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

Yang Li*

Columbia University

April 1, 2012

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 the major retail firms should employ a national pricing policy to maximize profits, rather than target prices in each local market. Fixing prices across markets allows the retailers to soften otherwise intense local competition by subsidizing competitive markets with profits 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 policies.

*Graduate School of Business, Columbia University. Email: yli14@gsb.columbia.edu. I wish to sincerely 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 each of the two chains, 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

---

¹Evidence 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.
²In the remainder of the paper, I use the terms national, uniform, and fixed interchangeably to refer to the policy of national pricing.
³Appendix 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.

To flexibly recover local consumer preferences, I estimate an aggregate model of demand with random coefficients separately in each of the more than 1,500 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 with 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 is necessary because they propagate into the supply-side estimation.

Although consumer preferences are estimated without any equilibrium assumptions, 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 vs. 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 profits 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 relate 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; Duan and Mela 2009). Previous studies on geographic price discrimination, however, generally neglect the effect of pricing competition in the multimarket context. 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), Bronnenberg (2008), 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 presents robustness tests of local market definition and counterfactual outcomes. Section 7 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) online vs. offline shopping statistics from Mintel, (4) store location data from AggData, (5) digital camera sales across channels from Euromonitor, and (6) 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 United States 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 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 almost double other quarterly sales. The data set contains nearly 60
unique camera brands. I focus my analysis on the largest seven brands, which account for
approximately 80% of sales, as reported in Table 2. The NPD data also contain a detailed
description of the product characteristics for each camera model, such as mega-pixels, optical
zoom, thickness, weight, display size, and face detection.

Second, I use consumer survey data from the market research firm PMA to augment
the estimation with micro moments, which relate digital camera purchases with household
income. 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

Table 1: Descriptive Statistics of the Store Sales Data

<table>
<thead>
<tr>
<th>Quarter</th>
<th>Total Revenue ($ billion)</th>
<th>Total Sales (million units)</th>
<th>Sales Weighted Average Price ($)</th>
<th># Camera Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007 Q1</td>
<td>0.617</td>
<td>2.834</td>
<td>215.79</td>
<td>940</td>
</tr>
<tr>
<td>2007 Q2</td>
<td>0.787</td>
<td>3.580</td>
<td>209.76</td>
<td>1016</td>
</tr>
<tr>
<td>2007 Q3</td>
<td>0.718</td>
<td>3.253</td>
<td>210.98</td>
<td>1060</td>
</tr>
<tr>
<td>2007 Q4</td>
<td>1.446</td>
<td>8.289</td>
<td>177.25</td>
<td>1092</td>
</tr>
<tr>
<td>2008 Q1</td>
<td>0.616</td>
<td>3.113</td>
<td>185.18</td>
<td>1143</td>
</tr>
<tr>
<td>2008 Q2</td>
<td>0.853</td>
<td>4.047</td>
<td>189.35</td>
<td>1149</td>
</tr>
<tr>
<td>2008 Q3</td>
<td>0.669</td>
<td>3.268</td>
<td>187.33</td>
<td>1167</td>
</tr>
<tr>
<td>2008 Q4</td>
<td>1.164</td>
<td>7.346</td>
<td>154.13</td>
<td>1172</td>
</tr>
<tr>
<td>2009 Q1</td>
<td>0.521</td>
<td>2.850</td>
<td>161.78</td>
<td>1203</td>
</tr>
<tr>
<td>2009 Q2</td>
<td>0.647</td>
<td>3.351</td>
<td>179.32</td>
<td>1284</td>
</tr>
<tr>
<td>2009 Q3</td>
<td>0.533</td>
<td>2.677</td>
<td>186.14</td>
<td>1162</td>
</tr>
<tr>
<td>2009 Q4</td>
<td>1.055</td>
<td>6.711</td>
<td>155.76</td>
<td>1132</td>
</tr>
<tr>
<td>2010 Q1</td>
<td>0.475</td>
<td>2.408</td>
<td>163.25</td>
<td>1118</td>
</tr>
</tbody>
</table>
Table 2: Annual Market Shares (%) of Top Camera Brands

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand 1</td>
<td>21.4</td>
<td>21.5</td>
<td>21.6</td>
</tr>
<tr>
<td>Brand 2</td>
<td>17.0</td>
<td>19.1</td>
<td>20.6</td>
</tr>
<tr>
<td>Brand 3</td>
<td>7.3</td>
<td>11.1</td>
<td>12.8</td>
</tr>
<tr>
<td>Brand 4</td>
<td>16.4</td>
<td>13.7</td>
<td>12.2</td>
</tr>
<tr>
<td>Brand 5</td>
<td>6.4</td>
<td>6.1</td>
<td>5.6</td>
</tr>
<tr>
<td>Brand 6</td>
<td>5.5</td>
<td>5.4</td>
<td>5.2</td>
</tr>
<tr>
<td>Brand 7</td>
<td>3.8</td>
<td>4.5</td>
<td>5.2</td>
</tr>
<tr>
<td>Total</td>
<td>77.8</td>
<td>81.4</td>
<td>83.2</td>
</tr>
</tbody>
</table>

of households at different income levels that bought a new digital camera. I use these proportions to construct the micro moments for demand estimation. Note that the PMA data aggregate across online and offline purchases, whereas the current study focuses on purchases in brick-and-mortar stores. Therefore, I bring in an online shopping report from the market research firm Mintel. The report provides probabilities of buying offline versus online by household income across categories. I use the statistics for household electronics to scale the PMA survey data to obtain the likelihoods of offline camera purchases at different income levels.

Third, I use the store location data from AggData to help define and validate the competitive selling areas. NPD splits the United States into 2,100 distinct geographic markets called store selling areas (SSAs), which define competitive markets. Ninety-five percent of SSAs contain only one store of each major retailer. The median distance between competing stores within an SSA is 0.58 miles, whereas the median and the bottom 5th-percentile distance to competing stores in neighboring SSAs are 10.20 and 3.45 miles, respectively. Thus, to some extent, the SSA definition captures distinct geographic markets, with retail stores located nearby within a market and relatively farther from stores outside their SSAs. In the robustness check section, I present a more structured test on market definition. 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, one must carefully control for geographic difference when modeling cameras sales at retail 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., Berry et al. 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 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 eventually choose not to purchase a camera. To estimate this group, I obtain 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 the analytical model in Appendix A illustrates, the relative advantage between national and local pricing relies on the characteristics of market structure, in particular, the size of competitive markets versus monopoly markets, and degree of local competition. Next,

\(^5\)For example, in a homogeneous logit model, the mark-up across all products of a firm is a constant and it is negatively related to market size.
I describe the patterns we observe in the data regarding these characteristics.

The retail digital camera market is concentrated with three national chains, A, B, and D, accounting for 70% of U.S. sales. Other retailers had shares below 3%. Chains A and B are specialty retailers of consumer electronics, whereas 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% of the shares of the U.S. market, respectively. At the end of 2008, chain B terminated operations and liquidated all stores within three months (for reasons mostly independent of camera sales). The market share B left was immediately taken up by A, making A the dominant national player with almost 60% of the entire U.S. digital camera market. Chain D maintained an approximate 9% share throughout the period. As a result of the concentrated market structure, in the current study, I focus on the competition between these three big-box retail chains. Accordingly, I remove the local markets in which none of the three firms exist, thereby leading to the three major chains accounting for almost 90% of the market share in the areas in which they operate. I group all small sellers in these areas into a single chain L.

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 operated in a mixture of monopolist-like 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 did not coexist and only faced competition either from Chain D or those small retailers, which are not shown in this table.

