Inferring Competitor Pricing with Incomplete Information

Marcel Goic
University of Chile

Alan Montgomery
Carnegie Mellon University, alm3@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/tepper

Part of the Marketing Commons
Inferring Competitor Pricing with Incomplete Information

Marcel Goić and Alan L. Montgomery

Working Paper

January 2011

Marcel Goić (email: mgoic@dii.uchile.cl) is an Assistant Professor at the Department of Industrial Engineering, University of Chile. Alan L. Montgomery (e-mail: alan.montgomery@cmu.edu) is an Associate Professor at the Tepper School of Business, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213. This research was partially completed while Marcel Goić was a doctoral candidate at Carnegie Mellon University’s Tepper School of Business.

Copyright © 2011 by Marcel Goić and Alan L. Montgomery All rights reserved
Inferring Competitor Pricing with Incomplete Information

Abstract:

We study how business customers make multi-product purchase decisions and how the distributors who sell those products can make inferences about their demand functions with incomplete information. The problem is that distributors rarely observe a competitor's price directly, and must infer competitor response indirectly from their own observations about customer purchases. In this research we propose that customers make their product orders by minimizing procurement costs. Using the first order conditions from this optimization problem we characterize the regions of the parameter space where consumers buy from each distributor. We use these conditions to estimate a model of purchase behavior that enables us to identify the likelihood of each consumer buying from the competitor versus a direct change in consumption patterns.

Our proposed model is applied to a wholesale food distributor and we find widespread heterogeneity in purchase patterns. The empirical results shed light on the competitive elements of customer demand that cannot be studied with traditional reduced form response models. For example, we found that some customers satisfy most of their requirement from one of their distributors, while others consistently split their demands across suppliers. Also, we found that price sensitivity of customers making most of their purchases with the focal supplier are less affected by the volume of purchases in previous periods. We expect this result to provide valuable information for vendors to negotiate prices with the customers. It allows the distributor to make efficient inferences about competitors in those occasions when competitor price is not directly observed.

Keywords: Competitive Pricing, Incomplete Information, Bayesian Methodology.
1. Introduction

A fundamental precept in marketing is to understand competition. Unfortunately in many situations it is difficult to directly observe competitors’ quantity sales and price. Instead the manager has to infer competitive effects by looking at the behavior of their customers. For example, suppose a manager holds their marketing mix constant. If a customer purchases less than typical, the manager may be able to infer that the current drop in sales comes from a competitor’s unobserved price decrease. Alternatively, an increase in purchases may be attributed to a competitor’s unobserved price increase.

The challenge is that there are many competing reasons for variation in purchases that may have nothing to do with a competitor’s actions, such as a drop in the underlying demand for the product, seasonal effects that depress purchases, or simply unattributable, random fluctuations in demand. Moreover, if the customer has shifted purchases from one supplier to another this change may be due to a more attractive price or increased loyalty for the other supplier. From a managerial perspective, it is important to disentangle these underlying reasons since it could suggest quite different strategies for dealing with a customer. For example, a price promotion to spur demand may be profitable if a demand drops due to lower competitive prices, but it would not be desirable if demand dropped due to random fluctuations.

The goal of this research is to model demand with incomplete information about competitors’ sales and prices. We propose a structural model of a customer’s buying patterns using a firm’s internal data about their customers and augment it with customer characteristics to predict unobserved competitive behavior. More specifically we develop a new econometric model to describe customer purchase decisions based upon costs associated with purchasing. This contrasts with a reduced form model that would treat the competitor’s price as a random unobserved variable.
that is absorbed into the error function. Reduced form models can provide good forecasts, but they are silent about key competitive questions relevant to managers that our structural model can answer: How competitive are our prices? Are our prices competitive across all customer profiles? And if not, for which customer segments and product categories are our prices more attractive?

