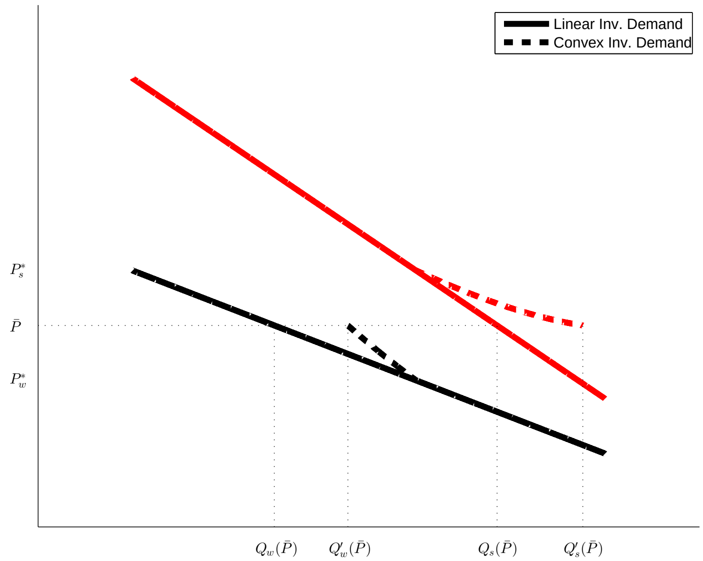

Figure 3: Illustration of Bias With Convex Inverse Demand Curves

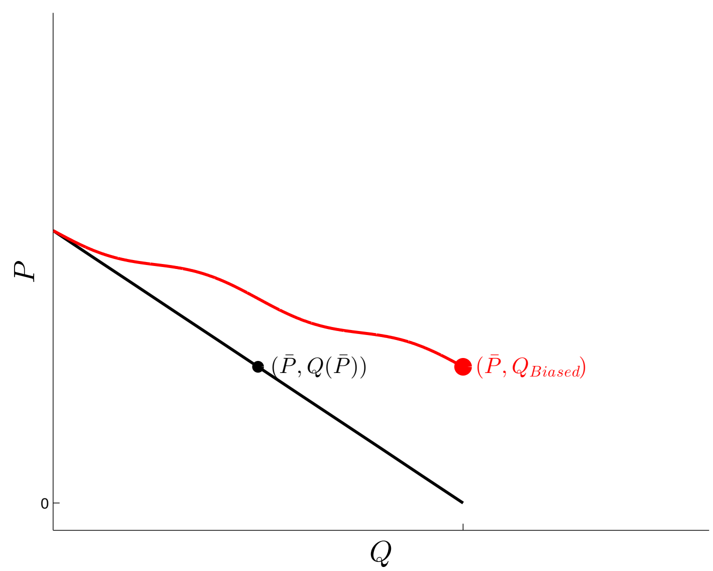

Figure 4: Graphical Example - Bias Can Vary at Points, Depending on True Shape of Inverse Demand