Given these market conditions, whether a firm would prefer a national or local pricing policy is unclear. On the 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 a
national pricing policy to ease the competition. Both the distributions 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 A would find switching to local pricing following Chain B’s exit optimal depends on the relative size of these markets and the intensity of competition in its other markets. Although Chain A gained monopoly markets, it 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
Table 3: Market Type, Number of Markets, and Average Annual Sales

<table>
<thead>
<tr>
<th>Market Type</th>
<th>Before B Left</th>
<th>After B Left</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># SSAs Sales</td>
<td># SSAs Sales</td>
</tr>
<tr>
<td>A only</td>
<td>101 0.62</td>
<td>165 1.27</td>
</tr>
<tr>
<td>A &amp; D</td>
<td>315 2.51</td>
<td>839 7.60</td>
</tr>
<tr>
<td>B only</td>
<td>79 0.33</td>
<td>— —</td>
</tr>
<tr>
<td>B &amp; D</td>
<td>118 0.71</td>
<td>— —</td>
</tr>
<tr>
<td>A &amp; B</td>
<td>59 0.76</td>
<td>— —</td>
</tr>
<tr>
<td>A, B, &amp; D</td>
<td>402 5.60</td>
<td>— —</td>
</tr>
<tr>
<td>D only</td>
<td>525 0.85</td>
<td>600 1.10</td>
</tr>
</tbody>
</table>

Note: Sales are in million units.

A tended to differentiate itself from Chain B when the two 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.

2.3 Pricing Policies

Both retailers A and B used nearly national pricing policies: prices for each product are almost 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 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.

Figure 3 presents the coefficients of variation in the sales-weighted price across stores for all products relative to their cumulative share of lifetime sales.\(^6\) Each dot in the graph represents a camera model in a month. 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

\(^6\)To determine the cumulative sales of the products that entered prior to January 2007, I use national sales data from NPD aggregated over stores from January 2000 to March 2010.
constant over a product’s lifecycle. The little observed dispersion for Chains A and B can be attributed to three sources under a national pricing policy. First, I must derive unit prices from the monthly sales data containing product-level revenue and volume in each store. This aggregation leads to differences in monthly average product price across stores. Second, 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. Third, measurement error in either the revenue or volume would generate apparent price variation. All these errors will be absorbed into an unobservable demand shock term 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. Also, the chains adopting national pricing claimed (e.g., on their website)
Figure 3: Price Dispersion in Chain A (top), B (middle), and D (bottom)
company policy dictates store price and online price should generally match. In addition, these chains offered price-match guarantees that compensate price difference for the sales from their own stores. In contrast, the chains with local pricing policies made no claims regarding price uniformity and said explicitly that online prices were excluded from their companies’ price-match guarantees.

3 Model

This section provides a market-specific aggregate demand model 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-camera purchasing patterns, and (2) a “congestion” term that addresses variation in assortment size 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

\[
U_{ijt} = (y_i - p_{jt})^\alpha G(x_{jt}, \xi_{jt}, \beta_i) e^{\epsilon_{ijt}},
\]

where \( t=1,\ldots,T \) is the index for month and \( j=1,\ldots,J_t \) denotes the set of products at \( t \). \( x_{jt} \) are observed product characteristics with coefficients \( \beta_i \).\(^7\) \( \xi_{jt} \) 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 \( \alpha \) is the price coefficient indicating the marginal utility of expenditures. For

\(^7\)Bold fonts denote vectors or matrices. All vectors are by default column vectors.
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.\footnote{One issue of using a Cobb-Douglas utility is that income must be larger than price before taking logs. With simulated income draws, some of these draws could fall below price and violate this condition. In the current study, digital camera price is much lower than average monthly income, so the estimation bias caused by the sample selection on income is negligible.}

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

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

Accordingly, the utility for the outside option $j=0$ is

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

Assuming $\epsilon$’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

$$s_{jt} = \int_{\forall i} s_{ijt} = \int_{\forall i} \frac{\exp[x'_{jt}b_i + \alpha \log(1 - p_{jt}/y_i) + \xi_{jt}]}{1 + \sum_{k=1}^J \exp[x'_{kt}b_i + \alpha \log(1 - p_{kt}/y_i) + \xi_{kt}]}dP(b_i)dP(y_i), \tag{4}$$

where $P(b_i)$ and $P(y_i)$ are probability density functions of heterogeneous tastes and household income, respectively. Following the literature, I assume $b_i$ is normally distributed and estimate the distribution of $y_i$ from the Census data. The normality assumption on consumer heterogeneity may cause estimation bias if the actual distribution is heavily tailed or multi-mode, as demonstrated by Li and Ansari (2012). To allow for flexible heterogeneity distribution, I estimate the demand model separately for each local 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 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
Table 4: Percent of Households that Purchased a New Camera

<table>
<thead>
<tr>
<th>Year</th>
<th>&lt; $29,999</th>
<th>$30,000–$49,999</th>
<th>$50,000–$74,999</th>
<th>&gt; $75,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>8%</td>
<td>16%</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td>2008</td>
<td>8%</td>
<td>12%</td>
<td>14%</td>
<td>18%</td>
</tr>
<tr>
<td>2009</td>
<td>7%</td>
<td>11%</td>
<td>14%</td>
<td>15%</td>
</tr>
</tbody>
</table>

not be sufficient to estimate a full set of random coefficients, I further decompose \( x_{jt} \) into \( x_{jt}^{fc} \) and \( x_{jt}^{rc} \), and assign random coefficients only to \( x_{jt}^{rc} \). \( x_{jt}^{rc} \) includes mega-pixels, store affiliation, and camera brand. The other three non-price attributes are included in \( x_{jt}^{fc} \). 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 \( x_{jt}^{fc} \) 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 constructing nested choice models impossible. Also, it is not clear that consumers actually follow the nested process and choose stores before selecting products. Moreover, Berry (1994) shows that nested logit is a special case of random coefficient logit in modeling aggregate demand, and the latter allows more complicated correlation patterns between products. Thus I treat store affiliation as an additional product attribute that additively enters into consumer utility function.

3.2 Micro Moments

Leveraging information that links average consumer demographics to consumers’ purchase behavior can improve estimates from aggregate models (Petrin 2002). I divide each market into \( R \) distinct income tiers, with varying price coefficients across these tiers:

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

15
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 these tiers (Table 4). In demand estimation, I construct additional micro moments according to

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

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, three modifications are necessary before constructing the micro moments. First, the PMA survey adds up both online and offline sales; therefore, it is inconsistent with the NPD data and the store demand model. I use additional statistics from Mintel regarding online versus offline buying probabilities by household income to calibrate the PMA survey responses. Second, 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. Third, 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 $\epsilon$'s) changes with the number of products. These restrictions can bias elasticity estimates if substantial variation exists in the number of products across markets.
or over 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 happens in the observed characteristics space but not in the unobservable characteristics space. With each new product, a new i.i.d. $\epsilon$ is added, and the dimensionality of the unobservable characteristics space is expanded. Price sensitivity can be estimated without price variation and solely based on the 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 equally differentiated along unobserved characteristics but constrained by the bound. The bound is implemented as a congestion term $\log(R_{jt})$, where $R_{jt} = J^\gamma_t$, and $\gamma$ is a parameter to estimate. I add the congestion term to the model and report the comparison with and without this correction in Section 5.

### 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 in every market. In each month, the chain sets prices according to the overall policy.

The demand estimation is free of equilibrium assumptions in order to accurately recover

---

9An alternative specification is $R_{jt} = \gamma/J_t + 1 - \gamma$. 
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 a 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 local market in which the chain operates. \( M_m \) represents the size of market \( m \) and \( s_{jm} \) is the share of product \( j \) in market \( m \).

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

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

Stacking prices and costs and aligning simulated shares across markets, the pricing equation
can be written in matrix notation for all competing chains:

$$c = p - \Delta^{-1}q,$$

(8)

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

$$\Delta_f = -\sum_{\forall m} M_m \int_{i \in m} \left[ \frac{\partial \mu_i(p)}{\partial p} \left( \text{diag}(s_i) - s_i s'_i \right) \right].$$

(9)

Here the integration is specific to the demographic distribution in market $m$.

