An Empirical Study of National vs. Local Pricing under Multimarket Competition

Yang Li*

Job Market Paper

August 30, 2011

Abstract

Geographic price discrimination is generally considered beneficial to firm profitability. Firms can extract higher rents by varying prices across markets to match consumers' preferences. This paper empirically demonstrates, however, that a firm may instead prefer a national pricing policy that fixes prices across geographic markets, foregoing the opportunity to customize prices. Under appropriate conditions, a national pricing policy helps avoid intense local competition due to targeted prices. I examine the choice of national versus local pricing under multimarket retail chain competition using extensive data from the digital camera market. I estimate a highly flexible model of aggregate demand that incorporates additional micro purchase moments and semi-parametric heterogeneity. Counterfactual analyses show that the major retail firms should employ a national pricing policy to maximize profits. Fixing prices across markets allows the retailers to soften otherwise intense local competition by subsidizing competitive markets with profit from less competitive markets. Additional results explore how market factors could affect the pricing policy decision and assist retail managers in choosing their geographic pricing policy.

*Graduate School of Business, Columbia University. Email: yli14@gsb.columbia.edu. I wish to thank my advisors Brett Gordon and Oded Netzer for their guidance and comments. I gratefully acknowledge Professor Vithala Rao for his generous support in obtaining the data. This project benefits from the grant awarded by the Eugene M. Lang Support Fund at Columbia Business School. All errors are my own.

1 Introduction

Geographic price discrimination is generally considered beneficial to firm profitability. Varying prices across markets with different socio-economic characteristics allows a firm to extract more consumer surplus by matching prices to consumers' local willingness to pay. Prior empirical work on geographic price discrimination documents such profit-enhancing effects (Chintagunta, Dubé and Singh, 2003). Many large retail chains, such as Walmart, Starbucks, and McDonald's, implement a form of region-based pricing that permits them to target prices to local market conditions.[1] In this study I argue, and empirically demonstrate, that in competitive settings, retailers may be better off forsaking the flexibility of local pricing in favor of a national pricing policy that fixes prices across geographic markets.[2]

The rationale behind such a national pricing policy is that targeted prices intensify local competition and increase the risk of a price war (Wells and Haglock, 2007). To illustrate the basic intuition in support of a national pricing policy, consider a simple example with two retail chains selling in three independent markets. The first two markets are monopolized by one of the two chains, respectively, and the third market is a duopoly in which both chains compete. Assuming similar price sensitivity across markets, under local pricing the chains set high prices in the monopoly markets and low prices in the duopoly market. If the duopoly market is relatively large the firm can increase its profits by committing to a single price across markets. The optimal national price falls between the otherwise high monopoly and low duopoly prices, thus softening the duopoly market competition (Dobson and Waterson, 2005). National pricing is optimal if the profit gain from softened competition in the duopoly market exceeds the profit loss sacrificed in the monopoly markets.[3] In effect the national pricing policy can be thought of as a mechanism that facilitates implicit price coordination.

The objective of this paper is to empirically examine a firm's choice of national versus

[1]Evidence can be found at, for example, http://walmartstores.com/317.aspx, and "Coffee talk: Starbucks chief on prices, McDonald's rivalry," *The Wall Street Journal*, March 7, 2011.

[2]In the remainder of the paper I interchangeably use the terms *national*, *uniform*, and *fixed* to refer to the policy of national pricing.

[3]Appendix A provides an analytical model in which I formalize this intuition.

local pricing in a multimarket competitive setting. I examine the multimarket pricing policy decisions in the context of the U.S. digital camera market, which generated $3 billion in sales in 2009. Point-of-sales data from the NPD Group provide a near census of the U.S. retail sales of digital cameras, including multiple large chains and rich geographic variation in market conditions. Two of the three largest chains in the data employed primarily national pricing policies.[4] Thus, this data set provides an excellent setting to study national versus local pricing, and the insights from this investigation could generalize to other industries evaluating their chain-level pricing policies. I focus on how a chain's choice of pricing policy results from balancing profits with competitive pressures across markets. Firms may have additional reasons to pursue a national pricing policy, such as a desire to avoid the organizational costs associated with local pricing or to maintain consistent prices offline and online.

In order to flexibly recover local consumer preferences, I estimate an aggregate model of demand with random coefficients separately in each of the over 1,000 markets in my data. Estimating the demand model separately across markets results in significantly more variation in elasticity estimates, particularly across markets with different market structures. To improve the estimation, I modify the model in Berry, Levinsohn, and Pakes (1995) in two ways: (1) I include micro-moments based on survey data that relate purchase behavior and consumer income levels (Petrin, 2002) and (2) I account for product congestion, which can confound estimation with unbalanced choice sets (Ackerberg and Rysman, 2005). Following Dubé, Fox and Su (2011), I formulate the demand estimation as a Mathematical Program with Equilibrium Constraints (MPEC), modifying it to include the additional micro-moments. Including the micro moments and correcting for product congestion improves the estimated substitution patterns and attenuates the price elasticities.

Given the demand estimates, I use the supply-side model to recover marginal costs. Estimating demand separately for each market is important for the supply side because pooled estimation across markets leads to an overestimation of the mean price-cost margin

[4] Due to a confidentiality requirement imposed by the data provider, I am prohibited from disclosing the names of retailers and camera brands in the data. Throughout the paper I denote chains and brands by generic letters and numbers.

by 31% and the median by 44%. Thus, addressing such biases on the demand side are necessary because they propagate into the supply-side estimation.

Although consumer preferences are estimated without any equilibrium assumptions, in order to recover cost estimates, I assume firms compete in a Bertrand-Nash equilibrium when setting prices. However, this equilibrium assumption only applies to the price-setting game and not to a chain's choice of its overall pricing policy (e.g., national versus local). This approach permits me to conduct several counterfactuals to assess the profitability of national and local pricing policies. First, a simulation demonstrates that the two major electronics retail chains in the data should employ national pricing policies to maximize profits. Uniform prices across markets allow the retailers to subsidize more competitive markets with profit from less competitive markets to soften the otherwise intense local competition. Compared to a situation in which both chains use local pricing policies, national pricing results in profit increases of 5.3% and 8.4%, respectively. Chen, Narasimhan, and Zhang (2001) discuss a similar finding in the context of targeting individual consumers. My results also relates to work on the coordination of retailer pricing strategies across channels (Zettelmeyer 2000) and choice of pricing formats across markets (Lal and Rao 1997; Ellickson and Misra 2008). Second, following the exit of one of the major retail chains, the remaining chain still prefers a national pricing policy due to competition from the remaining firms. Third, I investigate the boundary conditions under which a firm would prefer to stay with a national pricing policy. I find that the leading retailer would prefer local pricing if it were to close at least 29% of its stores in the competitive markets.

This paper broadly relates to the literature on retail pricing (Rao 1984; Eliashberg and Chatterjee 1985; Besanko, Gupta, and Jain 1998; Shankar and Bolton 2004), and in particular, on geographic price discrimination (Sheppard 1991; Hoch et al. 1995). Previous studies, however, generally neglect the multimarket structure of retail price competition. The closest existing paper to the present study is Chintagunta, Dubé, and Singh (2003), who study a single chain's zone-pricing policy across different neighborhoods in Chicago. The authors find that, by further localizing prices, a chain could substantially increase its profit without adversely affecting consumer welfare. Data limitations prevent the authors from incorporating information on competitors other than a distance-based proxy. Therefore, the counterfactual results do not account for competitive responses, whereas I explicitly model the interaction between retailers following a policy change. My findings provide empirical support to the theoretical literature on multimarket contact, such as Bernheim and Whinston (1990) and Dobson and Waterson (2005).

The rest of the paper is organized as follows. Section 2 introduces the data and overviews the market structure and pricing policies observed in the data. Section 3 describes the demand model and the chain pricing model. Section 4 details model estimation. Section 5 reports results of model estimation and counterfactual experiments. Section 6 concludes the paper with a discussion of its limitations and highlights areas of future research. All other details of the analysis are located in the Appendix.

2 Data and Industry Facts

In this section I discuss the data sets and the industry, and document the current market structure and pricing policies.

2.1 Data

The data in this paper come from a variety of sources: (1) store-level sales and price data on digital cameras from the NPD Group, (2) consumer survey statistics from PMA, (3) store location data from AggData, (4) digital camera sales across channels from Euromonitor, and (5) consumer demographics from the U.S. Census. Next I describe each of these data sets.

First, the NPD data is the main data set used in this study. It includes approximately 10 million monthly point-of-sales observations between January 2007 and April 2010. The data cover most stores in the U.S. that sell digital cameras. Each observation is at the month-store-camera model level, providing a highly granular picture of product-level sales across a large number of stores and time periods. Table 1 presents descriptive statistics for the store

Table 1: Descriptive Statistics of the Store Sales Data

Quarter	Total Revenue (billion \$)	Total Sales (million units)	Sales Weighted Average Price (\$)	# Camera Models
2007 Q1	0.617	2.834	215.79	940
2007 Q2	0.787	3.580	209.76	1016
2007 Q3	0.718	3.253	210.98	1060
2007 Q4	1.446	8.289	177.25	1092
2008 Q1	0.616	3.113	185.18	1143
2008 Q2	0.853	4.047	189.35	1149
2008 Q3	0.669	3.268	187.33	1167
2008 Q4	1.164	7.346	154.13	1172
2009 Q1	0.521	2.850	161.78	1203
2009 Q2	0.647	3.351	179.32	1284
2009 Q3	0.533	2.677	186.14	1162
2009 Q4	1.055	6.711	155.76	1132
2010 Q1	0.475	2.408	163.25	1118

sales data. Overall, after a long period of increase in sales, demand of digital cameras has generally declined since 2007. The industry is strongly seasonal, with sales in the fourth quarter nearly double other quarterly sales. There were nearly 60 unique camera brands in the data set. I focus my analysis on the largest seven brands that account for approximately 80% of sales, as reported in Table 2. The NPD data also contains a detailed description of the product characteristics for each camera model, such as mega-pixels, optical zoom, thickness, weight, display size, face detection and so forth.

Second, I use consumer survey data from the market research firm PMA to augment the estimation with micro moments, which relate consumer digital camera purchases with income levels. PMA conducts an annual survey each January on consumer purchasing and usage of digital cameras. The respondents consist of a rotating representative panel of 10,000 randomly selected U.S. households. From the survey responses I obtain the proportion of households at different income levels who bought a new digital camera. These proportions are used to construct the micro moments for demand estimation.

Third, I use the store location data from AggData to define and validate the competitive

Table 2: Annual Market Shares (%) of Camera Brands

	2007	2008	2009
Brand 1	21.4	21.5	21.6
Brand 2	17.0	19.1	20.6
Brand 3	7.3	11.1	12.8
Brand 4	16.4	13.7	12.2
Brand 5	6.4	6.1	5.6
Brand 6	5.5	5.4	5.2
Brand 7	3.8	4.5	5.2
Total	77.8	81.4	83.2

selling areas. NPD splits the U.S. into 2,100 distinct geographic markets called store selling areas (SSAs), which define competitive markets. 95% of SSAs contain only one store of each major retailer. The median distance between competing stores within an SSA is 0.58 miles while the median and the bottom 5th percentile distance to competing stores in neighboring SSA are 10.20 and 3.45 miles, respectively. Thus, the SSA definition indeed captures distinct geographic markets, with retail stores located nearby within a market and relatively farther from competing stores outside their SSA. Moreover, the correlation between the total number of households and the number of stores within an SSA is 0.66 ($p < 0.001$), and the correlation between the number of households and camera variety (i.e., distinct camera models) is 0.63 ($p < 0.001$). These strong positive correlations indicate that competition in this industry is highly localized; therefore, geographic difference must be carefully controlled for when modeling cameras sales at retailing stores.

Fourth, I use the channel sales data from Euromonitor to construct an appropriate market size definition. A proper measure of market size is important to accurately recover firms' mark-ups.[5] Common measures are population, number of households (e.g., BLP 1995), or total category demand (e.g., Song 2007). The use of population size as a proxy for demand is inconsistent with the observed seasonality in category sales. To correctly specify market size, I attempt to quantify all potential consumers including (1) those who bought cameras

[5]For example, in a homogeneous logit model, the mark-up across all of a firm's products is a constant and is negatively related to market size.

in the stores under investigation, (2) those who bought cameras through other channels (e.g., online), and (3) those who considered buying but chose not to. The first group of consumers directly corresponds to the store data assuming single-unit purchases per trip. For the second group, I estimate the share of consumers who purchased cameras outside of the retail chains using data on camera sales by distribution channel from Euromonitor International (2010). The third group represents consumers who are in the market but choose not to purchase a camera at all. To estimate this group, I obtained annual survey data on camera purchase intentions from PMA. The survey asked households about their purchase intentions in the next three-, six- or twelve-month periods. These percentages less the actual purchase probabilities from the PMA report of the following year yields a rough measure of the share of non-purchasers. In the demand model, I combine the second and third groups as the composite outside good.