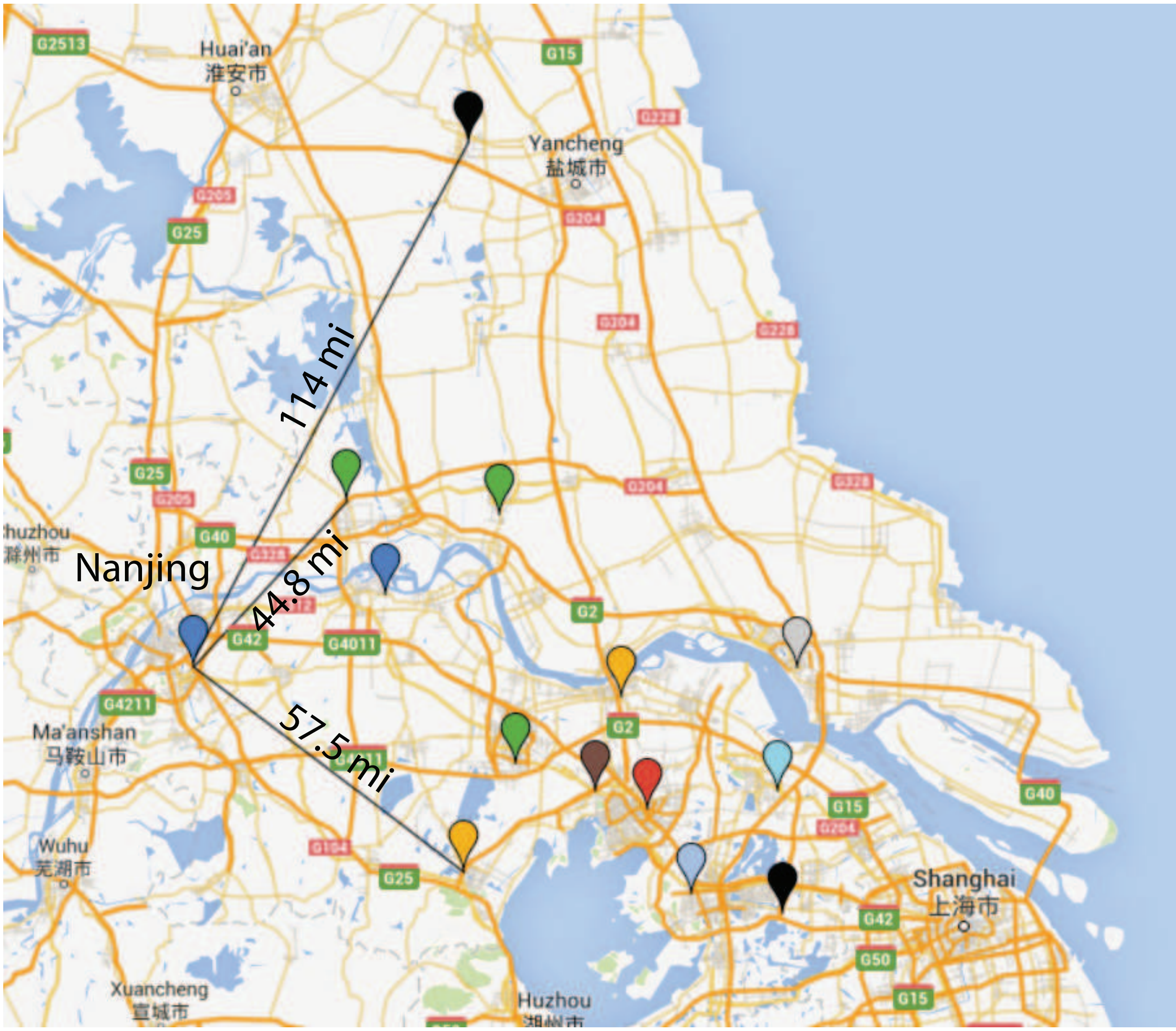

Figure 5: Volvo Dealerships in the Jiangsu Province

Each point represents a Volvo certified 4S dealer, and each color represents one owner. There were 14 Volvo dealers, owned by eight different companies. The black line measures the distance between the dealer of Nanjiang and the nearest three competing dealers.