Moon, Kamakura and Ledolter (2007) consider inferences about unobserved competitor promotions in the pharmaceutical industry. Their problem is similar to ours in that we wish to make inferences about competitive behavior, but limited to our own sales information. Additionally, Du, Kamakura and Mela (2007) consider a customer relationship manager problem for a bank that has customer transaction data, but lacks knowledge about their customer’s activities at other banks. However, they do have information about potential sales from other sources. This allows them to make predictions about the size and share of the customer. Our goal is to further study in this nascent area by considering inferences about competitive sales and prices when buyers are driven by cost minimization.

Our problem is drawn from a wholesale food distributor which sells to independent restaurants. This industry is characterized by a mature, intense competition and a high level of customer service. Distributors tend to employ their own salesforces, which typically contact customers every week, collect orders, arrange deliveries, and provide price and promotion information. The salesforce tends to have a hybrid compensation scheme, in which some income comes from a salary and the remainder from commissions. Most competitors have a similar structure, although some have only salaried or hourly employees. Additionally some competitors only employ fixed-posted prices, although most allow for negotiated prices. According to First Research, this industry is comprised of 33,000 distributors, but it is also highly concentrated with the fifty largest companies generating about half of the annual combined industry revenue of $670
billion. The largest distributors are McLane Company, Supervalu, Sysco Corp, and US Foodservice (First Research 2010). Although wholesale food distribution is an important industry for the U.S. economy, we are unaware of any academic studies of it.

The wholesale food distributor does not directly observe competitors’ prices or their sales, which is the focus of our research. Surprisingly, there is a high level of variability of price across customers for a particular distributor. This is largely due to the autonomy given the salesforce in negotiating prices and the salesforce compensation structure. Potentially it is also due to the large seasonal movements of certain foodstocks, the larger number of products purchased, and asymmetric competition between the distributor and restaurant. The size of customer facilities (space constraints) and the nature of the business (perishability of many items) imply a limited customer capability of strategic stockpiling. Ultimately, restaurants do not consume the products they purchase but instead derive their demand from their patrons’ demand. This places our problem within the context of a business-to-business market that is quite different than consumer retailing.

Our problem of not observing competitive information is atypical in marketing research. Consider the plethora of research on data rich environments like supermarket retailing. Syndicated data providers like Nielsen and IRI provide detailed sales, price, and promotional information about competitors at a product by weekly level for each retailer. This allows manufacturers such as P&G and Unilever or retailers like Kroger, Safeway, or SuperValue to precisely measure price and promotional competition. For example, Blattberg and Wisniewski (1989) and Allenby (1989) all describe how to directly measure consumer’s response to price and promotional changes through sales response models. Potentially they may provide data about every store within a retailer to infer store-level demand (Montgomery 1997). In contrast to these situations with dense competitive
information, research on inferring competition from internal sales data is brief (Moon et al 2007, Du et al 2007).

2. Research on Demand under Unobserved Competition

The goal for this research is to develop and estimate a model of customer demand with incomplete information using internal sales transaction data. We can characterize our competitors as suppliers and customers as industrial buyers. Given the nature of our customers it is useful to contextualize our problem with respect to the literature on industrial buyer behavior and dual sourcing. Additionally, our problem shares elements with models proposed in consumer buying which we all discuss.

2.1. Industrial Buyer Behavior

As has been pointed out by early work on industrial buyers, there are important differences with final consumer buying (cf., Webster 1978, Sheth 1996). Given the large variations we observe in industrial buyer behavior, it is useful to define the domain of our study. According to the classic Buyclass framework, we restrict our attention to modified rebuy situations (Anderson, Chu and Weitz 1987). In our empirical setting, small adjustments in quantities required are needed for each purchase occasion to satisfy weekly demand and balance short term inventories. Also, variations in the mix of products are introduced a few times in a year to accommodate seasonal variations in the final consumer demand.