2.2 Market Structure and Major Retailers

As illustrated by the analytical model in Appendix A, the relative advantage between national and local pricing relies on the characteristics of the market structure, in particular, the size of competitive markets versus monopoly markets, and competition. Next, I describe the patterns observed in the data regarding these characteristics.

The retail digital camera market is moderately concentrated with three national chains, A, B, and D, accounting for 70% of U.S. sales. Other retailers had shares below 3%. Accordingly, I focus on competition between these three big-box retail chains. Chains A and B are speciality retailers of consumer electronics, while Chain D is a discount retail chain. Figure 1 depicts the market shares of the three chains, which shows that before 2009, A and B accounted for approximately 40% and 16% shares of the market, respectively. At the end of 2008, chain B terminated operations and liquidated all stores within three months (for reasons independent of camera sales). The market share left by B was immediately taken up by A, making A the dominant player with almost 60% of the entire U.S. digital camera market. Chain D maintained approximately 9% share throughout the period.

Figure 1: Market Shares of Major Retailers

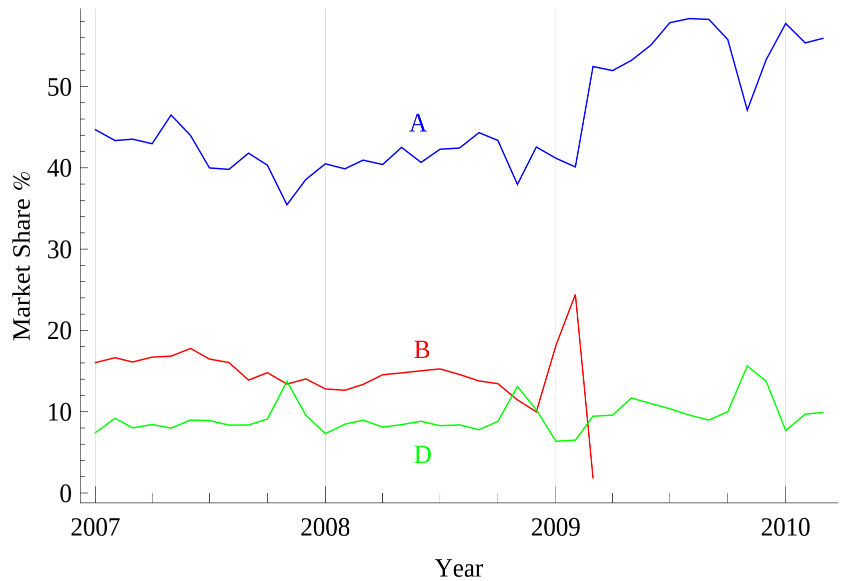


Table 3 presents the distribution of market structures across SSAs before and after chain B exited and the associated average annual sales. All three chains operate in a mixture of monopoly markets and oligopoly markets. The leading chain, A, had approximately 800 stores in 2007 and expanded to around 1,000 stores by early 2010. The second largest chain, B, operated approximately 600 stores until its bankruptcy. Before Chain B's exit, Chains A and B competed in more than half the markets in which they operated. At the same time, in many markets Chains A and B were either monopolists or only faced competition from Chain D.

Given these market conditions, it is unclear whether a firm would prefer a national or local pricing policy. On one hand, the retailer could leverage its power in the monopoly or low competition markets by employing a local pricing policy. On the other hand, the relatively large proportion of duopoly and triopoly markets may push the retailer to use national pricing policy to ease the competition. Both the distribution of market sizes and structures determine the optimal chain-level policy. After Chain B's exit, the number of monopoly markets for Chain A increased by approximately 65%. Again, whether Chain

Table 3: Market Type, Number of Markets, and Average Annual Sales (million units)

	Before B left		After B left	
Market type	# SSAs	Sales	# SSAs	Sales
A only	101	0.62	165	1.27
A & D	315	2.51	839	7.60
B only	79	0.33	—	—
B & D	118	0.71	—	—
A & B	59	0.76	—	—
A, B & D	402	5.60	—	—
D only	525	0.85	600	1.10

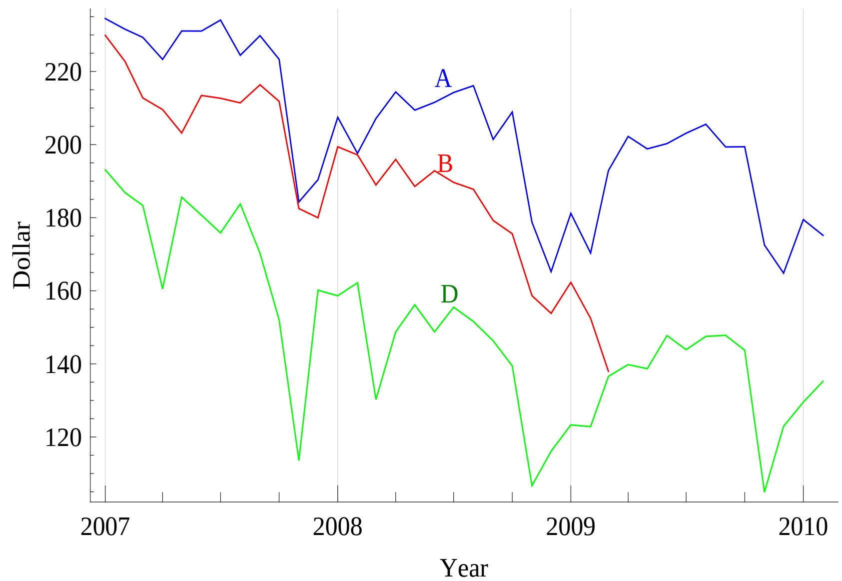
A would find it optimal to switch to local pricing following Chain B's exit depends on the relative size of these markets and the intensity of competition in its other markets. Although Chain A gained monopoly markets, the chain still faces competition from Chain D in many markets. Thus, a firm's choice of pricing policy is an empirical question, which I investigate in the next section using a structural empirical model of chain competition.

Besides differences in market structure, the three retailers also differentiate themselves according to price and product mix. Figure 2 plots the sales-weighted average price of each chain between 2007 and 2010. Chain A was the (relatively) "premium" retailer, charging a higher average price than its rivals. As expected, the discount chain, D, was the least expensive store, due to both lower prices and lower-end cameras sold by that retailer. Chain A tended to differentiate itself from Chain B when they coexisted in a local market. Chain A shelved 3.52 ($p < 0.01$) more camera models on average than a competing Chain B in the same market. For the same product mix, Chain A's store charged \$8.42 ($p < 0.01$) more per camera on average than Chain B's store did.

2.3 Pricing Policies

Both retailers A and B used nearly national pricing policies: prices for each product are nearly identical across geographic locations until the product reaches approximately 80% of its cumulative lifetime sales. For the remaining lifetime of the product sales, the products

Figure 2: Sales Weighted Average Price by Chain



often go on clearance and each local store can decide on price promotions. In contrast, Chain D implemented localized pricing throughout the life of the product.

To demonstrate these patterns, I select two popular camera models sold at each of the three chains. Figure 3 plots each product's sales-weighted price and its standard deviation across stores, and the cumulative share of lifetime sales against product age (in months).[6] The vertical lines at each price point represent the variation in the price across stores in a particular month. For Chains A and B, the cross-store price variation is minimal until cumulative sales exceeds approximately 80%. In contrast, the prices in Chain D exhibit substantial dispersion across stores from the time these products are introduced.

Figure 4 presents the coefficients of variation in the sales-weighted price across stores for all products relative to their cumulative share of lifetime sales. Each dot in the graph represents a camera model. For Chains A and B, before the cumulative share reaches approximately 80%, a product's price exhibits little to no variation across stores. In contrast, for Chain D, the price variation across stores is much higher and relatively constant over

[6]To determine the cumulative sales of products that entered prior to January 2007, I use national sales data from NPD aggregated over stores from January 2000 to March 2010.

Figure 3: Cross-store Price Dispersion as a Function of Product's Age and Cumulative Sales

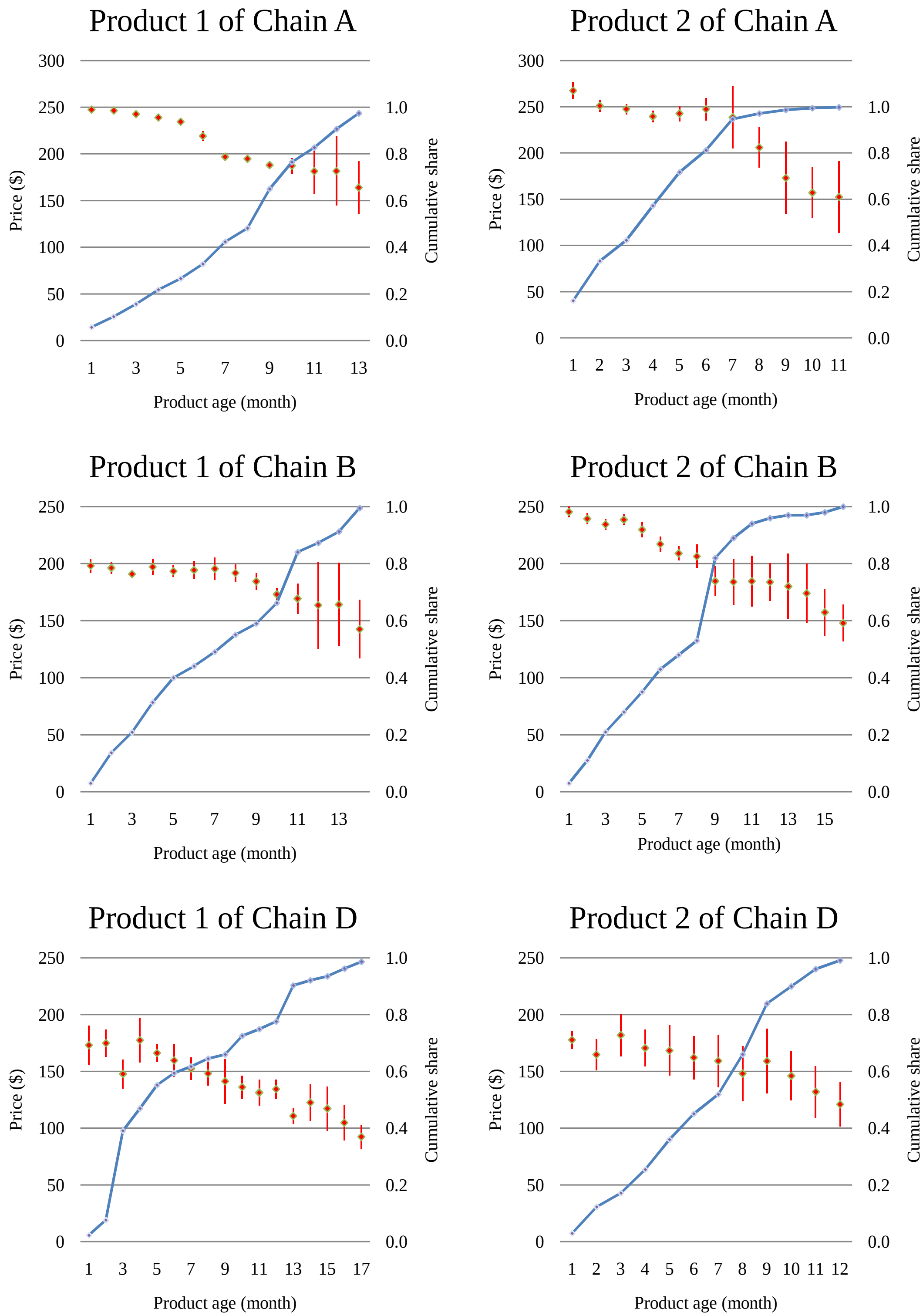
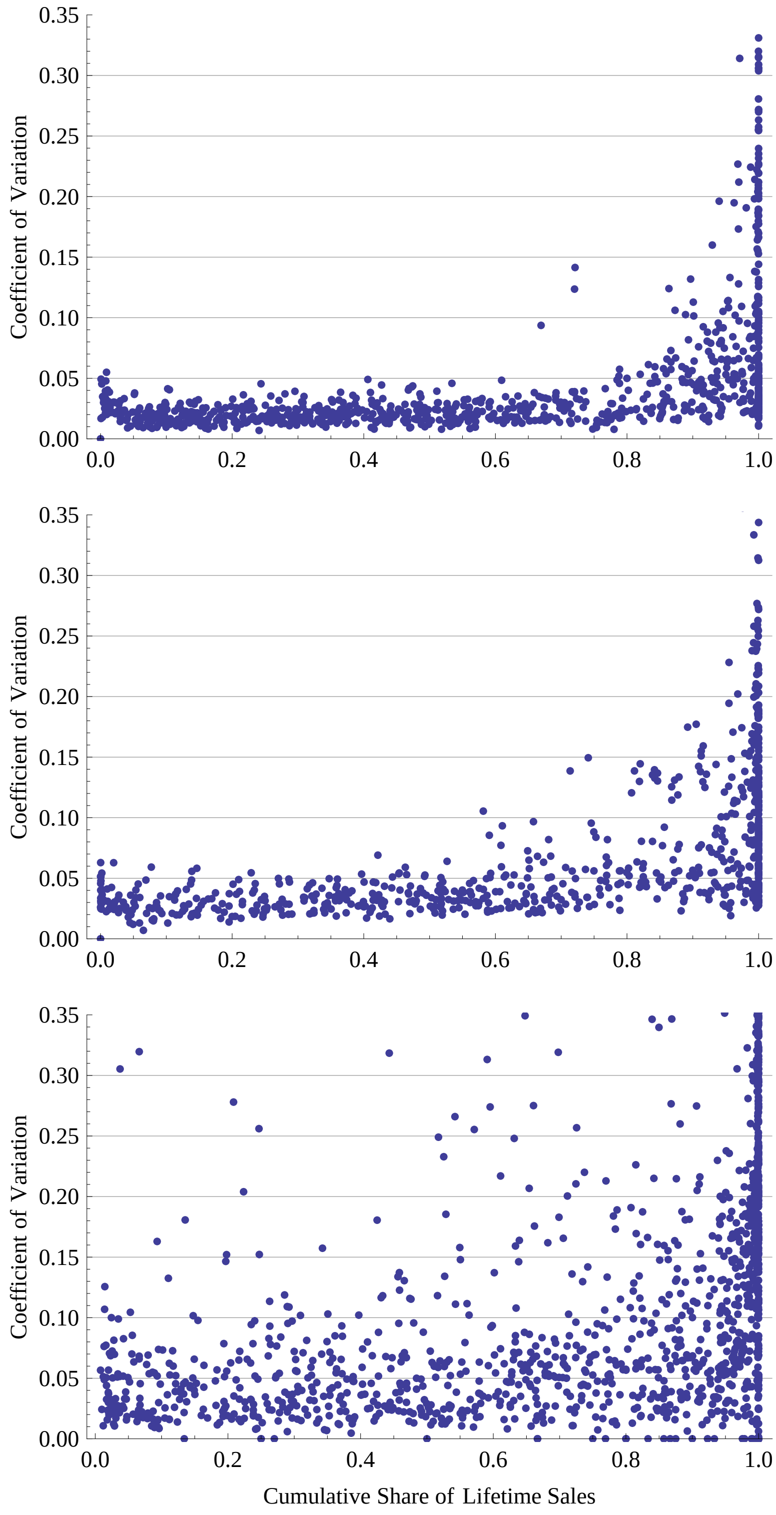