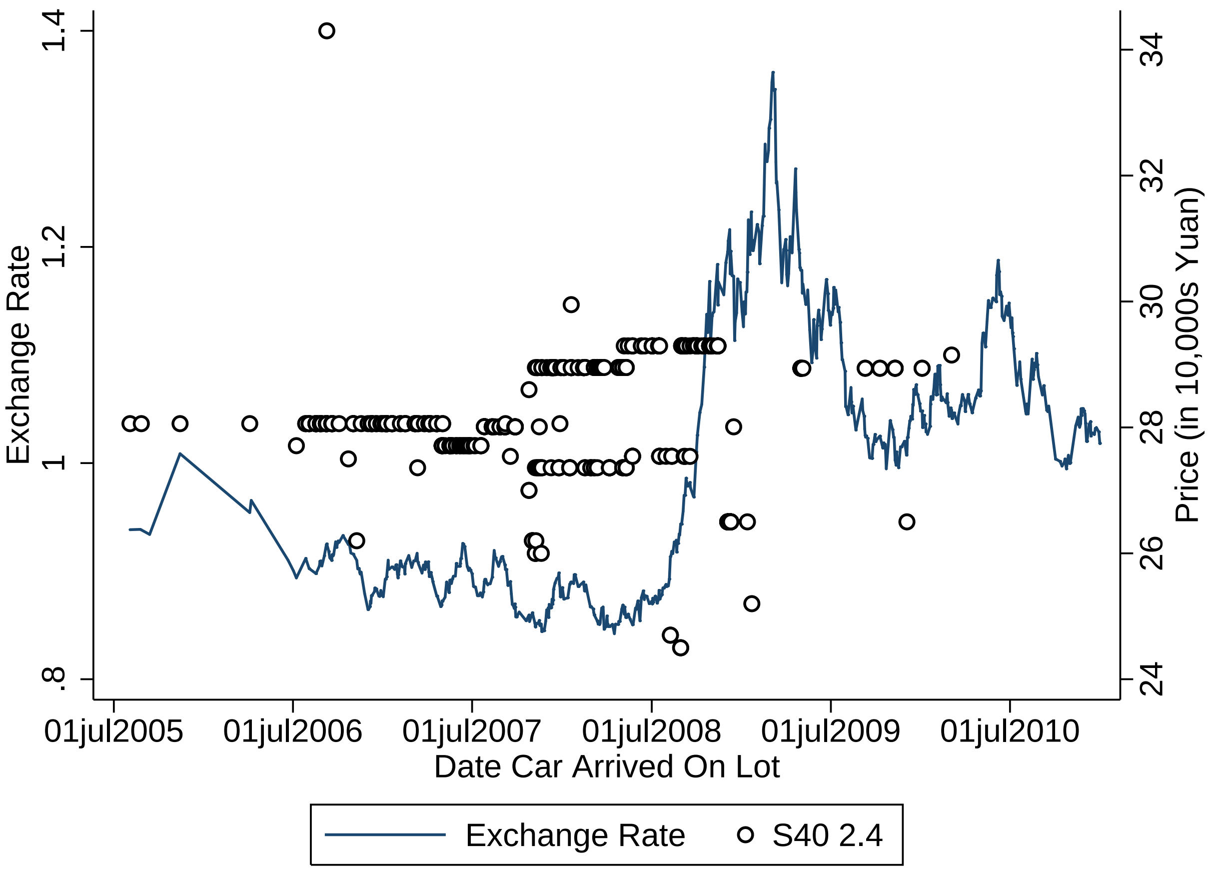

Figure 6: Exchange Rate and Arrival Prices vs. Date For S40 2.4 Trim

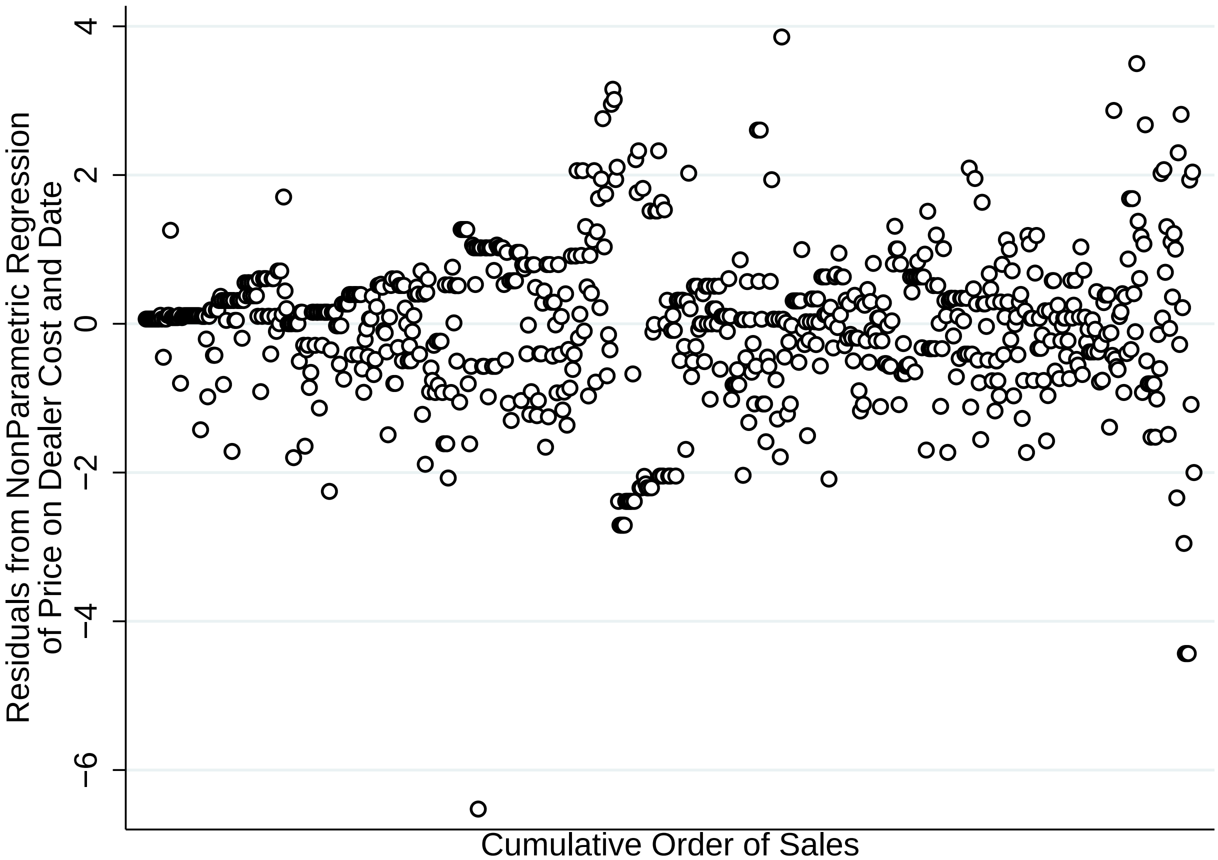

Figure 7: Evidence of Price Discrimination

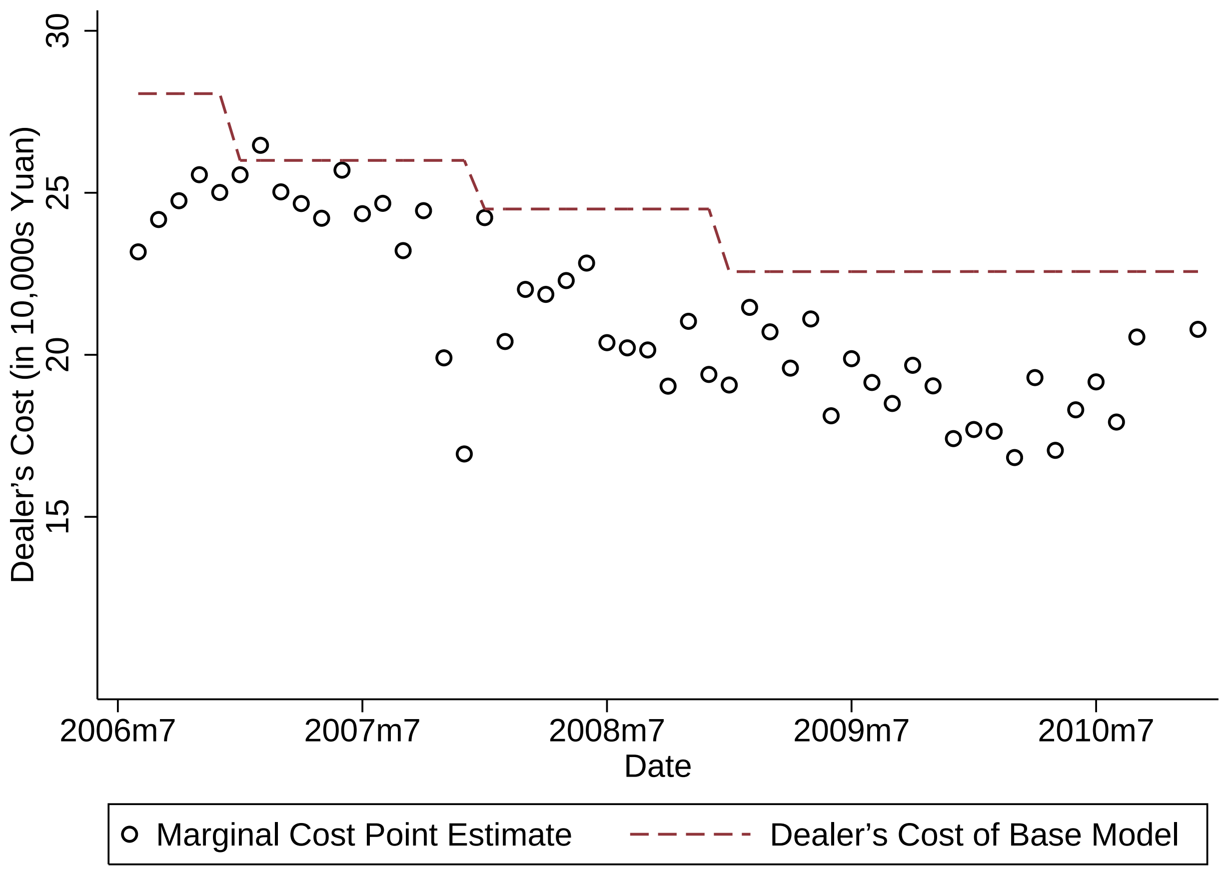

Figure 8: Monthly Marginal Cost Point Estimates

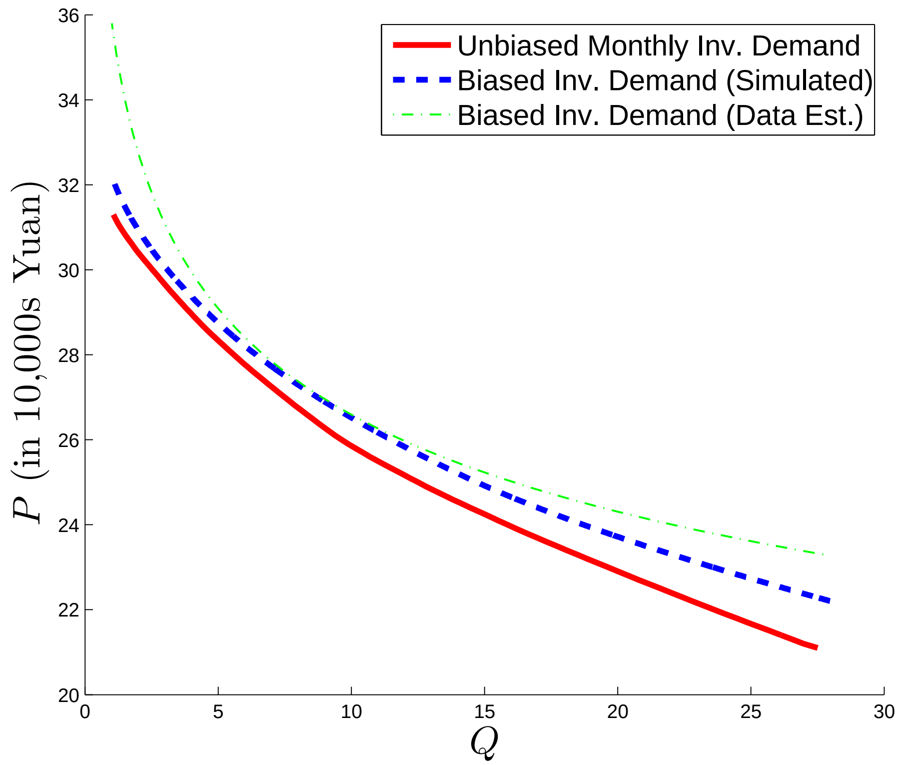