While dealing with industrial buyers most of the literature assumes that decision makers decide their purchases to minimize cost instead of the dual utility maximization problem typically assumed in the analysis of final consumer decision. The use of cost minimization framework
facilitates the interpretation of the result and it is in a better accordance with well established managerial practices. In modern manufacturing processes, most of the characteristics of the raw materials are well defined exhibiting little variation in the optimal consumption utility that can be achieved and therefore the focus is in the cost of acquiring such materials. This approach has being used by previous studies (Noordewier, John and Nevin 1990) and it is the base for material requirements planning systems (MRP) that many companies use to support their procurement decisions (e.g. Banker and Kauffman 2004)

Dual sourcing is a practice in which a company splits the production between different firms (Lyon 2006), and represents a stream of literature on industrial buying that is particularly relevant for our study. Previous research documents several reasons why companies would want to follow this procurement strategy including the protection during times of shortage, to keep a market feeling and to enhance competition between suppliers (Klotz and Chatterjee 1995). In our setting, dual sourcing occurs when a customer decide to buy their raw material from more than one supplier as a result of strategic considerations such as the continuous monitoring of qualities and prices of multiple vendors or to reduce the possibility of facing out of stocks. In private communication with the salesforce we found these to be important considerations in routine procurement decisions and therefore are explicitly included in our model. Thus, our study contributes a relatively scarce empirical example on industrial buyer behavior, while most of research is either based on survey data or descriptive in nature (see for example Puto, Patton and King 1985 or Rauyruen et al 2007)

2.2. Consumer Buying Behavior

*Multi-category buying:* From a methodological standpoint, we are interested in the estimation of demand functions when the customer purchases a basket of products but the competitor prices are
not observed. In this sense, we need to discuss the literature on multi-category choice behavior including models of store choice, incidence, brand choice and the quantity decision (Seetharaman et al. 2005). In recent years great progress has been achieved in this area, but the vast majority of this literature analyzes final consumer behavior which poses important differences with our industrial buyers. For example, given the relatively fixed material requirements of industrial buyers, there is very limited substitution between products which is a central focus of the multi-category choice behavior literature (e.g., a restaurant cannot substitute fish for beef when the customer orders the former, but end consumers are free to make these substitutions). It is possible that if a consumer purchases one product it might motivate the purchase of a complementary product during the same shopping trip. However, this is less likely to occur for industrial buyers with well defined material requirements.

Another difference comes when studying purchase incidence. While final customer purchase timing is heavily influenced by heterogeneity in consumption, industrial buyers have a relatively fixed purchase frequency determined by their inventory policies. Finally, as is pointed out by Seetharaman et al (2005) there is almost no research on customer-level multi-category models of quantity outcomes which is one of the central questions for organization buyers acquiring large volumes. To the best of our knowledge this is the first model that empirically investigates how industrial buyers make simultaneous purchases in several categories.

Store Choice: Among all studies on multi-category purchases, store choice models most closely resemble our problem. In both store choice and our problem customers need to decide where to buy (or the identity of the firm who is going to provide the product) and the quantities they are going to buy of each of the products being offered. However, there are profound differences that complicate the translation of store choice model ideas to our setting. First, most of store choice
models analyze the tradeoff between the geographic distances of consumer to store and basket attractiveness of the store offering (Bell and Lattin 1998), while in our problem the marginal cost of store switch is almost negligible. Second most of store choice models reasonably assume a hierarchical decision structure where actual quantities to be purchased are made conditional on store choice (Bodapati and Srinivasan 2001). In our problem the vendors are those who visit the customers every period, and therefore we consider the hierarchical decision structure is not appropriate.

The estimation of demand systems with unobserved marketing effort has been an issue extensively studied in the marketing literature. The focus here is in the quantification and the characterization of the nature of the biases in parameter estimates resulting of ignoring marketing information (Erdem, Keane and Sun 1999). We take a different approach where we use economic theory to describe the nature of the relationship of observed and unobserved information and directly incorporate that relationship into the estimation (Shugan 2004). This is because we are precisely interested in making inferences about those unobservables.