Figure 4: Price Dispersion in Chain A (top), B (middle) and D (bottom)



a product's lifecycle. The little observed dispersion for Chains A and B can be attributed to two sources under a national pricing policy. First, some sales are made using store-level coupons, open-box sales, or other local promotions that are independent of a chain's national pricing policy. Second, I must derive unit prices from the monthly sales data which contain product-level revenue and volume. Measurement error across stores in either the revenue or volume would generate apparent price variation. Both of these errors will be absorbed into an unobservable demand shock in the model.

In addition to these descriptive patterns in the data, discussions with a senior pricing director at one of the chains confirmed that Chains A and B both follow national pricing policies for most of a product's lifecycle, and then transition to local (clearance) pricing when they predict the product has reached a considerable portion (e.g., 80%) of its cumulative lifetime sales.

3 Model

This section provides a market-specific aggregate demand model in order to estimate consumer preferences. I then compute marginal costs for the counterfactual simulations using a supply side model. To facilitate demand estimation, I incorporate two important features: (1) a set of micro moments that relate income to digital cameras purchasing patterns, and (2) a "congestion" term that address the variation in products over time and across markets.

3.1 Aggregate Demand

I model consumer demand for digital cameras using an aggregate discrete choice model (Berry 1994; Berry et al. 1995). To incorporate demographic variation in income, I model consumer utility through a Cobb-Douglas function. The utility household i extracts from choosing product j at t is

$$U_{ijt} = (y_i - p_{jt})^{\alpha} G(\boldsymbol{x}_{jt}, \xi_{jt}, \boldsymbol{\beta}_i) e^{\epsilon_{ijt}} \quad (1)$$

where t=1, ..., T is the index for month and j=1, ..., J_t denotes the set of products at t. $\boldsymbol{x}_{jt}$ are observed product characteristics with coefficients $\boldsymbol{\beta}_i$.[7] ξ_{jt} represent unobservable shocks common to all households. These shocks may include missing product attributes, unquantifiable factors such as camera design and style, and measurement errors due to aggregation or sampling. y_i is the income of household i, p_{jt} is the price of product j at month t, and α is the price coefficient indicating the marginal utility of expenditures. For the income distribution y_i, I use zip-code level demographics from the U.S. Census adjusted by the CPI inflation data from the U.S. Bureau of Labor Statistics to match the periods under investigation.[8]

$G(\cdot)$ is assumed to be linear in logs, and the transformed utility for j=1, ..., J_t is

$$u_{ijt} = \boldsymbol{x}'_{jt}\boldsymbol{\beta}_i + \alpha \log(y_i - p_{jt}) + \xi_{jt} + \epsilon_{ijt} \tag{2}$$

Accordingly, the utility for outside option j=0 is

$$u_{i0t} = \alpha \log(y_i) + \epsilon_{i0t} \tag{3}$$

Assuming that ϵ's are distributed type-I extreme value, the market share of product j at month t is simply the logit choice probabilities aggregated over all households in the market

$$s_{jt} = \int_{\forall i} s_{ijt} = \int_{\forall i} \frac{\exp[\boldsymbol{x}'_{jt}\boldsymbol{\beta}_i + \alpha \log(1 - p_{jt}/y_i) + \xi_{jt}]}{1 + \sum_{k=1}^{J_t} \exp[\boldsymbol{x}'_{kt}\boldsymbol{\beta}_i + \alpha \log(1 - p_{kt}/y_i) + \xi_{kt}]} d\mathrm{P}(\boldsymbol{\beta}_i) d\mathrm{P}(y_i) \tag{4}$$

where $\mathrm{P}(\boldsymbol{\beta}_i)$ and $\mathrm{P}(y_i)$ are probability density functions of heterogeneous tastes and household income, respectively. Following the literature, I assume $\boldsymbol{\beta}_i$ is normally distributed and y_i is log-normally distributed. The normality assumption on consumer heterogeneity may cause estimation bias if the actual distribution is heavily tailed or multi-mode, as demon-

[7] Bold fonts denote vectors or matrices. All vectors are by default column vectors.

[8] One issue in using a Cobb-Douglas utility is that income must be larger than the price after taking logs. With simulated income draws, some random draws could violate this condition. In this paper, prices of digital cameras are quite low relative to income, so the estimation bias caused by such sample selection on income is negligible.

Table 4: Percent of Households Which Purchased A New Camera

Year	< \$29,999	\$30,000–\$49,999	\$50,000–\$74,999	> \$75,000
2007	8%	16%	20%	20%
2008	8%	12%	14%	18%
2009	7%	11%	14%	15%

strated by Li and Ansari (2011). To allow for flexible heterogeneity distribution I estimate the demand model separately for each market, leading to a semi-parametric estimation of national-level consumer heterogeneity.

Similar to prior work (e.g., Zhao, 2006; Lou et al., 2008), the set of observed camera attributes that I use includes price and five key attributes: camera brand, mega-pixels, optical zoom, thickness, and display size. Given that the sale observations in each market may not be sufficient to estimate a full set of random coefficients, I further decompose $\boldsymbol{x}_{jt}$ into $\boldsymbol{x}_{jt}^{fc}$ and $\boldsymbol{x}_{jt}^{rc}$, and assign random coefficients only on $\boldsymbol{x}_{jt}^{rc}$. $\boldsymbol{x}_{jt}^{rc}$ includes mega-pixels, store and camera brand. The other three non-price attributes are included in $\boldsymbol{x}_{jt}^{fc}$. Heterogeneity in price sensitivity is estimated through income distribution. As seasonality is strong in this industry, I add a "November-December" dummy and a "June" dummy in $\boldsymbol{x}_{jt}^{fc}$ to capture possible seasonal effects such as a temporary expansion of market size. In this demand model a product j is defined as a particular camera sold in a particular store. The lack of information on store characteristics makes it impossible to construct nested choice models. It is also not clear that consumers actually follow the nested process and choose stores before selecting products. Thus, I treat store affiliation as an additional product attribute that additively enters into the consumer utility function.

3.2 Micro Moments

Estimates from aggregate models can be improved by leveraging information that links average consumer demographics to their purchase behavior (Petrin, 2002). I divide each market into R distinct income tiers, with varying price coefficients across these tiers

$$\alpha_r = \begin{cases} \alpha_1, & \text{if } \; y_i < \bar{y}_1 \\ \alpha_2, & \text{if } \; \bar{y}_1 \leq y_i < \bar{y}_2 \\ \vdots \\ \alpha_R, & \text{if } \; y_i > \bar{y}_{R-1} \end{cases} \tag{5}$$

where $\bar{y}_1, \bar{y}_2, ..., \bar{y}_{R-1}$ are the cutoffs on income. PMA defined four income tiers from its consumer surveys and reports average purchase probabilities of households at different income tiers (Table 4). In demand estimation, I construct additional micro moments according to

$$\mathrm{E}[\{\text{household } i \text{ bought a new camera at } t\}| \; \{i \text{ belongs to income tier } r \text{ at } t\}],$$

where r=1, ..., R, and match these moments to the variation of purchase probabilities across income groups in the PMA data. The function of micro moments is different from hierarchically adding demographics via parameter heterogeneity. The latter approach only provides extra flexibility in the model, whereas the micro moments entail a process that restricts the GMM estimator to match additional statistics, making the estimated substitution pattern directly reflect demographic-driven differences in choice probability. Also, the variation in purchase probabilities across income groups provides new information that facilitates parameter identification.

To apply the PMA data, two modifications are needed before constructing the micro moments. First, the PMA data provide digital camera purchase likelihood by income tier at the national level, whereas my analysis is at the local market level. Thus, I scale the PMA data to make them consistent with the geographic differences in demographics and with the actual market size underlying the demand model. I discuss the details of the scaling procedure in Appendix C. Second, given the store data are at the monthly level, I linearly interpolate the PMA yearly data to convert them to monthly observations.

3.3 Product Congestion

Logit choice models impose strong restrictions on how the space of unobserved characteristics (the ϵ's) changes with the the number of products. These restrictions can bias elasticity estimates if there is substantial variation in the number of products across markets or time. The camera market in particular undergoes frequent product entry and exit due to the seasonal pattern of sales. The average number of products across markets varies from 25 to 64, and the within-market variation is about 23%.

In classical demand models (e.g., the Hotelling model), product "congestion" occurs because the space of product characteristics is bounded and a new product makes the characteristic space more crowded. However, in logit models, product congestion occurs in the observed characteristics space but not in the unobservable characteristics space. With each new product, a new i.i.d. ϵ is added. Price sensitivity can be estimated without price variation and solely based on variation in the number of products across markets, leading to biased elasticity estimates.

To accommodate congestion in logit models, Ackerberg and Rysman (2005) propose a modified specification in which a bound is imposed on the space of unobserved characteristics, thereby allowing for congestion in this space. The bound is a function of the number of products in a market, and the products are considered being equally differentiated along unobserved characteristics but constrained by the bound. The bound is implemented as a congestion term $\log(R_{jt})$, where $R_{jt} = J_t^{\gamma}$, and γ is to be estimated.[9] I add such congestion term to the model and test a more flexible specification in the robustness check section.

3.4 Chain-Level Pricing Model

This subsection presents the supply-side model of chains engaged in multimarket price competition. A chain operates under either a national pricing policy that fixes the same price for a product across markets, or a local pricing policy that customizes prices across

[9] An alternative specification is $R_{jt} = \gamma/J_t + 1 - \gamma$.

markets. In each month, the chain sets prices according to the overall policy.

The demand estimation was free of equilibrium assumptions in order to recover consumer preferences. To conduct counterfactual simulations, I must obtain estimates of each chain's marginal costs. These costs are assumed constant for a given product across markets and independent of the chain-level pricing policy. Constant marginal costs seem reasonable given the efficient distribution of consumer electronics and the chain-controlled sales force compensation schemes. I use the demand parameter estimates and observed prices to recover the marginal costs under the assumption that the chains compete in a Bertrand-Nash equilibrium. It should be noted that the equilibrium assumption only applies to the period price setting game and not to each chain's overall choice of pricing policy (national vs. local). This approach permits me to test a chain's pricing policy choice in a set of counterfactual analyses. Further, estimating the supply side under either national or local pricing for the firms yields nearly identical estimates of marginal costs, suggesting the ability to recover costs is not sensitive to this assumption.