Figure 9: Inverse Demand Function Estimates - Linear Group Demand

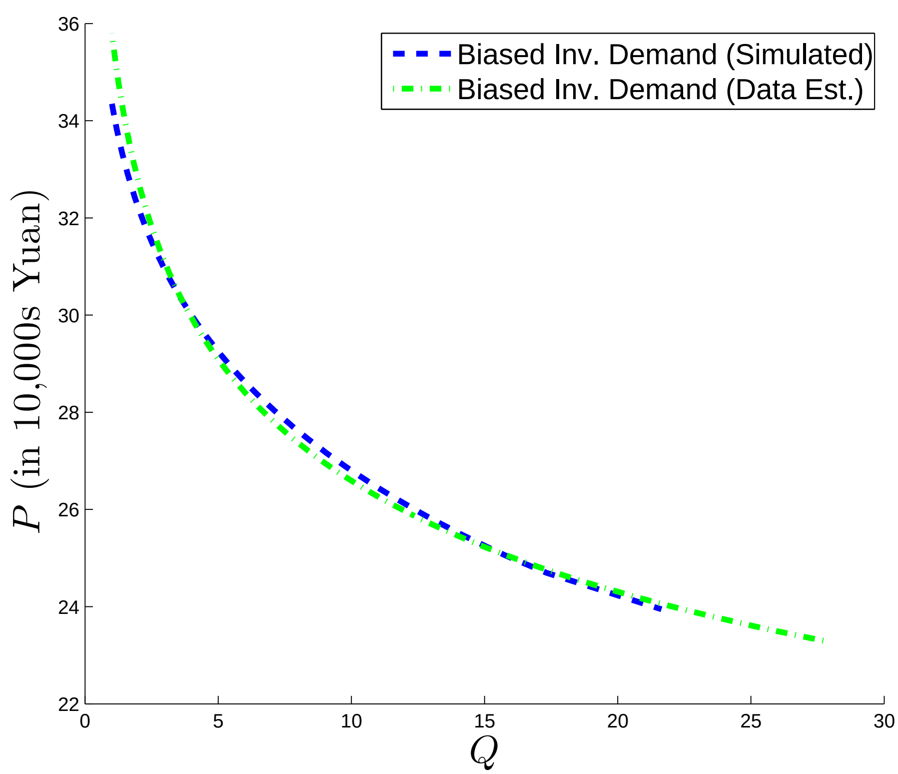

Figure 10: Inverse Demand Function Estimates - Curved Group Demand

Table 1: Volvo S40 Sale's Summary Statistics

Trim	Observations	Average Sale Price†	Average Cost to Dealer†	Average Pct. Markup
S40 2.4	417	28.8	28.2	2.31%
S40 T5	53	35.8	33.7	6.13%
S40 2.5T	12	30.6	32.6	−6.18%
S40 2.5	4	33.1	33.7	−2.01%
S40 2.0	251	25.1	25.4	−1.05%
Total	737	28.1	27.7	1.28%

†Unit of price: 10,000 Chinese Yuan (RMB)

Table 2: Exchange Rate's Impact on Markups

	Markup Over Cost (Yuan)				
	(i)	(ii)	(iii)	(iv)	(v)
Exchange rate when car sold	-7.601	-4.678	-4.706		-7.430
	(0.547)	(0.768)	(0.638)		(1.074)
Exchange rate when arrived on lot				-1.049	3.979
				(0.869)	(1.070)
Exchange rate when ordered by dealer					0.582
					(0.728)
Year FE		Y	Y	Y	Y
Month FE			Y	Y	Y
Constant	7.933	6.690	7.233	3.836	5.594
	(1.181)	(1.255)	(0.788)	(1.659)	(0.992)
N	694	694	694	694	671
R^2	0.311	0.526	0.604	0.567	0.632