As mentioned earlier our problem share some commonalities with Moon, Kamakura and Ledolter (2007) who propose a hidden Markov model, which takes the unobserved promotion level by competitors as a latent variable driven by a Markov process with two states (“promotion” and “no promotion”). They estimated this model simultaneously with a promotion response model for a pharmaceutical product. There are important drawbacks to apply this approach to our problem. First, it would need an inconveniently large state space to account for the large price variation we observe in our problem, but most importantly we believe that unlike price promotion, negotiated prices are not properly described as a first order Markov process.
Share of Wallet: Finally we consider the literature on share of wallet investigating the relative intensity of the relation of the customer with the firm with respect to the competitors. As in our research, a major main challenge here is how to make inferences about customer market potential with sparse information about the customer transactions with the competitor. To estimate the share of wallet, Du. et al (2007) augment the database by collecting survey information for a sample of customers regarding their service usage in the competitor firms and link the data to transactions within the firm. This enables the authors to evaluate the power of several competitive models to predict market potential to then project it to the whole database. The methodology is applied to a proprietary data set provided by a major US bank and concludes among others that longer relationships are not associated to larger share of wallet, that customers with high share in one category also tend to have high share in others and that share and total purchase might be negatively correlated.

Fox and Thomas (2006) use a hierarchical Bayesian Tobit model to predict customer expenditure separately at each of competitor in a retail environment. Several sets of regressors are tested including shopping behavior of the focal firm, customer demographics and retailer geographic variables. Empirical results suggest that geographic information (e.g. travel time to the store) has the highest explanatory power. Based on individual retailer expenditure estimates, a second model is suggested to estimate expenditures in other retailers conditional on the data the retailers most usually observe: the transactions in their own stores. Unlike ours all these studies identify the market potential by directly observing the intensity of the transaction with the competitors. Glady and Croux (2009) propose a statistical model to estimate share of wallet using only transactional data from the focal company. However, given the limited data requirement, model needs to make relatively strong distributional assumptions.
Analysis of share of wallets for industrial buyers is even scarcer. Keiningham, Perkins-Munn and Evans (2003) use regression and Chi-square automatic interaction detection (CHAID) methods to analyze the connection between customer satisfaction and share of wallet finding a positive, but nonlinear relationship. Anderson and Narus (2003) outline how share of wallet could be strategically use to generate value, but they do not propose any formal method to estimate share of wallet beyond a database of customer’s shares populated with self reported information.

3. An economic based model for supplier selection

Following a long tradition in inventory management (see for example Silver 1981), we assume that customers minimize the total cost $TC$ associated with purchases across $M$ product categories for each period. Cost minimization is subject to the constraint that minimum levels of raw materials are available for production (Puto, Patton and King 1985).

$$\min \quad TC \left( q_{int}, q^c_{int} \right)_{m=1}^M$$

s.t. $\quad q_{int} + q^c_{int} \geq \tau_{int} \quad t = 1...T, m = 1...M$

where $q_{int}$ and $q^c_{int}$ are the quantity purchased by customer $i$ in category $m$ in period $t$ from the focal and the competitor companies respectively and $\tau_{int}$ is the minimum quantity required by the customer to cover his consumption in the following period.

We conjecture that total costs can be approximated by a quadratic equation, which is the sum of expenditures from each supplier and an interaction term between the quantities purchased:

$$TC \left( q_{int}, q^c_{int} \right)_{m=1}^M = \sum_{m=1}^M \left( p_{int} q_{int} + p^c_{int} q^c_{int} - \psi_{int} q_{int} q^c_{int} \right)$$

Where $p_{int}$ and $p^c_{int}$ are the prices charged by the focal and competitor firm respectively and parameters $\psi_{int}$ capture customer propensity to have multiple suppliers. We impose the condition:
\( \psi_{\text{int}} > 0 \), which implies that if prices are equal, then the firm always prefers to buy from more than one distributor (Lyon 2006). This interaction term between quantities ordered from each vendor is an important driver of consumer behavior and does not represent a direct cost as the first terms, but an indirect benefit associated with maintaining multiple suppliers which may improve price knowledge or increases competition amongst the suppliers to the benefit of the buyer. Notice that if a corner solution occurs (i.e., all purchases are made at one supplier or another) than the interaction term vanishes and we are left with the usual cost minimization problem.