Each chain f sells some subset of J_{ft} of the total J_t products. With national pricing policy, a chain has a profit function that sums up local profits with uniform prices (t is suppressed in the rest of this section)

$$\Pi_f = \sum_{j=1}^{J_f} (p_j - c_j) \sum_{\forall m} s_{jm} M_m \tag{6}$$

where m denotes a market where the chain operates. M_m represents the size of market m and s_{jm} is the share of product j in m.

Given that s_{jm} is a function of price p_j, the first-order condition with respect to p_j is

$$\sum_{\forall m} s_{jm} M_m + \sum_{r=1}^{J_f} (p_r - c_r) \sum_{\forall m} \frac{\partial s_{rm}}{\partial p_j} M_m = 0, \text{ for } j = 1, ..., J_f\ . \tag{7}$$

Stacking prices and costs and aligning simulated shares across markets, the pricing equation

(7) can be written in matrix notation for all competing chains

$$\boldsymbol{c} = \boldsymbol{p} - \boldsymbol{\Delta}^{-1}\boldsymbol{q} \tag{8}$$

where $\boldsymbol{q} = \sum_{\forall m} M_m \int_{i \in m} \boldsymbol{s}_i$, is a vector of total unit sales of each product, and $\boldsymbol{\Delta}$ is a block diagonal matrix in which each block, $\boldsymbol{\Delta}_f$, corresponds to a chain. Let $\mu_i(\boldsymbol{p}) = \alpha_r \log(1-\boldsymbol{p}/y_i)$, and so $\partial\mu_i(\boldsymbol{p})/\partial\boldsymbol{p}$ is a diagonal matrix. Then,

$$\boldsymbol{\Delta}_f = -\sum_{\forall m} M_m \int_{i \in m} [\frac{\partial\mu_i(\boldsymbol{p})}{\partial\boldsymbol{p}}(\text{diag}(\boldsymbol{s}_i) - \boldsymbol{s}_i\boldsymbol{s}_i')] \tag{9}$$

Here the integrations are specific to the demographic distribution in market m.

Under local pricing, the profit in one market is independent from another market. The summation over m in (9) drops out, and market size cancels out as well. That is,

$$\boldsymbol{c} = \boldsymbol{p} - \boldsymbol{\Delta}^{-1}\boldsymbol{s} \tag{10}$$

where $\boldsymbol{s} = \int \boldsymbol{s}_i$ is a vector of product shares, and

$$\boldsymbol{\Delta}_f = -\int [\frac{\partial\mu_i(\boldsymbol{p})}{\partial\boldsymbol{p}}(\text{diag}(\boldsymbol{s}_i) - \boldsymbol{s}_i\boldsymbol{s}_i')] \tag{11}$$

Using (8) and (10), I compute the marginal costs using the demand estimates as input. Then in the counterfactual simulation, I use the same formulas to calculate the new equilibrium prices under alternative pricing policies. Based on pricing pattern observed in the data section, I assume A and B fixed price across markets for each camera before the sales of the camera hit the 80% threshold of its lifetime sales. In calculating marginal costs, I combine (8) and (10) to capture this transition. The marginal costs in the last 20% of sales are solve by constrained optimization on Equation (10), constraining the cost to be the same across markets for each product, in order to be consistent with cost uniformity assumption made above for these two chains.

4 Estimation

In this section I discuss the details of model estimation. The digital cameras market is characterized by rich geographic variation in market structure, product mix, and consumer demographics. Thus the method of estimating demand needs to take into account the local variation in market conditions. To this end, I estimate the demand model separately for each of the over 1,000 markets in which A, B and D operated. Because a market contains approximately 1,200 observations on average, separate estimation for each market permits me to include heterogeneity within a market, but does not constrain the shape of preference heterogeneity across markets. For comparison purposes, I also estimate a single model that pools the data across markets.

4.1 Moments

In each market, the demand system has the following two components

$$s_{jt} = \int_{\forall i} \frac{\exp(V_{ijt})}{1 + \sum_{k=1}^{J_t} \exp(V_{ikt})} d\mathrm{P}(\boldsymbol{\beta}_i) d\mathrm{P}(y_i) \tag{12}$$

$$\tilde{s}_{rt} = \int_{i \in r} \sum_{j=1}^{J_t} s_{ijt} \tag{13}$$

where (12) is market share equation with the systematic utility

$$V_{ijt} = \boldsymbol{x}_{jt}^{fc'} \boldsymbol{\beta}_{fc} + \boldsymbol{x}_{jt}^{rc'} \boldsymbol{\beta}_i + \alpha_r \log(1 - p_{jt}/y_i) + \rho \log(R_{jt}) + \xi_{jt}$$

and (13) is implemented as micro moments with $\tilde{s}_{rt}$ denoting the percent of households at income tier r who purchased new cameras at t. The integrals in these equations are numerically computed through Monte-Carlo simulation. For each dimension, I use $I = 2000$ pseudo-random draws generated from Sobol sequence to approximate the integrals (Train 2003).

Append four identical price terms, $\log(1 - p_{jt}/y_i)$ in $\boldsymbol{x}_{jt}^{rc'}$ to form $\boldsymbol{x}_{ijt}^{rc}$. Then stack observations $\forall j$ and then $\forall t$ as rows into matrices and rewrite the systematic utility V_{ijt} as

$$\boldsymbol{V}_i = \boldsymbol{X}\boldsymbol{\theta}_1 + \boldsymbol{X}_i^{rc}\boldsymbol{\theta}_2\boldsymbol{v}_i + \boldsymbol{\xi} \tag{14}$$

where $\boldsymbol{\theta}_1$ is a vector combining the fixed (non-random) coefficients $\boldsymbol{\beta}_{fc}$, the means of the random coefficients, $\bar{\boldsymbol{\beta}} = \mathrm{E}[\boldsymbol{\beta}_i]$, as well as the coefficient of the congestion term ρ. $\boldsymbol{\theta}_2$ is the Cholesky root of the covariance matrix of the random coefficients appended with α_r's as the last four diagonal elements. $\boldsymbol{v}_i$ is a vector consisting of random draws from a standard multivariate normal distribution associated with $\boldsymbol{\beta}_i$, as well as four binary indicators of i's income level. The mean utility invariant across households is therefore

$$\boldsymbol{\delta} = \boldsymbol{X}\boldsymbol{\theta}_1 + \boldsymbol{\xi} \tag{15}$$

The demand system is estimated by GMM estimator. There are three sets of moments in this estimation: the share equations (12), the micro moments (13), and the demand-side orthogonality conditions, which I describe next. Assuming $\boldsymbol{\xi}$ is mean independent of some set of exogenous instruments $\boldsymbol{Z}$, the demand side moments are given by

$$\boldsymbol{g}(\boldsymbol{\delta}, \boldsymbol{\theta}_1) = \frac{1}{N_d}\boldsymbol{Z}'\boldsymbol{\xi} = \frac{1}{N_d}\boldsymbol{Z}'\left(\boldsymbol{\delta} - \boldsymbol{X}\boldsymbol{\theta}_1\right) = \boldsymbol{0} \tag{16}$$

where N_d denotes the number of sale observations.

I construct two sets of instruments to identify demand parameters. The first set follows the approximation to optimal instruments in Berry et al. (1995), which include own product characteristics, the sum of the characteristics across other own-firm products, and the sum of the characteristics across competing firms. The second set of instruments are obtained through the intuition that a product's price is partially determined by its proximity to rival products in characteristics space. I calculate the Euclidean distances from own product characteristics to every competing product and then average the distances to get the second set of instruments. The two sets of instruments together explain relatively large portion of price variation. The R^2 in the regression of price on the instruments is 0.72 on average.

4.2 MPEC Approach

In demand estimation the GMM estimator minimizes the 2-norm of $\boldsymbol{g}(\boldsymbol{\delta}, \boldsymbol{\theta}_1)$ in (16), subject to the constraints imposed by the share equations (12) and by the micro moments (13). Berry (1994) proposes a contraction mapping procedure to numerically invert the share in (12) within each GMM minimization iteration. This nested unconstrained optimization approach, as Dubé et al. (2011) point out, is slow and sensitive to errors which propagate from the unconverged contract mapping.

Following the work of Su and Judd (2010) and Dubé et al. (2011), I formulate the aggregate demand estimation as a mathematical program with equilibrium constraints. Further, I incorporate the micro moments (13) into the MPEC framework. Specifically, I treat the micro moments as additional nonlinear constraints to the estimation objective function, and solve the nested problems and the GMM minimization simultaneously by augmenting the Lagrangian. The constrained optimization can be written as,

$$
\begin{aligned}
\min_{\boldsymbol{\phi}} \quad & \mathrm{F}(\boldsymbol{\phi}) = \boldsymbol{\eta}'\boldsymbol{W}\boldsymbol{\eta} \\
\text{s.t.} \quad & \boldsymbol{s}(\boldsymbol{\delta}, \boldsymbol{\theta}_2) = \boldsymbol{S} \\
& \boldsymbol{\eta}_1 - \boldsymbol{g}(\boldsymbol{\delta}, \boldsymbol{\theta}_1) = \boldsymbol{0} \\
& \boldsymbol{\eta}_2 - \tilde{\boldsymbol{s}}(\boldsymbol{\delta}, \boldsymbol{\theta}_2) = -\tilde{\boldsymbol{S}}
\end{aligned} \tag{17}
$$

where $\boldsymbol{\phi}$=$\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\delta}, \boldsymbol{\eta}_1, \boldsymbol{\eta}_2\}$ contains the optimization parameters. $\boldsymbol{W}$ is the weighting matrix. $\boldsymbol{S}$ is a vector of actual shares. $\tilde{\boldsymbol{S}}$ is a vector of the micro data collected from the PMA consumer survey. $\boldsymbol{\eta} = (\boldsymbol{\eta}_1' \ \boldsymbol{\eta}_2')'$ includes the auxiliary variables that yield extra sparsity to the Hessian of the Lagrangian (Dubé et al., 2011). I choose to enter the micro moments into the objective function because the weights on these constraints for minimization can be adaptively determined by the data via a two-stage estimation process. Moreover, the sparsity patterns in the original constrained optimization is unchanged after adding these micro moments, as these moments involve only shares that are independent across t. Therefore, the computing memory requirement of this MPEC is relatively mild.[10]

Denote the set of constraints as $\mathcal{G}(\boldsymbol{\phi})$, the constrained optimization problem (17) results in the following Lagrangian function

$$\mathcal{L}(\boldsymbol{\phi};\boldsymbol{\lambda}) = \mathrm{F}(\boldsymbol{\phi}) - \langle \boldsymbol{\lambda}, \mathcal{G}(\boldsymbol{\phi}) \rangle \tag{18}$$

where $\boldsymbol{\lambda} \in \mathbb{R}$ is a vector of Lagrange multipliers. Then the solution to (17) satisfies the following Karush-Kuhn-Tacker condition for $\mathcal{L}$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\phi}} = 0, \qquad \mathcal{G}(\boldsymbol{\phi}) = \mathbf{0} \tag{19}$$

The model estimation proceeds in two stages. In the first stage, identity matrix is used as the weighting matrix $\boldsymbol{W}$ in (17). In the second stage, equal weighting is replaced by the inverse matrix of the second moments $\boldsymbol{\Phi}$, which is a function of first-stage estimates. The micro moments (over i and r) are sampled independently from demand moments (over j and t), therefore $\boldsymbol{\Phi}$ has a block diagonal structure (Petrin, 2002). Accordingly, the asymptotic variance matrix for parameter estimates is given as

$$\boldsymbol{\Gamma} = \frac{1}{N_d + I} (\boldsymbol{J}'\boldsymbol{W}\boldsymbol{J})^{-1} \boldsymbol{J}'\boldsymbol{W}\boldsymbol{\Phi}\boldsymbol{W}\boldsymbol{J}(\boldsymbol{J}'\boldsymbol{W}\boldsymbol{J})^{-1} \tag{20}$$

where $\boldsymbol{J}$ is the Jacobian matrix of (16) and (13) with respect to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$.

In Appendix B, I derive the close-form Jacobian and Hessian formulas for the objective function, the demand moments, and the micro moments. The derivation follows the rules of matrix calculus and so the formulas are compactly written in matrix notation, which facilitates vectorization in actual coding.

[10] The transition from unconstrained optimization to constrained optimization usually increases the need on computing memory due to the added constraints. For dense problems, unconstrained optimization may be preferred over constrained optimization (Nocedal and Wright, 1999).

5 Results

In this section, I first present the estimates of the demand parameters and elasticities under alternative model specifications. Then, I report the results of the counterfactual experiments in which demand estimates are used to calculate the firms' profit under local and national pricing policies and under varying competitive market conditions.

5.1 Parameter Estimates

This subsection discusses the parameter estimates, elasticities, and margins across various model specifications. First, I present the parameter estimates from the pooled (across markets) demand model, which makes it easier to discuss the implications of each parameter. Second, I discuss the results from estimating the demand model separately across the 1,177 markets. Note that the parameter estimates are not directly comparable across markets because the scale utility is different (Swait and Louviere 1993). To facilitate comparison, I calculate the elasticities in both the separate and pooled estimation.