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

$$c = p - \Delta^{-1}s,$$

(10)

where $s = \int s_i$ is a vector of product shares, and

$$\Delta_f = -\frac{\partial \mu_i(p)}{\partial p} \left( \text{diag}(s_i) - s_i s'_i \right).$$

(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.\(^\text{10}\) Based on the pricing patterns observed in the data section, I assume A and B fixed price across markets for each camera before sales of the camera hit the 80% threshold of its lifetime sales. In calculating marginal costs,

\(^{10}\)Here I assume the price equilibrium to both (8) and (10) exists and the equilibrium is unique. For a homogeneous logit demand model, Caplin and Nalebuff (1991) lay out the set of conditions under which equilibrium exists for single-product firms. Berry et al. (2004) point out the same set of conditions cannot be generalized to establish existence for multiproduct firms. Recently, Konovalov and Sandor (2010) prove equilibrium existence for the multiproduct case by employing a different set of conditions. However, researchers have not yet established the existence of an equilibrium for a random coefficient logit demand model. In addition, simulations show that in a random coefficient logit model, the price equilibrium is not unique in general (Konovalov and Sandor 2010). Given these theoretical challenges, an investigation of equilibrium existence and uniqueness in the random coefficient logit model involving multiproduct firms and multimarket competition is far beyond the scope of the current paper.
I combine (8) and (10) to capture this transition. The marginal costs in the last 20% of sales are solved by constrained optimization on Equation (10), constraining the cost to be the same across markets for each product, in order to be consistent with the 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 more than 1,500 markets in which A, B, and/or D operated.

Because a market contains approximately 1,200 observations on average, separate estimation for each market permits the inclusion of 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.\textsuperscript{11}

4.1 Moments

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

\[ s_{jt} = \int \frac{\exp(V_{ijt})}{\sum_{k=1}^{J_t} \exp(V_{ikt})} dP(\beta_i) dP(y_i) \quad (12) \]

\[ \bar{s}_{rt} = \int \sum_{i \in r} \sum_{j=1}^{J_t} s_{ijt} \quad (13) \]

\textsuperscript{11}Note that separate estimation, although more flexible, generally leads to larger standard errors than pooled estimation. In Section 6, I will show the reduced efficiency in parameter estimates does not jeopardize the robustness of the conclusion from the counterfactual experiments.
where (12) is a market share equation with the systematic utility
\[
V_{ijt} = x_{jt}^{rE} \beta_{fc} + x_{jt}^{rC} \beta_i + \alpha_r \log(1 - p_{jt}/y_i) + \rho \log (R_{jt}) + \xi_{jt},
\]
and (13) is implemented as micro moments with \( \tilde{s}_{rt} \) denoting the percentage of households at income tier \( r \) that 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 \( x_{jt}^{rE} \) to form \( x_{ijt}^{rE} \). Then stack observations \( \forall j \) and then \( \forall t \) as rows into matrices and rewrite the systematic utility \( V_{ijt} \) as
\[
V_i = X\theta_1 + X_i^{rE} \theta_2 v_i + \xi,
\]
where \( \theta_1 \) is a vector combining the fixed (non-random) coefficients \( \beta_{fc} \), the means of the random coefficients, \( \bar{\beta} = \text{E}[\beta_i] \), as well as the coefficient of the congestion term \( \rho \). \( \theta_2 \) is the Cholesky root of the covariance matrix of the random coefficients appended with \( \alpha_r \)’s as the last four diagonal elements. \( v_i \) is a vector consisting of random draws from a standard multivariate normal distribution associated with \( \beta_i \), as well as four binary indicators of \( i \)’s income level. The mean utility invariant across households is therefore
\[
\delta = X\theta_1 + \xi.
\]

The demand system is estimated by GMM estimator. This estimation has three sets of moments: the share equations (12), the micro moments (13), and the demand-side orthogonality conditions, which I describe next. Assuming \( \xi \) is mean independent of some set of exogenous instruments \( Z \), the demand-side moments are given by
\[
g(\delta, \theta_1) = \frac{1}{N_d} Z' \xi = \frac{1}{N_d} Z'(\delta - X\theta_1) = 0,
\]
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 a 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 $g(\delta, \theta_1)$ 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 propagating from 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

$$\min_{\phi} \quad F(\phi) = \eta' W \eta$$

s.t. $$s(\delta, \theta_2) = S$$

$$\eta_1 - g(\delta, \theta_1) = 0$$

$$\eta_2 - \tilde{s}(\delta, \theta_2) = -\tilde{S},$$

where $$\phi = \{\theta_1, \theta_2, \delta, \eta_1, \eta_2\}$$ contain optimization parameters. $$W$$ is the weighting matrix in the optimization. $$S$$ is a vector of actual shares. $$\tilde{S}$$ is a vector of the micro-data collected from the PMA consumer survey and scaled by the Mintel statistics and local demographic distributions. $$\eta = (\eta_1', \eta_2')'$$ are 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 weighting these constraints during minimization can be adaptively determined by the data via a two-stage estimation process. Moreover, the sparsity pattern in the original constrained optimization is unchanged after adding these micro moments, as these moments only involve shares that are independent across $$t$$. Therefore, the requirement on computing memory is relatively mild in this MPEC.\(^{12}\)

Denoting the set of constraints as $$G(\phi)$$, the constrained optimization problem (17) results in the following Lagrangian function:

$$L(\phi; \lambda) = F(\phi) - \langle \lambda, G(\phi) \rangle,$$

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

$$\frac{\partial L}{\partial \phi} = 0, \quad G(\phi) = 0 .$$

\(^{12}\)The transition from unconstrained optimization to constrained optimization increases the need for computing memory due to the added constraints. As a result, unconstrained optimization is usually preferred over constrained optimization for dense problems (Nocedal and Wright 1999).
The model estimation proceeds in two stages. In the first stage, identity matrix is used as the weighting matrix $W$ in (17). In the second stage, equal weighting is replaced by the inverse of the second moments $\Phi$, which is a function of the first-stage estimates. The micro moments (over $i$ and $r$) are sampled independently from demand moments (over $j$ and $t$); therefore, $\Phi$ has a block diagonal structure (Petrin 2002). Accordingly, the asymptotic variance matrix for parameter estimates is given as

$$
\Gamma = \frac{1}{N_d + I} (J'WJ)^{-1} J'W\Phi W J (J'WJ)^{-1},
$$

where $J$ is the Jacobian matrix of (16) and (13) with respect to $\theta_1$ and $\theta_2$.

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

5 Results

In this section, I first present the estimates of 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’ profits under local and national pricing policies and under varying competitive market conditions.

5.1 Parameter Estimates

This section discusses parameter estimates, elasticities, and margins across various model specifications. First, I present the parameter estimates from the pooled (across markets) demand model, which makes discussing the implications of each parameter easier. Second, I discuss the results from estimating the demand model separately across the 1,507 markets. Note that the parameter estimates are not directly comparable across markets, because the scale in utility is different (Swait and Louviere 1993). To facilitate comparison, I calculate
Table 5: Parameter Estimates of the Pooled Demand Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS</th>
<th>2SLS</th>
<th>Random Coefficients &amp; Microdata</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price Coefficients (α’s)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α₁</td>
<td>5.014</td>
<td>18.162</td>
<td>32.005</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.027)</td>
<td>(0.806)</td>
</tr>
<tr>
<td>α₂</td>
<td></td>
<td></td>
<td>29.689</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(6.781)</td>
</tr>
<tr>
<td>α₃</td>
<td></td>
<td></td>
<td>63.408</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(8.467)</td>
</tr>
<tr>
<td>α₄</td>
<td></td>
<td></td>
<td>82.339</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(12.296)</td>
</tr>
<tr>
<td>Other Parameters</td>
<td>Mega-pixels</td>
<td>0.049</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>Mega-pixels s.d.</td>
<td>0.119</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.005)</td>
</tr>
<tr>
<td></td>
<td>Optical Zoom</td>
<td>0.004</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>Thickness</td>
<td>-0.159</td>
<td>-0.183</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>Display Size</td>
<td>0.340</td>
<td>0.457</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>Nov-Dec</td>
<td>-0.137</td>
<td>-0.159</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td>June</td>
<td>0.100</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td></td>
<td>Congestion</td>
<td>-0.898</td>
<td>-1.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Note: The data contain $1.78 \times 10^6$ observations. Standard errors are in round brackets. All specifications include year fixed effects and brand-chain interactions.
Table 6: Elasticity Estimates

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

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

Elasticities in both the separate and the pooled estimation.