The assumptions underlying the functional form to describe the propensity to split purchases have an important impact on our model. We make two implicit assumptions in our model. First, we assume that we can aggregate all relevant competitors into a single agent. Considering a unique firm to model competition is a common practice in both the academic literature and in industry practices (Chakravarthi 1988 and Moon et al. 2007). Unlike other applications that interpret the aggregate as the average behavior of the competitor (e.g Shankar and Bolton 2004), in our investigation we interpret that the aggregation represents the competitor with the most attractive value proposition for the customer. With respect to the functional form, we adopt a simple multiplicative form that has been extensively used in the Economics and Marketing literature to model demand interactions (e.g. Goic et al 2010, Arya et al 2008). There are certainly alternative specifications that can be tested. For example, we could consider an interaction term for each pair of potential competitor. However that model would be appropriate only if the customers have motivations to buy from each of the supplier distributing their consumption evenly among them. Instead, the customers in our empirical setting only want to avoid the situation where they concentrate all their purchases from a single distributor.
There are two further generalizations that we have considered but not implemented since industry managers suggest that these are not important drivers of buyer behavior. First, we could add nonlinear discounts, so that as quantity increases unit prices decrease. However, our conversations with managers suggest that given the limited inventory capacity of the customers and the high frequency of the customer purchases, the desire of buying from multiple suppliers is much more important than nonlinear discounts (see identification discussion in §3.1). Second, we could introduce a fixed ordering cost to reflect buyer transaction costs. However, in our industry the suppliers’ salesforce actively seeks out buyers on weekly visits to collect order and negotiate prices, hence the transactions costs are born largely by the supplier and not the buyer. Hence, these transactions can be considered small and irrelevant to supplier selection.

3.1. Deriving Demand from Cost Minimization

To characterize firm’s optimal behavior we derive the first order conditions on the minimization cost problem defined in (2.1). Given our assumptions about the cost function, the problem is separable between categories. Let $\lambda_{imt}$ be the shadow price associated to the minimum quantity constraints. Then, the optimal solution is characterized by the following system of equations.

$$0, \quad \text{with equality if } 0 \psi_{imt} q_{imt} \geq \lambda_{imt}$$  \hspace{1cm} (2.3)

$$0, \quad \text{with equality if } 0 \psi_{imt} q_{imt} \geq \lambda_{imt}$$  \hspace{1cm} (2.4)

$$c_{imt} = q_{imt}$$  \hspace{1cm} (2.5)

The identification conditions where interior and boundary solutions occur is a key element to the estimation of a model with imperfectly observed demand. In every period the customer could be in one of three cases: (a) the whole demand is satisfied from the focal retailer ($q_{imt} = \tau_{imt}$), (b) the
whole demand is satisfied from the competitor \( q_{int}^c = \tau_{int} \) and (c) the interior solution where customer buys from the focal and competitor firms in which case demands are given by:

\[
q_{int} = \frac{1}{2} \left( \tau_{int} - \frac{\delta_{int}}{\psi_{int}} \right) \quad q_{int}^c = \frac{1}{2} \left( \tau_{int} + \frac{\delta_{int}}{\psi_{int}} \right)
\]

where \( \delta_{int} = p_{int} - p_{int}^c \) is the price difference between the focal form and its competitors.

We point out that our solution has a similarity with Kim et al (2002) in that we predict optimizing behavior through first order conditions. However, they are interested in estimating product choice conditional on expenditure and therefore they take differences of the first order conditions to reflect the budget constraint. In our application, the value of the demand function is important and needs to be estimated.

The decision of buying from one or other supplier depends on the prices, but also on the parameters \( \psi_{int} \). The conditions characterizing the boundaries of each scenario can be derived by intersecting the corresponding first order conditions (see section 1 in Technical Appendix 2). We define sets \( \Omega_1 \) and \( \Omega_2 \) as the regions in the parameter space where the customer buy from the competitor and focal firm respectively while the set \( \Omega_3 \) correspond to the case where the customer buys from both.