Table 5 reports the parameter estimates from the pooled demand model. The price coefficient triples when moving from OLS to 2SLS with instrumental variables, suggesting price endogeneity is present in the demand specification. The random coefficients model with micro data identifies that the price coefficients vary substantially across income tiers. Similar to the findings in Petrin (2002), I find that the marginal utility of expenditures on other goods and services increases with income. Consumers on average favor cameras with higher mega-pixels, longer optical zoom, and larger display, and they dislike cameras that are thick in size. Yet the taste for mega-pixels is highly heterogeneous across consumers. Some consumers in the market appear to have little valuation for resolution, consistent with the industry trend that the pursuit of higher resolution in the compact point-and-shoot sector has declined since 2007 (Euromonitor 2010).

Table 6 reports the elasticities from estimating the model separately across markets, and

Table 5: Parameter Estimates of The Pooled Demand Model

Variable	OLS	2SLS	Random Coefficients	Random Coefficients & Microdata
	Price Coefficients (α's)			
α_1	4.621	16.235	30.878	9.064
	(0.033)	(0.030)	(0.846)	(1.021)
α_2				26.465
				(6.742)
α_3				75.788
				(9.793)
α_4				85.334
				(13.007)
	Other Parameters			
Mega-pixels	0.051	0.057	0.547	1.350
	(0.001)	(0.000)	(0.025)	(0.014)
Mega-pixels s.d.			0.049	0.856
			(0.006)	(0.074)
Optical Zoom	0.004	0.014	0.015	0.065
	(0.001)	(0.000)	(0.007)	(0.001)
Thickness	-0.159	-0.183	-0.177	-0.371
	(0.003)	(0.001)	(0.002)	(0.005)
Display Size	0.340	0.457	0.564	1.101
	(0.002)	(0.000)	(0.002)	(0.041)
Nov-Dec	-0.137	-0.159	-0.361	-0.114
	(0.002)	(0.002)	(0.000)	(0.000)
June	0.100	0.117	0.056	0.035
	(0.003)	(0.003)	(0.000)	(0.000)
Congestion	-0.934	-0.931	-0.910	-0.852
	(0.002)	(0.000)	(0.024)	(0.017)

Note: Standard errors are in round brackets. All specifications include year fixed effects and brand-chain interactions.

Table 6: Elasticity Estimates

	Separate Estimation		Pooled Estimation	
Elasticity	2SLS	Random Coefficients & Microdata	2SLS	Random Coefficients & Microdata
Price	-1.507	-1.921	-1.345	-1.618
	[0.251]	[0.437]	[0.171]	[0.254]
Mega-pixels	0.390	0.424	0.460	0.576
	[0.179]	[0.307]	[0.099]	[0.181]
Optical Zoom	0.080	0.067	0.051	0.064
	[0.169]	[0.181]	[0.059]	[0.043]
Thickness	-0.206	-0.305	-0.228	-0.367
	[0.270]	[0.264]	[0.126]	[0.132]
Display Size	0.212	0.179	0.118	0.253
	[0.559]	[0.532]	[0.012]	[0.094]

Note: Standard deviations are computed across markets and put in square brackets.

compares them to elasticities from the pooled estimation.[11] Figure 5 plots the density of the price elasticity and mega-pixel elasticity under either the pooled or separate estimation. From Table 6 and Figure 5, it is evident that for both homogeneous and random coefficients specifications, estimating demand separately for each market generates more dispersion in elasticities than the pooled estimation. The separate estimation relaxes the assumption made in the pooled estimation that coefficients across markets share a common heterogeneity distribution. Therefore, the estimates of the market specific models should better reflect local market conditions and geographic variations embedded in the data. In addition, the congestion term leads to a decrease of approximately 10% in price elasticity due to the varying number of products rather than consumer substitution.

Table 7 reports chain-specific elasticities averaged across markets with different competitive conditions. Overall, demand is more elastic in markets with more competing stores. When chains A and B compete in a market, their elasticities both increase about 11% compared to when each operates as a local monopolist. Under a market structure of A-D or

[11]Elasticities are only calculated for continuous variables. In each separate model, I include year, brand and chain dummies. Brand-chain interaction fixed effects are dropped due to data sparsity at the SSA level.

Table 7: Average Elasticities by Different Market Types

Chain	A Monopoly	B Monopoly	A-B Duopoly	A-D Duopoly	B-D Duopoly	A-B-D Triopoly
A	-1.761	—	-1.948	-1.824	—	-2.041
B	—	-1.659	-1.852	—	-1.796	-1.843
D	—	—	—	-1.713	-1.679	-1.702

Figure 5: Histograms of Market Specific Elasticities of Price and Mega-pixels
(upper: pooled estimation; lower: separate estimation)

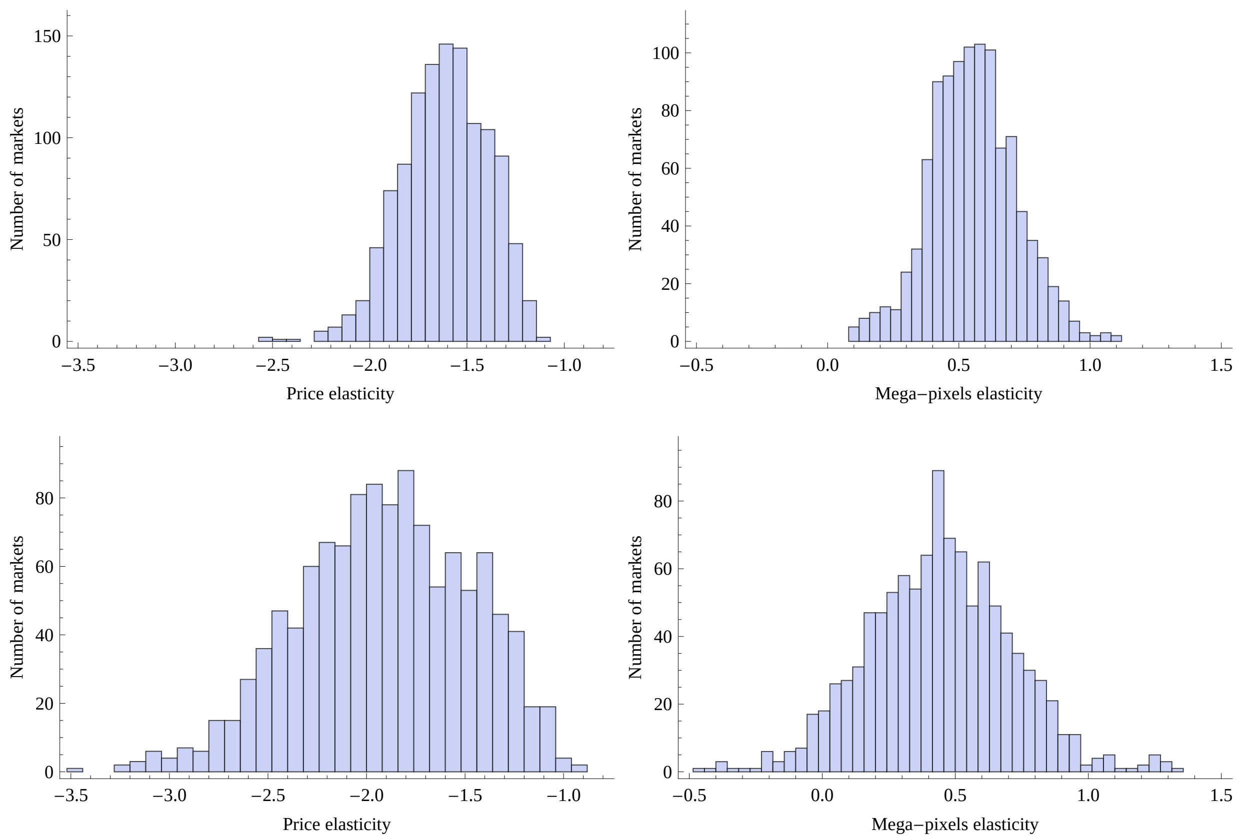


Table 8: Inferred Price Margin from Demand Estimates and Pricing Equilibrium

	Separate Estimation		Pooled Estimation	
Margin	2SLS	Random Coefficient & Microdata	2SLS	Random Coefficient & Microdata
Mean	69.62%	34.53%	82.05%	45.13%
Median	63.19%	28.59%	78.36%	41.08%
10% percentile	45.46%	21.24%	74.42%	33.87%
90% percentile	93.82%	42.89%	87.63%	58.40%

Note: Margin is defined as $(p - c)/p$.

B-D, chains A and B still have more elastic demand, although the increase is less relative to monopoly. As further evidence of the relationship between market competitiveness and elasticity, I calculate the physical distance between rival stores within a market and correlate it with the market-specific price elasticities. The correlations are -0.25 ($p < 0.01$) for A-B duopoly markets, -0.18 ($p < 0.01$) for A-D duopoly markets, -0.13 ($p < 0.01$) for B-D duopoly markets, and -0.22 ($p < 0.01$) for A-B-D triopoly markets. These negative and statistically significant correlations indicate that price elasticity tends to be higher in more competitive markets. In contrast, Hoch et al. (1995) find that local demographics explain more of the variation in store elasticities (estimated at the product category level) compared to local competitive conditions. The difference between the two results might be due to the fact that consumers shop differently for grocery products compared to electronics, and that consumers may be more likely to comparison shop for electronics since they represent more expensive purchases.

Table 8 compares the estimated price margins across alternative demand estimations. Marginal costs (assumed to be constant across geographical markets) are computed according to Equations (8) and (10) for every product given the corresponding pricing policy (i.e., national or local). The 2SLS estimates imply average price margins of approximately 70% and 82%, in the separate and pooled estimation, respectively. These margins are unrealistic for the digital cameras retail industry, reflecting the underestimated price elasticities in the

Table 9: Counterfactual Profits (π_A, π_B) under Alternative Pricing Policies (in $ millions)

		Chain B: Local	Chain B: National
Chain A	Local	(307.60, 104.06)	(320.58, 105.17)
	National	(310.03, 110.47)	(323.91, 112.78)

Table 10: Decompose Profit Difference between (National, National) and (Local, Local)

	$\Delta\pi$ Overall		$\Delta\pi$ in Markets where A and B do *not* compete		$\Delta\pi$ in Markets where A and B compete	
Chain	million $	percent	million $	percent	million $	percent
A	-16.31	-5.04%	4.09	3.00%	-20.40	-11.07%
B	-8.72	-7.74%	2.91	7.91%	-11.63	-15.33%

Note: $\Delta\pi = \pi(\text{Local}, \text{Local}) - \pi(\text{National}, \text{National})$

homogeneous models.[12] The pooled estimation leads to higher margins than the separate estimation. The congestion terms increase margins by approximately 10%. Overall, the random coefficients model with the micro moments and congestion leads to average price margins of approximately 35%, which are the closest to the reported margin in public reports. These results correcting for biases in demand estimates incorporating micro moments and congestion helps the demand model better estimate substitution patterns.

5.2 Counterfactual Simulation

I conduct counterfactual experiments to assess the impact of alternative pricing policies on firm profitability. First I consider the period prior to Chain B's exit. Specifically, I simulate equilibrium prices and profits when A and B choose between national and local pricing. Throughout the simulation I assume that Chain D, the large discount retailer, continues to use local pricing. This assumption seems reasonable given that digital camera sales make up a small portion of Chain D's overall sales, and so would be unlikely to change its general

[12]According to industry reports, such as Euromonitor (2010), the average margin for point-and-shoot cameras usually ranges from 25% to 35%.

product pricing strategy. Table 9 reports profits for A and B under the four possible pricing policy scenarios: Local-Local, Local-National, National-Local, and National-National. The results show that, under the existing market conditions, it was optimal for both firms A and B to employ national pricing. Consistent with Figure 4, which shows that both Chains A and B used nearly national pricing policy in the data, the profit increase between the purely national pricing and the observed national pricing policy is very small (less than 1%). The increase in profits between the purely national pricing and local pricing is 5.3% for Chain A and 8.4% for Chain B. Moreover, neither A nor B would find it profitable to deviate unilaterally from national pricing strategy. In a game between A and B in which they first choose a pricing policy and then set prices each period, the results in Table 9 constitute a sub-game perfect equilibrium with a national pricing policy.

Table 10 decomposes the difference in profits between National-National and Local-Local in order to highlight the rationale behind the enhanced profit under the national pricing policy. Moving from national to local pricing, both Chains A and B garner higher profits in markets in which they are a monopolist or only compete with D. Yet the chains lose profits due to the intensified competition in other markets in which they compete. Because the portion of the competitive markets is sufficiently large relative to the markets in which the two chains do not compete, for both chains the loss from the intensified competition is excessive and cannot be offset by the gains from the markets in which they have more market power. Accordingly, both chains become worse off by employing a local pricing policy.