Table 5 reports 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 shows the price coefficients vary substantially across income tiers. Similar to the findings in Petrin (2002), I find the marginal utility of expenditures on other goods and services increases with income. Consumers on average favor cameras with higher mega-pixels, longer optical zooms, and larger displays, 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 compares them to elasticities from the pooled estimation. Elasticities are only calculated for continuous variables. In each separate model, I include year, brand, and so on.
of the price elasticity and mega-pixel elasticity under either the pooled or the separate estimation. From Table 6 and Figure 4, we see 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 compet-
Table 7: Average Elasticities by Different Market Types

<table>
<thead>
<tr>
<th>Chain</th>
<th>A Monopoly</th>
<th>B Monopoly</th>
<th>A-B Duopoly</th>
<th>A-D Duopoly</th>
<th>B-D Duopoly</th>
<th>A-B-D Triopoly</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-1.761</td>
<td>—</td>
<td>-1.948</td>
<td>-1.824</td>
<td>—</td>
<td>-2.041</td>
</tr>
<tr>
<td>B</td>
<td>—</td>
<td>-1.659</td>
<td>-1.852</td>
<td>—</td>
<td>-1.796</td>
<td>-1.843</td>
</tr>
<tr>
<td>D</td>
<td>—</td>
<td>—</td>
<td>-1.852</td>
<td>—</td>
<td>-1.796</td>
<td>-1.843</td>
</tr>
</tbody>
</table>

itive conditions. Overall, demand is more elastic in markets with more competing stores. When Chains A and B compete in a market, their elasticities increase about 11% compared to when each operates as a local monopolist. Under a market structure of A-D or 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 price elasticity tends to be higher in more competitive markets. In contrast, Hoch et al. (1995) find 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 than for electronics, and that consumers may be more likely to comparison shop for electronics because they are 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 the pooled estimation, respectively. These margins are unrealistic for the digital cameras retail industry, reflecting the underestimated price elasticities.
Table 8: Inferred Price Margin from Demand Estimates and Pricing Equilibrium

<table>
<thead>
<tr>
<th>Margin</th>
<th>Separate Estimation</th>
<th>Pooled Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random Coefficient &amp; Microdata</td>
<td>Random Coefficient &amp; Microdata</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>69.62%</td>
<td>34.53%</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>63.19%</td>
<td>28.59%</td>
</tr>
<tr>
<td><strong>10%-percentile</strong></td>
<td>45.46%</td>
<td>21.24%</td>
</tr>
<tr>
<td><strong>90%-percentile</strong></td>
<td>93.82%</td>
<td>42.89%</td>
</tr>
</tbody>
</table>

Note: Margin is defined as $(p - c)/p$. The pooled estimation results in higher margins than the separate estimation. 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. The correction for biases in demand estimates by incorporating micro moments and congestion helps the demand model better estimate consumer substitution patterns.

5.2 Counterfactual Simulation

5.2.1 National vs. Local Pricing

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 Chain D, the large discount retailer, continues

---

14 According to industry reports, such as Euromonitor (2010), the average margin for point-and-shoot cameras usually ranges from 25% to 35%.
Table 9: Counterfactual Profits \((\pi_A, \pi_B)\) under Alternative Pricing Policies (in $ millions)

<table>
<thead>
<tr>
<th></th>
<th>Chain B</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Local</td>
<td>National</td>
<td></td>
</tr>
<tr>
<td>Chain A</td>
<td>(307.60, 104.06)</td>
<td>(320.58, 105.17)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(310.03, 110.47)</td>
<td>(323.91, 112.78)</td>
<td></td>
</tr>
</tbody>
</table>

to use local pricing. Also, I assume the smallest Chain L, consisting of tiny sellers, is a “dumb” firm that does not respond to any environmental changes. 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, employing national pricing was optimal for both firms A and B. Consistent with Figure 3, which shows 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 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 deviating unilaterally from a national pricing strategy profitable. In a game between A and B in which chains 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.

The counterfactual profits of Chain D under different policy scenarios are the following (in $ millions):

<table>
<thead>
<tr>
<th></th>
<th>A National</th>
<th>A Local</th>
<th>A National</th>
<th>A Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>D Local</td>
<td>47.21</td>
<td>45.29</td>
<td>46.57</td>
<td>44.65</td>
</tr>
<tr>
<td>D National</td>
<td>44.75</td>
<td>40.08</td>
<td>42.84</td>
<td>40.19</td>
</tr>
</tbody>
</table>

Switching to national pricing results in profit loss for Chain D, because (1) in more than half of its markets, D does not coexist with A or B, and (2) D mainly sells low-end products, which only leads to weak competition with A and/or B stores. On the other hand, from a managerial point of view, fixing D’s policy to local 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 D’s general product pricing strategy.
in their own stronghold markets in which they do not compete. Yet the chains lose profits due to the intensified competition in other markets in which they do 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. Overall, 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 national pricing, the prices under the local pricing policy are generally higher in the stronghold 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 one chain chooses national pricing and the other chain chooses local pricing, the latter chain would free ride the former chain. The chain with the national prices would lower its prices relative to the National-National scenario, whereas the chain with the local pricing policy could raise the 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. Whether a chain would unilaterally deviate to local pricing depends on the incremental profit gain in its stronghold market, relative to the profit loss in the contested market as the other chain also
Table 11: Average Price Difference from the National Pricing Policy across Market Types and Scenarios

<table>
<thead>
<tr>
<th></th>
<th>A Local</th>
<th>B Local</th>
<th>A Local</th>
<th>B National</th>
<th>A National</th>
<th>B Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chain</td>
<td>Type I</td>
<td>Type II</td>
<td>Type I</td>
<td>Type II</td>
<td>Type I</td>
<td>Type II</td>
</tr>
<tr>
<td>A</td>
<td>5.81%</td>
<td>-9.77%</td>
<td>5.81%</td>
<td>-6.23%</td>
<td>-2.54%</td>
<td>-2.63%</td>
</tr>
<tr>
<td>B</td>
<td>8.31%</td>
<td>-11.63%</td>
<td>-3.30%</td>
<td>-3.56%</td>
<td>8.31%</td>
<td>-7.17%</td>
</tr>
</tbody>
</table>

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

cuts its uniform price. Under the current market structure, the losses are again excessive for both chains; therefore, neither chain would unilaterally deviate to local pricing, as is evident in Table 9.

5.2.2 Market Structure as a Boundary Condition

Because market structure is an 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 every few stores, I solve the counterfactual profits under national and under local pricing. To mimic the business reality in which struggling chains first close poorly performing stores, I remove stores in the increasing order of profit. The results are plotted in Figure 5. 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 29% of its competitive markets, it 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 has no major competition in all markets, local pricing strictly dominates national pricing, which is consistent with previous findings (e.g., Chintagunta et al. 2003) where
competition is absent or not explicitly modeled.