Standard errors in parentheses

Clustered by date

Table 3: Biased IV Demand Estimation Results

	Dependent Variable is Logged Sales					
	OLS			IV		
	i	ii	iii	iv	v	vi
First Stage						
Lagged Dealer Inventory				-0.00333	-0.000999	-0.000841
				(0.00154)	(0.000759)	(0.000847)
Exchange Rate				-0.704	-0.183	-0.183
				(0.0823)	(0.0631)	(0.0458)
Second Stage						
Log (Avg. Price)	0.634	-0.976	-7.517	-0.0948	-7.248	-7.715
	(0.712)	(2.325)	(3.280)	(0.930)	(5.589)	(4.642)
Year Fixed Effects		Y	Y		Y	Y
Month Fixed Effects			Y			Y
Constant	0.470	5.541	28.40	2.799	26.47	29.00
	(2.280)	(7.809)	(11.21)	(2.976)	(18.75)	(15.85)
N	51	51	51	50	50	50
R^2	0.0159	0.135	0.395	.†	0.0122	0.400

Standard errors in parentheses

First stage rows show excluded instruments' coefficient values in first stage of 2SLS

Second stage rows show coefficient values in second stage of 2SLS

† Negative R^2 omitted

Table 4: Estimated Bias in Demand

	Price Percentile (Price in 10,000 Yuan)†				
Biased/Unbiased Est.	10th (24.51)	25th (26.4)	50th (29.55)	75th (32.7)	90th (34.59)
Biased Estimate of Q(P) ‡	19.51	11.60	4.95	2.19	1.38
	(27.25)	(4.71)	(2.17)	(1.16)	(0.27)
Unbiased Estimate of Q(P)	18.49	10.89	4.54	1.95	1.21
	(24.97)	(3.98)	(2.08)	(0.94)	(0.32)
Difference	1.02	0.71	0.41	0.24	0.17
	(2.29)	(0.74)	(0.13)	(0.24)	(0.32)
Percent Difference	5.51%	6.52%	9.03%	12.14%	13.84%
	(4.94)	(4.22)	(4.52)	(16.49)	(17.16)

† Price percentiles on unsimulated aggregate demand function, within range of observed monthly sales.

‡ Calculated by simulating the biased demand curve from the groups' unbiased demand estimates

Standard errors calculated via the delta method

Table 5: Estimated Bias in Slope of Demand

Biased/Unbiased Est.	10th (24.51)	25th (26.4)	50th (29.55)	75th (32.7)	90th (34.59)
	Price Percentile (Price in 10,000 Yuan)[†]				
Biased Estimate of $\frac{dQ}{dP}^{‡}$	-5.39	-3.18	-1.32	-0.55	-0.33
	(24.93)	(3.92)	(0.07)	(0.55)	(0.51)
Unbiased Estimate of $\frac{dQ}{dP}$	-5.18	-3.05	-1.25	-0.51	-0.30
	(23.68)	(3.43)	(0.07)	(0.61)	(0.54)
Difference	-0.21	-0.13	-0.07	-0.04	-0.03
	(1.25)	(0.50)	(0.02)	(0.06)	(0.03)
Percent Difference	3.97%	4.24%	5.63%	8.36%	10.64%
	(3.76)	(5.12)	(1.07)	(18.71)	(35.66)

[†] Price percentiles on unsimulated aggregate demand function, within range of observed monthly sales.
[‡] Calculated by simulating the biased demand curve from the groups' unbiased demand estimates
Standard errors calculated via the delta method

Table 6: **Appendix B** Biased IV Demand Estimation Regression - Regression-Based Quality-Adjusted Price

	OLS			IV		
	i	ii	iii	iv	v	vi
Log(Quality-Adjusted Avg. Price)	0.615	-0.886	-6.734	-0.0892	-6.346	-7.360
	(0.691)	(2.272)	(3.129)	(0.906)	(5.247)	(4.344)
Year Fixed Effects		Y	Y		Y	Y
Month Fixed Effects			Y			Y
Constant	0.542	5.229	25.65	2.780	23.37	27.69
	(2.201)	(7.602)	(10.65)	(2.880)	(17.54)	(14.78)
N	51	51	51	50	50	50
r2	0.0159	0.134	0.385	.	0.0416	0.391

Standard errors in parentheses