Table 11 shows the average difference between the optimal price under the National-National scenario and each of the alternative pricing policy scenarios in markets in which Chains A and B compete, and in markets in which they do not compete. Compared to the national pricing, the prices under the local pricing policy are generally higher in the monopoly markets, and lower in the contested markets. Switching to a local policy, regardless of the other chain's policy, leads to an increase in the average price in the less competitive markets and a decrease in the competitive markets. Another interesting observation is the free rider effect, which is revealed in the last four columns of Table 11 (see also in Table 9). That is, if

Table 11: Average Price Difference from the National Pricing Policy across Market Types and Scenarios

	A Local B Local		A Local B National		A National B Local	
Chain	Type I	Type II	Type I	Type II	Type I	Type II
A	5.81%	-9.77%	5.81%	-6.23%	-2.54%	-2.63%
B	8.31%	-11.63%	-3.30%	-3.56%	8.31%	-7.17%

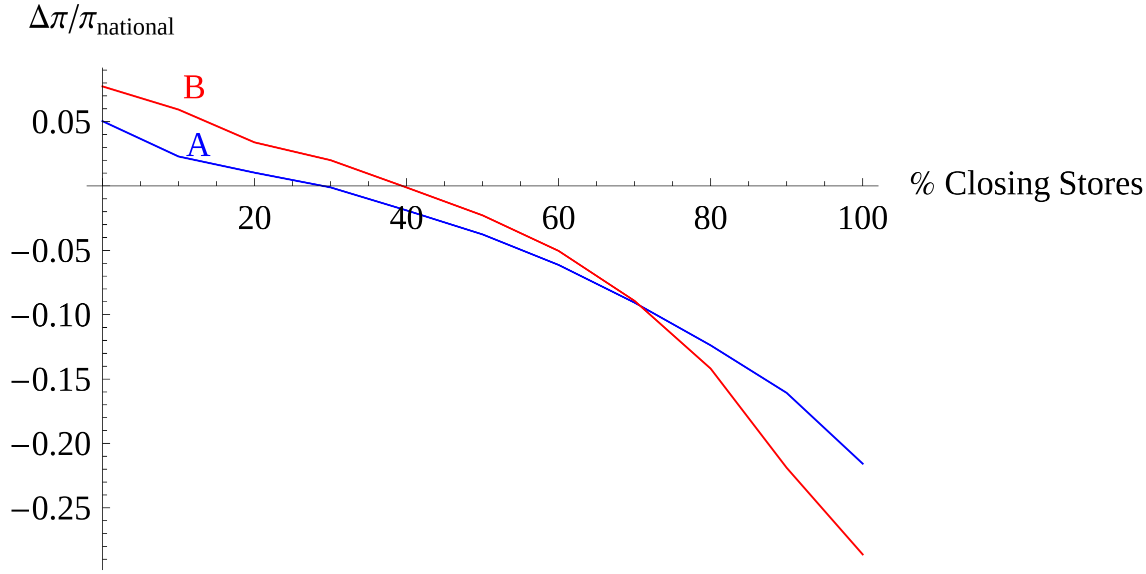
Note: Type I are markets in which A and B do *not* compete, Type II are markets in which A and B compete.

one chain chooses national pricing while the other chain chooses local pricing, the later chain would free ride the former chain. The chain with the national prices would lower its prices relative to the National-National scenario, while the chain with the local pricing policy could raise price in its own less competitive markets. Also, the profit of the chain that employs national prices would be close to the level of the Local-Local scenario.

The exit of Chain B eased the competitive landscape of the industry. The absence of such a large rival could create incentives for Chain A to localize prices as it became the single dominant chain. To investigate this possibility I simulate the optimal prices and profits after firm B exits. I find that under national pricing A's profit is $176.84 million, and under local pricing the profit is $174.60 million. I assume that Chain D maintains local pricing. The result implies that it is still optimal for Chain A to set prices uniformity across markets. The rationale behind this result is that Chain A still faces substantial competition from Chain D. As is evident from Table 3, Chain A faces competition from Chain D in 839 (84%) of the 1,004 markets in which it operates. Thus, the extent of competition between A and D is sufficient to justify national pricing even after Chain B's exit.

Because market structure is the most important factor affecting the decision to employ national or local pricing, I conduct another counterfactual experiment in which I directly vary the competitive landscape. Specifically, I gradually remove stores from the markets in which Chains A and B compete, leading to fewer competitive markets. After removing

Figure 6: Relative Profit Difference between National and Local Pricing as Market Structure Changes



every few stores, I solve the counterfactual profits under national and under local pricing and compare the results in Figure (6). As the number of competitive markets decreases, the profit gain from national pricing relative to local pricing due to softened competition declines. In particular, once Chain A retreats from at least 29% of its competitive markets, Chain A would benefit from employing local pricing. Similarly, Chain B would benefit from local pricing once it closes 40% of its stores in the competitive markets. The difference between Chains A and B is primarily due to the fact that Chain B originally had fewer stores operating in markets in which Chain A is not present. At the extreme, when a chain is a pure monopoly in all markets, then local pricing strictly dominates national pricing, which is consistent with previous findings (e.g., Chintagunta et al. 2003) where competition is absent. I also investigated a more realistic scenario under which the store-removing procedure removed store in increasing order of profit, thus removing low profit stores first. This procedure, which mimics the business reality in which struggling chains first close poor performance stores, provided qualitatively similar results.

6 Conclusion

In this paper I empirically examine a firm's choice of national versus local pricing policy in a multimarket competitive setting. To do so I estimate an aggregate model of demand with random coefficients separately in each of the over 1,000 markets. The separate estimation strategy leads to a significant increase in estimated heterogeneity across markets, reflecting the rich geographic variation in the data. I include a set of micro moments to improve model estimates and incorporate these moments into the recently proposed MPEC framework. I further control for product congestion to remove the confounds caused by varying number of products across markets and over time. The counterfactual policy simulation demonstrates that, relative to locally targeted pricing, national pricing results in substantially higher profit for the major retailers under the existing multimarket structure. The optimality of national pricing would hold as long as the ratio of competitive markets to non-competitive markets is high. These results have direct implications for the electronics retail industry. Furthermore, the insights from this investigation could generalize to other industries evaluating their chain-level pricing policies.

A few issues are left for future research. First, throughout the current analysis I assume marginal costs associated with the sales of digital cameras, and ignore any potential costs associated with the implementation of national or local pricing. For example, by moving from national to local, a chain may incur additional costs in customizing advertising to match locally varying prices.

Second, several recent papers have documented that durable goods buyers may strategically delay their purchases in anticipation of technology improvement and price decline (e.g., Song and Chintagunta 2003; Gordon 2009; Carranza 2010). Similarly, sellers may trade off between current and future profit by setting optimal price sequences (Zhao 2006). In this paper I ignore forward-looking dynamics on both the consumer and the retailer side. Given the nature of the research question, allowing for flexible consumer preferences at the market level is absolutely critical. Doing so in the context of a dynamic structural demand model is computationally challenging. Forward-looking behavior may also be less of a concern in this paper, given that the quality-adjusted prices in the period studied declined more slowly compared to the decline in earlier periods studied in previous research (e.g., Song and Chintagunta 2003).

Third, a more general model could endogenize the retailers' product assortment decisions. A retailer may have different incentives to stock a particular product under different pricing policies, and could also change the timing of a product's clearance period. This option would require an explicit model of multi-product retail assortment under competition. I plan to pursue this specific avenue in future research.

References

[1] Ackerberg, D., and M. Rysman (2005). "Unobserved product differentiation in discrete-choice models: estimating price elasticities and welfare effects," *RAND Journal of Economics*, 36, 771-788.

[2] Berry, S. (1994). "Estimating discrete-choice models of product differentiation," *RAND Journal of Economics*, 25, 242-262.

[3] Berry, S., J. Levinsohn and A. Pakes (1995). "Automobile prices in market equilibrium," *Econometrica*, 63, 841-890.

[4] Bernheim B. D. and M. D. Whinston (1990). "Multimarket contact and collusive behavior," *Rand Journal of Economics*, 21, 1-26.

[5] Borenstein, S. (1991). "Selling costs and switching costs: explaining retail gasoline margins," *Rand Journal of Economics*, 22, 354-369.

[6] Besanko, D., S. Gupta, D. Jain (1998). "Logit Demand Estimation Under Competitive Pricing Behavior: An Equilibrium Framework," *Management Science*, 44(11), Part 1, 1533-1547.

[7] Carranza, J. E. (2010). "Product innovation and adoption in market equilibrium: The case of digital cameras," *International Journal of Industrial Organization*, 28, 604-618.

[8] Chen, Y., C. Narasimhan, Z. J. Zhang (2001). "Individual Marketing with Imperfect Targetability," *Marketing Science*, 20(1), 23-41.

[9] Chintagunta, P. K., J. Dubé and V. Singh (2003). "Balancing profitability and customer welfare in a supermarket chain," *Quantitative Marketing and Economics*, 1, 111-147.

[10] Desai, P. S., O. Koenigsberg and D. Purohit (2010). "Forward buying by retailers," *Journal of Marketing Research*, 47 (February), 90-102.

[11] Dobson, P. W. and M. Waterson (2005). "Chain-store pricing across local markets," *Journal of Economics & Management Strategy*, 14, 39-119.

[12] Dubé J., J. T. Fox and C.-L. Su (2011). "Improving the numerical performance of BLP static and dynamic discrete choice random coefficients demand estimation," *Working paper, Booth School of Business, University of Chicago.*

[13] Eliashberg, J. and R. Chatterjee (1985). "Analytical Models of Competition with Implications for Marketing: Issues, Findings, and Outlook," *Journal of Marketing Research*, 22(3), 237-261.

[14] Ellickson, P. and S. Misra (2008). "Supermarket Pricing Strategies," *Marketing Science*, 27(5), 811-828.

[15] Euromonitor International (2010). "Country sector briefing, camera, U.S."

[16] Gordon, B. R. (2009). "A dynamic model of consumer replacement cycles in the PC processor industry," *Marketing Science*, 28(5), 846–867.

[17] Gordon B. R., A. Goldfarb and Y. Li (2011). "Does price elasticity vary with economic growth? A cross-category analysis," *Working Paper*.

[18] Greene, W. H. (2008). *Econometric analysis*. Pearson/Prentice Hall, Upper Saddle River, NJ.

[19] Heath, M. T. (2002). *Scientific computing: an introductory survey*. McGraw-Hill, Boston, MA.

[20] Hoch, S. J., B.-D. Kim, A. L. Montgomery and P. E. Rossi (1995). "Determinants of store-level price elasticity," *Journal of Marketing Research*, 32, 17-30.

[21] Lal, R. and R. Rao (1997). "Supermarket Competition: The Case of Every Day Low Pricing," *Marketing Science*, 16(1), 60-80.

[22] Li, Y. and A. Ansari (2011). "A semi-parametric Bayesian approach for endogeneity and heterogeneity in consumer choice models," *Working Paper*.

[23] Lou, W., D. Prentice and X. Yin (2008). "The effects of product ageing on demand: the case of digital cameras," *Working paper*.

[24] McGuire T. W. and R. Staelin (1983)."An industry equilibrium analysis of downstream vertical integration," *Marketing Science*, 2, 161-191.

[25] Nevo, A. (2000a). "Mergers with differentiated products: the case of the ready-to-eat cereal industry," *RAND Journal of Economics*, 31, 395-421.

[26] Nevo, A. (2000b). "A Practitioner's guide to estimation of random coefficients logit models of demand," *Journal of Economics & Management Strategy*, 9, 513-548.

[27] Petrin, A. (2002). "Quantifying the benefits of new products: the case of the minivan," *Journal of Political Economy*, 110, 705-729.

[28] PMA (2008). "U.S. consumer photo buying report."

[29] PMA (2009). "U.S. consumer photo buying report."

[30] PMA (2010). "U.S. consumer photo buying report."

[31] Rao, V. (1984). "Pricing Research in Marketing: The State of the Art," *Journal of Business*, 57(1), S39-S60.

[32] Shankar, V. and R. N. Bolton (2004). "An Empirical Analysis of Determinants of Retailer Pricing Strategy," *Marketing Science*, 23(1), 28-49.

[33] Sheppard, A. (1991). "Price Discrimination and Retail Configuration," *Journal of Political Economy*, 99(1), 30-53.

[34] Shubik, M. and R. Levitan (1980). *Market structure and behavior*. Harvard University Press, Cambridge, MA.

[35] Song I. and P. K. Chintagunta (2003). "A micromodel of new product adoption with heterogeneous and forward-looking consumers: application to the digital camera category," *Quantitative Marketing and Economics*, 1, 371-407.

[36] Song M. (2007). "Measuring consumer welfare in the CPU market: an application of the pure-characteristics demand model," *RAND Journal of Economics*, 38, 429-446.

[37] Su, C.-L. and K. L. Judd (2010). "Constrained optimization approaches to estimation of structural models," *Working Paper, University of Chicago*.