5.2.3 Competitive Intensity as a Boundary Condition

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 local prices and profits for firm A after firm B exits. I find that under national pricing, A's profits are $176.84 million, and under local pricing, the profits are $174.60 million.\textsuperscript{16} The result implies setting prices uniformly across markets is still optimal for Chain A. The rationale behind this result is that Chain A still faces substantial competition from Chain D. As is evident from Table 3, Chain A coexists with 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, although the advantage of national pricing over local pricing has largely

\textsuperscript{16}Similar to the period before B exits, maintaining local pricing is better for D.
disappeared because of the eased competitive landscape.

5.2.4 Hybrid Pricing Policy

Besides purely local and national pricing, a retail chain may adopt a mixture of the two policies. That is, a chain allows price to be set locally in some markets, and, in the rest of its markets, the chain maintains uniform pricing. In this way, the chain can exploit many possible geographic combinations. Here I simulate the outcome of two candidate policies and discuss the implication of the results.

The first candidate policy is motivated by the business reality that firms sometimes play different strategies in large and influential markets as opposed to other markets. In terms of geographic pricing policy, a similar scenario occurs when a national chain customizes prices in populous markets and keeps prices uniform elsewhere. To examine this possibility, I select the top five metropolitan areas in the United States to be the local pricing zone of Chains A and B. According to the recent census, the five most populous U.S. cities are New York City, Los Angeles, Chicago, Houston, and Philadelphia. These five cities together account for about 6% of the total U.S. population. The retail sales of digital cameras in these cities constitute 8.6% of the national demand. Using the data prior to B’s exit, I simulate the outcome as if both A and B adopted local pricing in these cities.

Table 12 presents the relative changes in profit and price if Chains A and B replace the observed policy with the proposed hybrid pricing scheme. Consistent with the mechanism discussed earlier, switching to local pricing intensifies price competition between the two

<table>
<thead>
<tr>
<th>Chain</th>
<th>% changes in local pricing zone profit</th>
<th>% changes in uniform pricing zone profit</th>
<th>overall profit change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-12.29%</td>
<td>0.57%</td>
<td>-0.76%</td>
</tr>
<tr>
<td>B</td>
<td>-15.33%</td>
<td>0.72%</td>
<td>-1.23%</td>
</tr>
</tbody>
</table>
rivals in the five largest cities. Both firms would lower prices in response to the policy change. In 2007-2008, the counterfactual local prices led to a relative profit loss of 12.29% for Chain A and 15.33% for Chain B in the local pricing zone. On the other hand, excluding the five biggest competitive markets slightly improves the profitabilities of both firms in the uniform pricing zone, thanks to the reduced “downward” force on the uniform prices. Aggregating across the two pricing zones, however, the proposed hybrid policy results in overall profit declines for both firms A and B, because of the excessive loss in the areas in which local pricing is applied.

Although the above proposal does not increase profits, many other alternative policies remain. Instead of localizing prices in large competitive markets, a chain could localize them in its stronghold markets where it faces little competition from major rivals, thereby leading to profit gains in these markets. Hence, I simulate the outcome as if A and B set price locally in some of their own stronghold markets while maintaining uniform pricing elsewhere. I start by ranking the SSAs in which A and B do not compete with each other, according to the total sales volume in 2007-2008 of local A and B stores, respectively. Then I let the top 10% of these markets be local pricing zones for A and for B. The relative profit and price changes are reported in Table 13. As a comparison, Table 14 presents the outcome as if local pricing is applied to the top 20% stronghold markets.

Table 13: Relative Profit and Price Changes if Chains A and B Adopt Local Pricing in Top 10% of Stronghold Markets

<table>
<thead>
<tr>
<th>Chain</th>
<th>% changes in local pricing zone</th>
<th>% changes in uniform pricing zone</th>
<th>overall profit change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>profit</td>
<td>price</td>
<td>profit</td>
</tr>
<tr>
<td>A</td>
<td>7.12%</td>
<td>6.61%</td>
<td>-0.23%</td>
</tr>
<tr>
<td>B</td>
<td>9.75%</td>
<td>9.52%</td>
<td>-0.99%</td>
</tr>
</tbody>
</table>

After switching to local pricing in its top 10% stronghold markets, firm A generates 7.12% more profit in these SSAs relative to the observed policy. In the rest of A’s markets in which uniform pricing is maintained, prices drop because of the reduced market power
A could leverage from the excluded stronghold markets. The decrease in price intensifies competition slightly and leads to a profit decline that can be offset by the gain in the local pricing zone. Overall, chain profitability improves for Chain A, although the improvement is rather small (0.06%). On the other hand, Chain B obtains incremental profit in its local pricing zone, similar to Chain A. However, the profit loss in B’s uniform pricing zone is large, and the chain profitability deteriorates slightly under the proposed hybrid policy. The different changes in overall profitabilities are due to the distinct distributions of market structures between the two chains. According to Table 3, Chain A has many more stronghold markets than B does, and the total sales in A’s stronghold markets also surpass those in B by a large margin. As a result, when 10% of the stronghold markets are excluded, prices in B’s uniform pricing zone would decrease by a larger amount than prices in A’s, thanks to the relatively heavier “downward” force on B’s uniform prices. In Table 14, the changes in chain profits remain in the same direction, but the magnitude of the changes increases.

Once we allow for hybrid pricing, the search for an equilibrium policy becomes much more complicated. I am now pursuing the equilibrium existence in both analytical models and empirical models using the current data. Of course, even if such a policy exists theoretically, whether it works in a specific industry depends on the industry’s multimarket structure and is therefore subject to empirical investigation. More importantly, the current model does not account for organizational costs associated with local pricing, nor does it consider the negative consumer perception for inconsistent online and offline prices. Therefore, whether hybrid pricing is a managerially and institutionally viable policy remains an open question for future research.

6 Robustness Check

6.1 Market Definition

Properly defined local markets are important to the current analysis of multimarket competition. The results obtained thus far are dependent on how neighboring stores are
Table 14: Relative Profit and Price Changes if Chains A and B Adopt Local Pricing in Top 20% of Stronghold Markets

<table>
<thead>
<tr>
<th>Chain</th>
<th>% changes in local pricing zone profit</th>
<th>% changes in local pricing zone price</th>
<th>% changes in uniform pricing zone profit</th>
<th>% changes in uniform pricing zone price</th>
<th>overall profit change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5.86%</td>
<td>5.93%</td>
<td>-0.40%</td>
<td>-0.87%</td>
<td>0.11%</td>
</tr>
<tr>
<td>B</td>
<td>7.94%</td>
<td>8.62%</td>
<td>-1.35%</td>
<td>-2.06%</td>
<td>-0.83%</td>
</tr>
</tbody>
</table>

grouped into the same market or different markets. The key requirement in delineating markets is that stores within the same market should serve as substitutes for each other, and stores in different (even adjacent) markets should have little competition. In other words, no spillover should be present between markets, and consumers in one market would not shop in other markets.

The data provider has already defined distinct competitive local markets as SSAs. According to this definition, ninety-five percent of SSAs contain only one store from each major retailer, meaning almost no internal competition exists between stores of the same chain. I use store-address data from AggData to construct distance measures to evaluate the SSA definition. First, the median distance between competing stores within an SSA is 0.58 miles, whereas the median and the bottom 5th-percentile distance to competing stores in neighboring SSAs are 10.20 and 3.45 miles, respectively. These model-free statistics suggest retail stores are located near each other within a market and relatively farther from stores outside their SSAs. However, distance measures alone are not sufficient support for SSAs being independent, because consumers may still travel a certain longer distance to buy cameras. A test that directly examines cross-store substitution should be a better way to verify the SSA definition.