[38] Subramanian U., J. S. Raju, S. K. Dhar and Y. Wang (2010). "Competitive consequences of using a category captain," *Management Science*, 56, 1739-1765.

[39] Swait, J. and J. Louviere (1993), "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models," *Journal of Marketing Research*, 30(3), 305-314.

[40] Train, K. (2003). *Discrete choice methods with simulation*. Cambridge University Press.

[41] Wells J. R. and T. Haglock (2007). "Best Buy Co., Inc.: competing on the edge," *Harvard Business School Publishing*.

[42] Zettelmeyer, F. (2000). "Expanding to the Internet: Pricing and Communications Strategies When Firms Compete on Multiple Channels," *Journal of Marketing Research*, 37(3), 292-308.

[43] Zhao, Y. (2006). "Why are prices falling fast? An empirical study of the US digital camera market," *Working Paper*.

A Analytical Model of Multimarket Competition

I present an analytical model of multimarket price competition between retail chains. I start by deriving demand functions across markets with consistent underlying utility specification. Then I build the duopoly chain competition model to investigate the conditions under which national pricing generates more profit than localized pricing does. Building on Dobson and Waterson (2005), I allow for more flexibility and asymmetry cross markets.

The duopoly demand function is derived based on the quadratic utility specification introduced by Shubik and Levitan (1980), which has been widely used in the marketing literature to study duopoly competition (e.g., McGuire and Staelin, 1983; Desai et al., 2010; Subramanian et al., 2010). In the original specification both utility and demand are symmetric between the two competing goods. To accommodate asymmetry I follow the manipulation in Subramanian et al. (2010) to derive the demand function. Assume that, by consuming two goods a and b, a representative customer obtains the following quadratic utility of consumption, less the disutility of monetary expenditure,

$$U = \frac{1}{2}\left[\boldsymbol{\alpha}'\boldsymbol{\Theta}\boldsymbol{\alpha} - (\boldsymbol{\alpha} - \boldsymbol{q})'\boldsymbol{\Theta}(\boldsymbol{\alpha} - \boldsymbol{q})\right] - \beta\boldsymbol{p}'\boldsymbol{q} \tag{21}$$

where $\boldsymbol{q} = (q_a, q_b)'$ is the amount of consumption, $\boldsymbol{p} = (p_a, p_b)'$ is a vector of prices, and $\boldsymbol{\alpha} = (\alpha_a, \alpha_b)'$ denotes the amount of consumption that yields maximum utility. According to Subramanian et al. (2010), $\boldsymbol{\Theta}$ is a positive definite matrix and is normalized to be

$$\begin{pmatrix} \dfrac{1}{1+\theta} & \dfrac{\theta}{1+\theta} \\ \dfrac{\theta}{1+\theta} & \dfrac{1}{1+\theta} \end{pmatrix} \tag{22}$$

where $\theta \in [0, 1)$ denotes the degree of substitution between the two goods. When $\theta = 0$, they are completely independent of each other. When $\theta > 0$, the two goods are substitutable and the substitutability increases with θ. As $\theta \to 1$, the two goods approach perfect substitutes. In Desai et al. (2010) and Subramanian et al. (2010), the coefficient β on expenditure is normalized to one because these studies primarily focus on the difference between the two competing goods, for which β is a common multiplier. The current analysis, however, is to examine differences not only within a market but also across markets, and so I keep β in the demand model.

This representative customer maximizes her utility by setting the optimal amount of consumption, which results in the following duopoly demand function

$$
\begin{aligned}
q_a &= \alpha_a - \beta p_a + \frac{\beta\theta}{1-\theta}(p_b - p_a) \\
q_b &= \alpha_b - \beta p_b + \frac{\beta\theta}{1-\theta}(p_a - p_b)
\end{aligned} \tag{23}
$$

In a market where only one good is available, $\theta = 0$ and the utility function (21) reduces to

$$
U = \frac{1}{2}\left[\alpha^2 - (\alpha - q)^2\right] - \beta pq \tag{24}
$$

Accordingly, the monopoly demand is

$$
q = \alpha - \beta p \tag{25}
$$

Similar to Dobson and Waterson (2005), I hypothesize an industry with two chains, a and b, and three *independent* and *isolated* markets, 1, 2 and 3. The first two markets are monopolized by a and b, respectively, whereas the third market is a duopoly market where a and b compete. Assuming both chains are single-product firms, then demand in the three markets follows (23) and (25).

Under local pricing a chain makes price decision independently across markets. For instance, with this policy chain a solves two unrelated pricing problems given chain b's price in market 3

$$
\underset{p_{a1}}{\text{Max}}\ \pi_{a1}(p_{a1}) \quad \text{and} \quad \underset{p_{a3}}{\text{Max}}\ \pi_{a3}(p_{a3}|p_{b3})
$$

where profit $\pi_{a1}(p_{a1}) = q_{a1}p_{a1}$ and $\pi_{a3}(p_{a3}|p_{b3}) = q_{a3}p_{a3}$. On the other hand, under national pricing a chain pools demand across markets and sets a single optimal price to maximize chain profit. For example, chain a solves

$$
\begin{aligned}
&\underset{p_a}{\text{Max}} \quad \pi_{a1}(p_{a1}) + \pi_{a3}(p_{a3}|p_{b3}) \\
&\text{s.t.} \quad p_{a1} = p_{a3} = p_a
\end{aligned}
$$

Having specified the model setup, the game of multimarket chain competition proceeds in two stages. In the first stage, chains choose between national and local pricing policies; in the second stage, chains set optimal prices according the policy they have chosen in the

Table 12: Payoffs of the Two Stage Game

		Chain b National	Chain b Local
Chain a	National	π_{aN}, π_{bN}	π'_{aN}, π'_{bL}
	Local	π'_{aL}, π'_{bN}	π_{aL}, π_{bL}

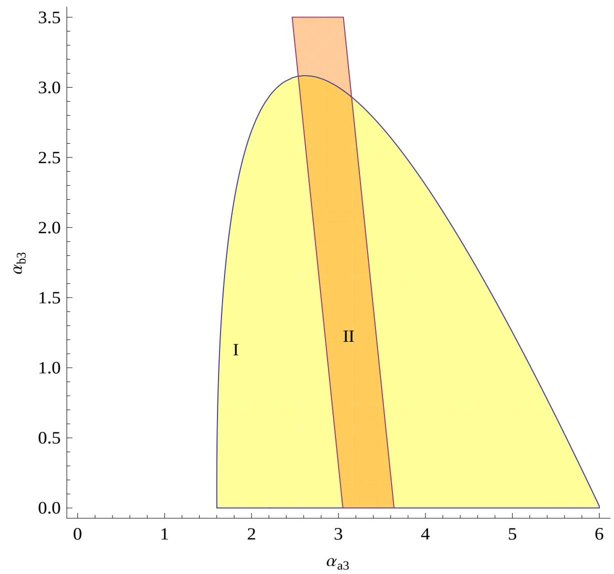
first stage. Table 12 summarizes all possible payoffs of the game, where the terms are given by:[13]

$$
\begin{aligned}
\pi_{1N} &= \frac{(\theta-1)[\beta_{a1}(\theta-1)-1][2(\alpha_{a1}+\alpha_{a3})(1+\beta_{b2})+\theta(\alpha_{b2}+\alpha_{b3}-2\beta_{b2}(\alpha_{a1}+\alpha_{a3})]^2}{[4(1+\beta_{a1})(1+\beta_{b2})-4\theta(\beta_{a1}+\beta_{b2}+\beta_{a1}\beta_{b2})+\theta^2(4\beta_{a1}\beta_{b2}-1)]^2} \\
\pi_{2N} &= \frac{(\theta-1)[\beta_{b2}(\theta-1)-1][2(\alpha_{b2}+\alpha_{b3})(1+\beta_{a1})+\theta(\alpha_{a1}+\alpha_{a3}-2\beta_{a1}(\alpha_{b2}+\alpha_{b3})]^2}{[4(1+\beta_{a1})(1+\beta_{b2})-4\theta(\beta_{a1}+\beta_{b2}+\beta_{a1}\beta_{b2})+\theta^2(4\beta_{a1}\beta_{b2}-1)]^2} \\
\pi'_{1L} &= \frac{\alpha_{a1}^2}{4\beta_{a1}} - \frac{(\theta-1)[\theta(\alpha_{b2}+\alpha_{b3})+2\alpha_{a3}(1+\beta_{b2}-\beta_{b2}\theta)]^2}{[\theta^2+4\beta_{b2}(\theta-1)-4]^2} \\
\pi'_{2N} &= \frac{(\theta-1)[\beta_{b2}(\theta-1)-1][2(\alpha_{b2}+\alpha_{b3})+\alpha_{a3}\theta]^2}{[\theta^2+4\beta_{b2}(\theta-1)-4]^2} \\
\pi'_{1N} &= \frac{(\theta-1)[\beta_{a1}(\theta-1)-1][2(\alpha_{a1}+\alpha_{a3})+\alpha_{b3}\theta]^2}{[\theta^2+4\beta_{a1}(\theta-1)-4]^2} \\
\pi'_{2L} &= \frac{\alpha_{b2}^2}{4\beta_{b2}} - \frac{(\theta-1)[\theta(\alpha_{a1}+\alpha_{a3})+2\alpha_{b3}(1+\beta_{a1}-\beta_{b1}\theta)]^2}{[\theta^2+4\beta_{a1}(\theta-1)-4]^2} \\
\pi_{1L} &= \frac{\alpha_{a1}^2}{4\beta_{a1}} - \frac{(\theta-1)(2\alpha_{a3}+\alpha_{b3}\theta)^2}{(\theta^2-4)^2} \\
\pi_{2L} &= \frac{\alpha_{b2}^2}{4\beta_{b2}} - \frac{(\theta-1)(2\alpha_{b3}+\alpha_{a3}\theta)^2}{(\theta^2-4)^2}
\end{aligned}
\tag{26}
$$

Since there are seven parameters in these profit functions, it is not possible to draw closed-form conclusion regarding the conditions under which one policy is better than the other. Therefore, in the remainder of the section I use numerical analysis to evaluate the analytical results.

[13] For the purpose of cross-market comparison, variable costs are set to zero, and the β in the duopoly market is normalized to one.

Figure 7: Contour Plot on $\Delta\pi = \pi_{aN} - \pi_{aL}$ against Varying Market Structure (α_{a1}=2, β_{a1}=1, α_{b2}=2, β_{b2}=1 and θ=0.5)



To show the profit enhancing effect of national pricing and see how such effect changes with market structure, I first examine the profit changes when a chain switches from national to local pricing, given the other chain does the same move. Figure 7 plots the profit difference for chain a ($\Delta\pi = \pi_{aN} - \pi_{aL}$) against both chains' strength in the duopoly market. The colored region I represents the range of α_{a3} and α_{b3} under which $\Delta\pi > 0$. The shape of the region presents several interesting implications. First, if national pricing is better than local pricing, the presence of chain a in the duopoly market cannot be too large or too small compared to its monopoly market. When the chain is very small, the profit gain in the duopoly market through national pricing cannot cover the loss in its monopoly market. On the other hand, when chain a is very large in market 3, the demand in its monopoly market is not sufficient to drive up the duopoly price, thereby hardly softening the competition. Further, if chain b is large in the duopoly market, it will be hard for chain a to raise the overall price level in this market. Hence, a chain prefers national pricing over local pricing only if this chain has a medium presence in the duopoly market, and the other chain is not too large either.

Next, I examine the conditions under which national pricing is an equilibrium of this game. When $\Delta\pi_a = \pi_{aN} - \pi'_{aL} > 0$ and $\Delta\pi_b = \pi_{bN} - \pi'_{bL} > 0$ both hold, national pricing becomes the dominant strategy for both chains. The colored region II in Figure 7 describes the ranges under which the equilibrium exists. This range is smaller than the previous case because there is a free-rider issue. Suppose a chain switches to local pricing while the other chain still plays national, the first chain will reap the maximal profit from its own monopoly markets, while benefiting from the eased competition in the duopoly market through the other chain's national pricing. However, the duopoly competition is anyway intensified in the duopoly market, so the first chain may still find it worthwhile to go back to national pricing. Again, the intuition behind the equilibrium outcome is basically the same as before.

The analytical model highlights the rationale on how geographic pricing policies affect chain profitability. However, as this is a simple hypothetical model, it does not reflect the complexity of multimarket chain competition in the real world. For example, real markets usually contain more than two competing firms, so the results from duopoly competition may not be generalizable to a market with more than 2 firms. Also, chain stores sell multiple differentiated products and thus internal competition exists within a chain. The simplification in the analytical model on single-product chains with linear demand may also have an impact on pricing policy choice. For these reasons, in this paper I primarily rely on real data to investigate chain store pricing.