To this end, I employ a “structural” test – the hypothetical monopolist (HM) test that antitrust literature first introduced to investigate market definition in the context of horizontal mergers (Katz and Shapiro 2003). Davis (2006) applies this test to delineate geographic markets in the study of spatial competition among movie theaters. To implement the HM
test, I start with narrowly defined markets with only one store. Each of these markets is therefore a monopolist market. If a store is unable to raise its price profitably, the market definition is considered sufficiently broad, because the local market power is already being exploited and consumer demand is already elastic. If the store generates incremental profit by raising its price, however, the observed price before the price increase is below the optimal level a monopolist store would charge, which must be due to competition from outside the monopolist market. In this case, the monopolist market needs to expand to include the currently best available substitute (store) to form a broader two-store market. In the newly defined market, I again raise price to assess profit change. If profit declines, the current market definition is deemed sufficiently broad and the test algorithm stops. Otherwise, the algorithm continues to include one more store and repeats the procedure until the price increase does not improve profitability.

Store profit in a local market is $\pi^0 = M \sum_{\forall j} \tau p^0_j s^0_j$, where $\tau$ denotes the average margin rate. After imposing a price increase $\kappa$, profit change will be $\pi^1 - \pi^0 = M \sum_{\forall j} ((\tau + \kappa)p^1_j s^1_j - \tau p^0_j s^0_j)$. Following Davis (2006), I calibrate $\tau = 0.3$ using data from industry reports such as Euromonitor (2010). I cannot use the margins estimated in Table 8 here, because they are based on the predetermined SSA definition, whereas the goal of the HM test is to find the right market definition. According to antitrust literature, the price increase for this test should be “small but significant and nontransitory” (usually around 5%-10%) and endure for a year (Katz and Shapiro 2003). Therefore, I set $\kappa = 5\%$ and let it persist for 12 months.

In the current analysis, two factors further complicate the implementation of the test. The first complication is that the two major chains, A and B, primarily adopted national pricing. The observed national prices in A and B stores are not intended to be locally optimal and so cannot be used in conducting the local market definition test. As a result, I have to rely on the observations from D stores, which used local pricing. Specifically, I start the narrowly defined monopolist markets with D stores and subsequently include A and B stores, if the hypothetical markets have to expand. In this way, however, the SSAs without Chain D are excluded from the test. According to Table 3, SSAs without D stores
only accounts for 14.9% of all SSAs before B exits and 10.3% after B exits. Therefore, the test already includes the majority of SSAs, and the excluded SSAs should not change the qualitative conclusion of the test.

The second complicating factor is the presence of small stores other than A, B, or D. In the preceding analysis, I grouped all small stores into a single chain, L, for simplicity. When delineating local markets, small stores located at various parts of a market blur the competition boundary of major stores, whereas their existence is unlikely to impact the substitution pattern between major stores. Therefore, I focus the test on the three major chain stores and combine small stores as the outside option.

Table 15 presents the test results. Initially, every D store is treated as a single monopolist market, and demand is estimated in each market over time. After a price increase on all products in a market, I compute profit change using the exogenously calibrated margin rate as well as the newly obtained demand estimates. A rise in profit after the price increase indicates the monopolist is not immune to competition outside the market; therefore, we must reject the hypothesis of a one-store market. For those D stores in SSAs without A or B, only 18.7% of them have the monopolist hypothesis rejected. That is, the majority (81.3%) of this type of SSAs are generally in line with the monopolist-like market structure. On the other hand, most D stores in SSAs with A and/or B (94.5%, 98.2%, and 91.9%) generate incremental profit after the price increase, suggesting these SSAs cannot be contracted into single-store markets.
Next, I add a competitor to the D stores in SSAs with A and/or B to form two-store hypothetical markets. Among the two-store markets in which the corresponding SSA definition is also two-store (D, A or D, B), only 13.4% and 14.7% of them have profitable price increases on D’s products. Therefore, most of the D, A-type and D, B-type SSAs are well defined. In SSAs in which A, B, and D are all present, the two-store hypothesis is rejected at a rate of 89.1%. Only after including the third store does the rejection rate (of a three-store hypothesis) drop to 7.6% for the D, A, B-type SSAs.

So far the test results support the assertion that most SSAs are properly defined during 2007, prior to B’s exit. To evaluate the SSA definition after B exits, I conduct the same test again, using the 2009 data. The results are shown in the last two columns of Table 15. Much like the situation in 2007, the one- and the two-store hypotheses are largely consistent with the one- and the two-store SSA definitions, respectively. The one-store hypothesis is rejected in 93.8% of the SSAs in which both D and A exist. Based on the results from the two separate periods, I conclude the SSA definition serves as a good approximation of distinct local competitive markets.

6.2 Robustness of Counterfactual Outcomes

The calculation of counterfactual outcomes in Section 5 uses the estimates of demand and supply systems as input. Uncertainty around these estimates may affect the counterfactual results and consequently the general conclusion of this paper. I obtain the demand estimates from a statistical model with standard errors as part of the estimation output. The cost estimates and counterfactual results, on the other hand, are not from statistical models, but from deterministic computing process – finding roots of the optimal pricing equations (8) and (10). Therefore, the standard errors from the demand side propagate into marginal costs and subsequently into the counterfactual calculation. Now the question concerns the amount of impact the uncertainty of demand estimates has on the final results. To assess the impact, I take \( I = 100 \) random draws from the asymptotic (normal) distribution of the demand estimates, and with each draw, I recompute marginal costs and the counterfactual
outcomes. Ninety-two percent of these draws lead to the same Nash equilibrium as in Table 9. This result suggests the counterfactual outcomes are rather robust given the current data sets.

7 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 more than 1,500 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 related to the implementation of national or local pricing. For example, by switching from national to local, a chain may incur additional costs in customizing advertising to match locally varying prices. Also, consumers may dislike inconsistent prices offline and online, and across different stores. Therefore, moving to local pricing could incur certain psychological costs for which the current model does not account.

Second, several recent papers have documented that durable goods buyers may strategi-
cally 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 critical. Doing so in the context of a dynamic structural demand model is generally intractable in computation, especially because the model involves hundreds of local markets and hundreds of distinct products. On the other hand, the focus of the current study is geographic pricing policy and the differences between markets primarily drive the conclusion. The effect of forward-looking dynamics, if relatively similar across markets, would be canceled out when examining cross-market variations and would therefore not influence the main result qualitatively. Lastly, 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


45

Appendix

A An Analytical Model of Multimarket Competition

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

I derive the duopoly demand function 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 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:

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

(21)

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

$$\begin{pmatrix}
1 & \theta \\
1 + \theta & 1 + \theta \\
\theta & 1 \\
1 + \theta & 1 + \theta
\end{pmatrix},$$

(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 $\theta$. As $\theta \to 1$, the two goods approach perfect substitutes. In Desai et al. (2010) and Subramanian et al. (2010), the coefficient $\beta$ on expenditure is set to one because these studies primarily focus on the difference between the two competing goods, for which $\beta$ is a common multiplier. The current analysis, however, examines differences not only within a market but also across markets (with different $\beta$’s), so I keep this parameter 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:

\[
q_a = \alpha_a - \beta p_a + \frac{\beta \theta}{1 - \theta} (p_b - p_a) \\
q_b = \alpha_b - \beta p_b + \frac{\beta \theta}{1 - \theta} (p_a - p_b).
\] (23)

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

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

Accordingly, the monopoly demand is

\[
q = \alpha - \beta p .
\] (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 in which $a$ and $b$ compete. Assuming both chains are single-product firms, demand in the three markets follows (23) and (25).

Under local pricing, a chain makes price decisions independently across markets. For instance, chain $a$ solves two unrelated pricing problems given chain $b$'s price in market 3:
Table 16: Payoffs of the Two-stage Game

<table>
<thead>
<tr>
<th>Chain a</th>
<th>National</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\pi_a^N, \pi_b^N$</td>
<td>$\pi_a^L, \pi_b^L$</td>
</tr>
<tr>
<td></td>
<td>$\pi_a^N, \pi_b^N$</td>
<td>$\pi_a^L, \pi_b^L$</td>
</tr>
</tbody>
</table>

Max $\pi_{a1}(p_{a1})$ and 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

$$\text{Max}_{p_{a}} \pi_{a1}(p_{a1}) + \pi_{a3}(p_{a3}|p_{b3})$$

s.t. $p_{a1} = p_{a3} = p_{a}$

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 to the policy they chose in the first stage. Table 16 summarizes all possible payoffs of the game, where individual payoffs are given by

$$\pi_{1N} = \frac{(\theta - 1)[\beta_{a1}(\theta - 1) - 1]2(\alpha_{a1} + \alpha_{a3})(1 + \beta_{b2}) + \theta(\alpha_{a2} + \alpha_{b3} - 2\beta_{b2}(\alpha_{a1} + \alpha_{a3}))^2}{[4(1 + \beta_{a1})(1 + \beta_{b2}) - 4\theta(\beta_{a1} + \beta_{b2} + \beta_{a1}\beta_{b2}) + \theta^2(4\beta_{a1}\beta_{b2} - 1)]^2}$$

$$\pi_{2N} = \frac{(\theta - 1)[\beta_{b2}(\theta - 1) - 1]2(\alpha_{b2} + \alpha_{b3})(1 + \beta_{a1}) + \theta(\alpha_{a1} + \alpha_{a3} - 2\beta_{a1}(\alpha_{b2} + \alpha_{b3}))^2}{[4(1 + \beta_{a1})(1 + \beta_{b2}) - 4\theta(\beta_{a1} + \beta_{b2} + \beta_{a1}\beta_{b2}) + \theta^2(4\beta_{a1}\beta_{b2} - 1)]^2}$$

$$\pi_{1L}' = \frac{\alpha^2_{a1}}{4\beta_{a1}} - \frac{(\theta - 1)[\theta(\alpha_{a2} + \alpha_{a3}) + 2\alpha_{a3}(1 + \beta_{b2} - \beta_{a2}\theta)]^2}{[\theta^2 + 4\beta_{b2}(\theta - 1) - 4]^2}$$

$$\pi_{2N}' = \frac{(\theta - 1)[\beta_{b2}(\theta - 1) - 1][2(\alpha_{b2} + \alpha_{a3}) + \alpha_{a3}\theta]^2}{[\theta^2 + 4\beta_{b2}(\theta - 1) - 4]^2}$$

\footnote{For the purpose of cross-market comparison, I set all variable costs to zero, and normalize the $\beta$ in the duopoly market to one.}
\[ \pi_{1N}' = \frac{(\theta - 1)[\beta a_1(\theta - 1) - 1][2(\alpha a_1 + \alpha a_3) + \alpha a_3 \theta]^2}{(\theta^2 + 4\beta a_1(\theta - 1) - 4)^2} \]

\[ \pi_{2L}' = \frac{\alpha b_2^2}{4\beta b_2} - \frac{(\theta - 1)[\theta(\alpha a_1 + \alpha a_3) + 2\alpha b_3(1 + \beta a_1 - \beta b_1 \theta)]^2}{(\theta^2 + 4\beta a_1(\theta - 1) - 4)^2} \]

\[ \pi_{1L} = \frac{\alpha a_1^2}{4\beta a_1} - \frac{(\theta - 1)(2\alpha a_3 + \alpha b_3 \theta)^2}{(\theta^2 - 4)^2} \]

\[ \pi_{2L} = \frac{\alpha b_2^2}{4\beta b_2} - \frac{(\theta - 1)(2\alpha a_3 + \alpha a_3 \theta)^2}{(\theta^2 - 4)^2} \].

(26)

Because these profit functions contain seven parameters, drawing a closed-form conclusion regarding the conditions under which one policy is better than the other is impossible. Therefore, in the remainder of this section, I numerically analyze the analytical results.

To show the profit-enhancing effect of national pricing and how such an effect changes with market structure, I first examine the profit change when a chain switches from national to local pricing, given the other chain does the same. Figure 6 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 ranges of \( \alpha a_3 \) and \( \alpha b_3 \) 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 can neither be too large nor too small compared to its monopoly market. When the chain is very small in the contested market, the local profit gain through national pricing cannot cover the profit loss in its monopoly market. When the chain is very small in the contested market, the local profit gain through national pricing cannot cover the profit loss in its monopoly market. On the other hand, when the chain is very large in the competitive market, the demand in its monopoly market is not sufficient to drive up the duopoly price, thereby barely softening competition and generating incremental profit. Further, if chain \( b \) is large in the duopoly market, chain \( a \) would have difficulty raising the price 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 also not too large.
Figure 6: Contour Plot on $\Delta \pi = \pi_{aN} - \pi_{aL}$ against Varying Market Structure

$(\alpha_{a1}=2, \beta_{a1}=1, \alpha_{b2}=2, \beta_{b2}=1$ and $\theta=0.5)$

Next, I examine the conditions under which national pricing is an equilibrium of this game. When $\Delta \pi_a = \pi_{aN} - \pi'_a L > 0$ and $\Delta \pi_b = \pi_{bN} - \pi'_b L > 0$ both hold, national pricing is the dominant strategy for both chains. The colored region II in Figure 6 describes the ranges under which the equilibrium exists. This range is smaller than the previous case because a free-rider issue is present. Suppose a chain moves to local pricing while the other chain sticks to national, the first chain will reap the maximal profit from its own monopoly market, while benefiting from the other chain’s national pricing in the contested market, which softens competition. However, the duopoly competition is still intensified in this market as the second chain would lower its national price; therefore, the first chain may still
experience excessive loss in the competitive market. The possibility of unprofitable unilateral deviation leads to a Nash equilibrium in this multimarket duopoly game.

The analytical model highlights the rationale on how geographic pricing policy affects chain profitability. However, because this hypothetical model is simplified, it does not reflect the complexity of multimarket chain competition in the real world. For example, real industries usually contain many local markets and multiple competing chains, so the conditions supporting national pricing are not just about three markets and two firms, but about the distribution of market structures and the distribution of competitive intensities across these markets and firms. Also, chain stores sell multiple differentiated products, and these products are substitutes for each other as well. Moreover, consumer demand is usually not in the linear fashion. For these reasons, in this paper, I rely on real data sets and investigate the choices of pricing policy through empirical models.

### B Analytic Derivatives for the MPEC Estimation

In this section, I derive the analytic derivatives for the optimization problem specified in (17). My derivation follows matrix calculus and employs tensor operators such as Kronecker product. Thanks to the sparsity pattern of this 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, thereby substantially saving computational time. All derivatives are formulated compactly in matrix notation to assist coding in computer programs.

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

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

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

The Jacobian matrices of the constraints imposed by the share equations are
\[
\frac{\partial s_t(\delta_t, \theta_2)}{\partial \theta_2} = \int_{\forall i} \text{diag}(s_{it}) [X_{it}^{rc} - 1_J s'_{it} X_{it}^{rc}] \text{diag}(v_i) \] (29)

\[
\frac{\partial s_t(\delta_t, \theta_2)}{\partial \delta_t} = \int_{\forall i} \text{diag}(s_{it}) - s_{it} s'_{it}, \] (30)

where \(1_J\) is a \(J_t\)-element column vector of ones.