B Analytic Derivatives for MPEC Estimation

In this section I derive the analytic derivatives of the optimization problem specified in (17). My derivation follows matrix calculus and employs tensor operators such as Kronecker product. Thanks to the sparsity of the optimization problem (i.e., shares being independent across markets), all Kronecker products that appear in the middle of the derivation drop out in the final results, so computing speed is not affected by these Kronecker products. I write the derivatives compactly in matrix notation so one can easily code them in computer programs.

The gradient and Hessian of the GMM objective function $F(\boldsymbol{\phi})$ are respectively

$$\frac{\partial F(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}} = (\boldsymbol{W} + \boldsymbol{W'})\boldsymbol{\eta} \tag{27}$$

$$\frac{\partial^2 F(\boldsymbol{\phi})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'} = \boldsymbol{W} + \boldsymbol{W'} \tag{28}$$

The Jacobian matrices of the constraints imposed by share equations are

$$\frac{\partial \boldsymbol{s}_t(\boldsymbol{\delta}_t, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2} = \int_{\forall i} \text{diag}(\boldsymbol{s}_{it})[\boldsymbol{X}_{it}^{rc} - \mathbf{1}_{J_t} \boldsymbol{s}_{it}^{'} \boldsymbol{X}_{it}^{rc}]\text{diag}(\boldsymbol{v}_i) \tag{29}$$

$$\frac{\partial \boldsymbol{s}_t(\boldsymbol{\delta}_t, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\delta}_t} = \int_{\forall i} \text{diag}(\boldsymbol{s}_{it}) - \boldsymbol{s}_{it} \boldsymbol{s}_{it}^{'} \tag{30}$$

where $\mathbf{1}_{J_t}$ is a J_t-element column vector of ones.

The Jacobian matrices of the constraints imposed by the demand side orthogonal conditions are

$$\frac{\partial[\boldsymbol{\eta}_1 - \boldsymbol{g}(\boldsymbol{\delta}, \boldsymbol{\theta}_1)]}{\partial \boldsymbol{\theta}_1} = \frac{1}{N_d}\boldsymbol{Z}'\boldsymbol{X} \tag{31}$$

$$\frac{\partial[\boldsymbol{\eta}_1 - \boldsymbol{g}(\boldsymbol{\delta}, \boldsymbol{\theta}_1)]}{\partial \boldsymbol{\delta}} = -\frac{1}{N_d}\boldsymbol{Z}' \tag{32}$$

$$\frac{\partial[\boldsymbol{\eta}_1 - \boldsymbol{g}(\boldsymbol{\delta}, \boldsymbol{\theta}_1)]}{\partial \boldsymbol{\eta}_1} = \boldsymbol{I}_{nz} \tag{33}$$

The Jacobian matrices of the constraints imposed by the micro moments are

$$\frac{\partial[\boldsymbol{\eta}_2 - \tilde{\boldsymbol{s}}_{rt}(\boldsymbol{\delta}_t, \boldsymbol{\theta}_2)]}{\partial \boldsymbol{\theta}_2} = -\int_{i\in r} s_{i0t}\boldsymbol{s}_{it}^{'}\boldsymbol{X}_{it}^{rc}\mathrm{diag}(\boldsymbol{v}_i) \tag{34}$$

$$\frac{\partial[\boldsymbol{\eta}_2 - \tilde{\boldsymbol{s}}_{rt}(\boldsymbol{\delta}_t, \boldsymbol{\theta}_2)]}{\partial \boldsymbol{\delta}_t} = -\int_{i\in r} s_{i0t}\boldsymbol{s}_{it}^{'} \tag{35}$$

The Hessian vector[14] of all the constraints at the $\boldsymbol{\theta}_2$ by $\boldsymbol{\theta}_2$ block

$$\sum_{\forall j,t}\lambda_{jt}\frac{\partial^2 s_{jt}(\boldsymbol{\delta}_t, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\theta}_2\boldsymbol{\theta}_2^{'}} = \sum_{t=1}^{T}\int_{\forall i}\mathrm{diag}(\boldsymbol{v}_i)[(\boldsymbol{X}_{it}^{rc'} - \boldsymbol{X}_{it}^{rc'}\boldsymbol{s}_{it}\boldsymbol{1}_{J_t}^{'})\mathrm{diag}(\boldsymbol{\lambda}_t) - \boldsymbol{\lambda}_t^{'}\boldsymbol{s}_{it}\boldsymbol{X}_{it}^{rc'}]\frac{\partial \boldsymbol{s}_{it}}{\partial \boldsymbol{\theta}_2} \tag{36}$$

$$\sum_{\forall r,t}\lambda_{rt}\frac{\partial^2[\boldsymbol{\eta}_2 - \tilde{\boldsymbol{s}}_{rt}]}{\partial \boldsymbol{\theta}_2\boldsymbol{\theta}_2^{'}} = \sum_{\forall r,t}\lambda_{rt}\int_{i\in r} s_{i0t}\mathrm{diag}(\boldsymbol{v}_i)\boldsymbol{X}_{it}^{rc'}[\boldsymbol{s}_{it}\boldsymbol{s}_{it}^{'}\boldsymbol{X}_{it}^{rc}\mathrm{diag}(\boldsymbol{v}_i) - \frac{\partial \boldsymbol{s}_{it}}{\partial \boldsymbol{\theta}_2}] \tag{37}$$

[14] The following linear operation is found particularly useful in deriving the Hessian from the Jacobian as one needs to take derivatives over the diagonal matrix of share vectors. For example, an n-by-n diagonal matrix $\mathrm{diag}(\boldsymbol{s})$ with a vector $\boldsymbol{s}$ on its diagonal can be transformed linearly by

$$\mathrm{diag}(\boldsymbol{s}) = \sum_{i=1}^{n}\boldsymbol{E}_i\boldsymbol{s}\boldsymbol{e}_i^{'}$$

where $\boldsymbol{E}_i$ is a n-by-n matrix of all zeros except the i-th diagonal entry is one, and $\boldsymbol{e}_i$ is a vector of all zeros except the i-th element is one. Since the transformation is linear, the derivative of the diagonal matrix can be compactly written as

$$\frac{\partial \mathrm{diag}(\boldsymbol{s})}{\partial \boldsymbol{\delta}} = \sum_{i=1}^{n}(\boldsymbol{e}_i \otimes \boldsymbol{E}_i)\frac{\partial \boldsymbol{s}}{\partial \boldsymbol{\delta}}$$

where $\otimes$ denotes Kronecker product.

where $\frac{\partial \boldsymbol{s}_{it}}{\partial \boldsymbol{\theta}_2}$ is calculated as in (29) without the integration. $\boldsymbol{\lambda}_t$ is a vector of the Lagrange multipliers corresponding to the share equations at t.

The Hessian vector of all the constraints at the $\boldsymbol{\delta}_t$ by $\boldsymbol{\theta}_2$ block

$$\sum_{\forall j,t} \lambda_{jt} \frac{\partial^2 s_{jt}(\boldsymbol{\delta}_t, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\delta}_t \boldsymbol{\theta}_2^{'}} = \sum_{t=1}^{T} \int_{\forall i} [\text{diag}(\boldsymbol{\lambda}_t) - \boldsymbol{\lambda}_t^{'} \boldsymbol{s}_{it} \boldsymbol{I}_{J_t} - \boldsymbol{s}_{it} \boldsymbol{\lambda}_t^{'}] \frac{\partial \boldsymbol{s}_{it}}{\partial \boldsymbol{\theta}_2} \tag{38}$$

$$\sum_{\forall r,t} \lambda_{rt} \frac{\partial^2 [\boldsymbol{\eta}_2 - \tilde{\boldsymbol{s}}_{rt}]}{\partial \boldsymbol{\delta}_t \boldsymbol{\theta}_2^{'}} = \sum_{\forall r,t} \lambda_{rt} \int_{i \in r} s_{i0t} [\boldsymbol{s}_{it} \boldsymbol{s}_{it}^{\prime} \boldsymbol{X}_{it}^{rc} \text{diag}(\boldsymbol{v}_i) - \frac{\partial \boldsymbol{s}_{it}}{\partial \boldsymbol{\theta}_2}] \tag{39}$$

The Hessian vector of all the constraints at the $\boldsymbol{\delta}_t$ by $\boldsymbol{\delta}_t$ block:

$$\sum_{\forall j,t} \lambda_{jt} \frac{\partial^2 s_{jt}(\boldsymbol{\delta}_t, \boldsymbol{\theta}_2)}{\partial \boldsymbol{\delta}_t \boldsymbol{\delta}_t^{'}} = \sum_{t=1}^{T} \int_{\forall i} [\text{diag}(\boldsymbol{v}_i) - \boldsymbol{\lambda}_t \boldsymbol{s}_{it}^{'} \boldsymbol{I}_{J_t} - \boldsymbol{s}_{it} \boldsymbol{\lambda}_t^{'}] \frac{\partial \boldsymbol{s}_{it}}{\partial \boldsymbol{\delta}_t} \tag{40}$$

$$\sum_{\forall r,t} \lambda_{rt} \frac{\partial^2 [\boldsymbol{\eta}_2 - \tilde{\boldsymbol{s}}_{rt}]}{\partial \boldsymbol{\delta}_t \boldsymbol{\delta}_t^{'}} = \sum_{\forall r,t} \lambda_{rt} \int_{i \in r} s_{i0t} [2 \boldsymbol{s}_{it} \boldsymbol{s}_{it}^{\prime} - \text{diag}(\boldsymbol{s}_{it})] \tag{41}$$

where $\frac{\partial \boldsymbol{s}_{it}}{\partial \boldsymbol{\delta}_t}$ is calculated as in (30) without the integration.

Once the optimization converges, standard errors of the parameter estimates are obtained through (20). The Jacobian matrix of the two sets of moments with respect to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ is

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\theta}_1} & \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\theta}_2} \\ \\ \frac{\partial \tilde{\boldsymbol{s}}}{\partial \boldsymbol{\theta}_1} & \frac{\partial \tilde{\boldsymbol{s}}}{\partial \boldsymbol{\theta}_2} \end{pmatrix} \tag{42}$$

where

$$\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\theta}_1} = -\frac{1}{N_d} \boldsymbol{Z}^{\prime} \boldsymbol{X} \tag{43}$$

$$\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{\theta}_2} = \frac{1}{N_d}\boldsymbol{Z}'(\frac{\partial \boldsymbol{s}_t}{\partial \boldsymbol{\delta}_t})^{-1}\frac{\partial \boldsymbol{s}_t}{\partial \boldsymbol{\theta}_2} \tag{44}$$

$$\frac{\partial \tilde{s}_{rt}}{\partial \boldsymbol{\theta}_1} = \left(\int_{i\in r} s_{i0t}\boldsymbol{s}'_{it}\right)\boldsymbol{X}_t \tag{45}$$

$$\frac{\partial \tilde{s}_{rt}}{\partial \boldsymbol{\theta}_2} = \int_{i\in r} s_{i0t}\boldsymbol{s}'_{it}\boldsymbol{X}^{rc}_{it}\mathrm{diag}(\boldsymbol{v}_i) \tag{46}$$

The second moments $\boldsymbol{\Phi}$ is

$$\begin{pmatrix} \boldsymbol{\Phi_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi_1} \end{pmatrix} \tag{47}$$

where

$$\boldsymbol{\Phi_1} = \frac{1}{N_d}\sum_{j,t}\xi_{jt}^2\boldsymbol{Z}_{jt}\boldsymbol{Z}'_{jt} \tag{48}$$

$$\boldsymbol{\Phi_2} = \frac{1}{I}\,\mathrm{diag}\left(\sum_{i}^{I}(\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{S}})^2\right) \tag{49}$$

C Scaling for the micro moments

As the survey information is not immediately usable in the estimation, scaling is needed to match the survey statistics to the geographic variation and the actual market size the model uses. Assume a survey gives average purchase probabilities for four income segments, A, B, C, and D, at the national level. I need to obtain a, b, c, and d for the corresponding four income segments in each market. First, from the market-specific income distribution $\mathrm{P}(y_i)$, I obtain the weights for each segments

$$w_r = \int_{i\in r} d\mathrm{P}(y_i)$$

where $r = 1, 2, 3, 4$. With the total share of all inside options $\tilde{S}_t$ observed in the sales data, I solve the following four equations to get a, b, c, and d

$$
\begin{aligned}
\tilde{S}_t &= w_1 a + w_2 b + w_3 c + w_4 d \\
a/b &= A/B \\
b/c &= B/C \\
c/d &= C/D
\end{aligned}
$$

D Data trimming

The raw data on store sales include nearly ten million observations. I trim the data before applying it to estimate the econometric model. I first remove SSAs where none of the three major chains had a presence. Then I delete all cameras that are not compact point-and-shoot. Third, I retain only sales of the top seven brands. Fourth, I get rid of all observations in 2010, due to the right truncation issue at the end of the observation period. Fifth, I remove observations with unreasonably high or low prices, as these are most likely data collection errors. Lastly, in each chain I sort camera models from largest to smallest market share and include models that yield a cumulative market share of at least 80%. I perform the last step year-by-year because of the frequent product entries and exits.