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

\[
\frac{\partial [\eta_1 - g(\delta, \theta_1)]}{\partial \theta_1} = \frac{1}{N_d} Z' X \] (31)

\[
\frac{\partial [\eta_1 - g(\delta, \theta_1)]}{\partial \delta} = -\frac{1}{N_d} Z' \] (32)

\[
\frac{\partial [\eta_1 - g(\delta, \theta_1)]}{\partial \eta_1} = I_{n_\eta}. \] (33)

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

\[
\frac{\partial [\eta_2 - \tilde{s}_{it}(\delta_t, \theta_2)]}{\partial \theta_2} = -\int_{i \in r} s_{it} s'_{it} X_{it}^{rc} \text{diag}(v_i) \] (34)

\[
\frac{\partial [\eta_2 - \tilde{s}_{it}(\delta_t, \theta_2)]}{\partial \delta_t} = -\int_{i \in r} s_{it} s'_{it}. \] (35)

The Hessian vector\(^{18}\) of all the constraints in the \(\theta_2\) by \(\theta_2\) block is

\(^{18}\)The following linear transformation is particularly useful in deriving the Hessian from the Jacobian due to the necessity of taking derivatives over the diagonal matrix of share vectors. For example, an \(n\)-by-\(n\) diagonal matrix \(\text{diag}(s)\) with a vector \(s\) on its diagonal can be transformed linearly by

\[
\text{diag}(s) = \sum_{i=1}^{n} E_i s_i e'_i, \]

where \(E_i\) is an \(n\)-by-\(n\) matrix of all zeros except the \(i\)-th diagonal entry equal to one, and \(e_i\) is a vector of
\[ \sum_{j,t} \lambda_{jt} \frac{\partial^2 s_{jt} (\delta_t, \theta_2)}{\partial \theta_2 \partial \delta_t} = \sum_{t=1}^{T} \int_{\mathcal{V}_i} \text{diag}(v_i) \left[ (X_{it} r' - X_{it} r' s_{it} I_{J_t}) \text{diag}(\lambda_i) - \lambda_i s_{it} X_{it} r' \right] \frac{\partial s_{it}}{\partial \theta_2} \quad (36) \]

\[ \sum_{r,t} \lambda_{rt} \frac{\partial^2 [\eta_2 - \tilde{s}_{rt}]}{\partial \theta_2 \partial \delta_t} = \sum_{r,t} \lambda_{rt} \int_{i \in r} s_{it} \text{diag}(v_i) X_{it} r' s_{it} X_{it} r' \text{diag}(v_i) \frac{\partial s_{it}}{\partial \theta_2} \quad (37) \]

where \( \frac{\partial s_{it}}{\partial \theta_2} \) is calculated as in (29) without the integral. \( \lambda_t \) is a vector of the Lagrange multipliers corresponding to the share equations at \( t \).

The Hessian vector of all the constraints in the \( \delta_t \) by \( \theta_2 \) block is

\[ \sum_{j,t} \lambda_{jt} \frac{\partial^2 s_{jt} (\delta_t, \theta_2)}{\partial \delta_t \partial \theta_2} = \sum_{t=1}^{T} \int_{\mathcal{V}_i} \text{diag}(v_i) \left[ (X_{it} r' - X_{it} r' s_{it} I_{J_t}) \text{diag}(\lambda_i) - \lambda_i s_{it} X_{it} r' \right] \frac{\partial s_{it}}{\partial \theta_2} \quad (38) \]

\[ \sum_{r,t} \lambda_{rt} \frac{\partial^2 [\eta_2 - \tilde{s}_{rt}]}{\partial \delta_t \partial \theta_2} = \sum_{r,t} \lambda_{rt} \int_{i \in r} s_{it} \text{diag}(v_i) X_{it} r' s_{it} X_{it} r' \text{diag}(v_i) \frac{\partial s_{it}}{\partial \theta_2} \quad (39) \]

The Hessian vector of all the constraints in the \( \delta_t \) by \( \delta_t \) block is

\[ \sum_{j,t} \lambda_{jt} \frac{\partial^2 s_{jt} (\delta_t, \theta_2)}{\partial \delta_t \partial \delta_t} = \sum_{t=1}^{T} \int_{\mathcal{V}_i} \text{diag}(v_i) \left[ (X_{it} r' - X_{it} r' s_{it} I_{J_t}) \text{diag}(\lambda_i) - \lambda_i s_{it} X_{it} r' \right] \frac{\partial s_{it}}{\partial \delta_t} \quad (40) \]

\[ \sum_{r,t} \lambda_{rt} \frac{\partial^2 [\eta_2 - \tilde{s}_{rt}]}{\partial \delta_t \partial \delta_t} = \sum_{r,t} \lambda_{rt} \int_{i \in r} s_{it} \text{diag}(v_i) X_{it} r' s_{it} X_{it} r' \text{diag}(v_i) \frac{\partial s_{it}}{\partial \delta_t} \quad (41) \]

all zeros except the \( i \)-th element equal to one. Because the transformation is linear, the derivative of the diagonal matrix with respect to \( s \) can be compactly written as

\[ \frac{\partial \text{diag}(s)}{\partial \delta} = \sum_{i=1}^{n} (e_i \otimes E_i) \frac{\partial s}{\partial \delta}, \]

where \( \otimes \) denotes Kronecker product.
where $\frac{\partial s_{it}}{\partial \delta_t}$ is calculated as in (30) without the integral.

After 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 $\theta_1$ and $\theta_2$ is

$$
J = \begin{pmatrix}
\frac{\partial g}{\partial \theta_1} & \frac{\partial g}{\partial \theta_2} \\
\frac{\partial \tilde{s}_r t}{\partial \theta_1} & \frac{\partial \tilde{s}_r t}{\partial \theta_2}
\end{pmatrix},
$$

(42)

where

$$
\frac{\partial g}{\partial \theta_1} = -\frac{1}{N_d} Z' X
$$

(43)

$$
\frac{\partial g}{\partial \theta_2} = \frac{1}{N_d} Z'(\frac{\partial s_t}{\partial \delta_t})^{-1} \frac{\partial s_t}{\partial \theta_2}
$$

(44)

$$
\frac{\partial \tilde{s}_r t}{\partial \theta_1} = \left( \int_{i \in r} s_{i0t} s_{it}' \right) X_t
$$

(45)

$$
\frac{\partial \tilde{s}_r t}{\partial \theta_2} = \int_{i \in r} s_{i0t} s_{it}' X_{it}^{rc} \text{diag}(v_i)
$$

(46)

The second moments $\Phi$ is

$$
\begin{pmatrix}
\Phi_1 & 0 \\
0 & \Phi_1
\end{pmatrix},
$$

(47)

where

$$
\Phi_1 = \frac{1}{N_d} \sum_{j,t} \xi_j^2 Z_{jt} Z_{jt}'
$$

(48)
\[ \Phi_2 = \frac{1}{I} \text{diag} \left( \sum_i (\bar{s}_i - \bar{S})^2 \right). \]  

(49)

C Scaling for the Micro Moments

Because the PMA survey information is available at the national level, scaling is needed to match the survey statistics to the geographic variation in demographics and market sizes. 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 local market. First, from the market-specific income distribution \( P(y_i), \) I obtain the weight of each segments in this market by

\[ w_r = \int_{i \in r} dP(y_i), \]

where \( r = 1, 2, 3, 4. \) Denoting \( \bar{S}_t = \sum_{j=1}^{H} s_{jt} \) as the sum of shares of all inside options observed in the sales data, I can solve the following equations to obtain \( a, b, c, \) and \( d: \)

\[ \bar{S}_t = w_1 a + w_2 b + w_3 c + w_4 d \]

\[ a/b = A/B \]
\[ b/c = B/C \]
\[ c/d = C/D. \]

D Data Trimming

The raw NPD data on store sales include nearly 10 million observations. I trim the data before applying them to estimate the econometric model. First, I remove SSAs in which none of the three major chains had a presence. Then I delete cameras that are not compact
point-and-shoot (e.g., digital SLRs, which account for less than 10% of the industry sales). Third, I retain only sales corresponding to the top seven brands. Fourth, I get rid of all 2010 observations, due to the right truncation issue in calculating cumulative sales. 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 in this industry.