\[
\begin{align*}
\Omega_1 &= \left\{ (\tau_{int}, \delta_{int}, \psi_{int}) : \psi_{int} \leq \delta_{int} / \tau_{int} \right\} \\
\Omega_2 &= \left\{ (\tau_{int}, \delta_{int}, \psi_{int}) : \psi_{int} \leq -\delta_{int} / \tau_{int} \right\} \\
\Omega_3 &= \left\{ (\tau_{int}, \delta_{int}, \psi_{int}) : \psi_{int} \geq \delta_{int} / \tau_{int} \right\}
\end{align*}
\]

A graphical depiction of the regions where each alternative is chosen in a plane \( \delta_{int} - \psi_{int} \) is shown in Figure 1. Remember we impose a positivity condition on \( \psi_{int} \) so that the customer always has a motivation to buy from more than one provider. If the difference in prices is small, then a
small value of \( \psi_{int} \) will suffice to motivate the customer to buy from multiple firms. However if the differences in prices is large, the firm will buy from a single firm unless the benefits of splitting the purchases is very large. In the extreme case where \( \psi_{int} \rightarrow 0 \), the firm has no incentives to buy from more than a single firm and the customer will always buy from the provider with the lower price.

\[
\psi_{int} = \frac{\delta_{int}}{\tau_{int}}
\]

**Figure 1**: Graphical description of first order condition cases

Therefore, the conditional likelihood of the observed demand \( q_{int} \) is defined by parts as follows:

\[
q_{int} = \begin{cases} 
\frac{\tau_{int}}{2} \left( \frac{\delta_{int}}{\psi_{int}} \right) & \text{if } (\tau_{int}, \delta_{int}, \psi_{int}) \in \Omega_2 \\
\frac{1}{2} \left( \tau_{int} - \frac{\delta_{int}}{\psi_{int}} \right) & \text{if } (\tau_{int}, \delta_{int}, \psi_{int}) \in \Omega_3 \\
0 & \text{if } (\tau_{int}, \delta_{int}, \psi_{int}) \in \Omega_1
\end{cases}
\]  

(2.8)

The probability of being in each region of the parameter space \( (\Omega_k, k \in \{1,2,3\}) \) depends on the probabilistic model we define to describe individual level parameters. We assume that the nonlinear cost parameter \( \psi_{int} = \psi_i \) is constant in time and across categories. On the other hand, we assume
that both the required quantity ($\tau_{imt}$) and the price difference ($\delta_{imt}$) are described by the following regression equations.

$$\tau_{imt} = \bar{\tau}_{imt} (\beta_i, XT_{imt}) + \varepsilon_{1imt} \quad (2.9)$$

$$\delta_{imt} = \bar{\delta}_{imt} (\gamma_i, XD_{imt}) + \varepsilon_{2imt} \quad (2.10)$$

Where $XT_{imt}$ and $XD_{imt}$ are matrices of observable covariates, $\beta_i$ and $\gamma_i$ are parameters to be estimated and $\varepsilon_{1imt}$ and $\varepsilon_{2imt}$ are random perturbances.

Several predictor variables can be included to describe customer level cost parameters. For example, we could postulate that the product requirement $\tau_{imt}$ depend on the type of product the customer sells or the volume of purchases in previous periods. Price differences could be described as function of firm size or promotional activity or the overall volume of the customer is buying. For simplicity we assume that $\varepsilon_{1imt}$ and $\varepsilon_{2imt}$ are normally distributed with 0 mean and variances $\sigma_1^2$ and $\sigma_2^2$ respectively. Under this normality assumption and assuming positive product requirements, the probability of being in each region of the parameter space is given by (Hinkley 1969):

$$p\left(\Omega_1 \mid \bar{\tau}_{imt}, \bar{\delta}_{imt}, \psi_i, \sigma_1^2, \sigma_2^2\right) = 1 - \Phi\left(\bar{\tau}_{imt} - \bar{\delta}_{imt}/\sqrt{\sigma_1^2/\sigma_2^2} - 1/\sigma_1^2\right)$$

$$p\left(\Omega_2 \mid \bar{\tau}_{imt}, \bar{\delta}_{imt}, \psi_i, \sigma_1^2, \sigma_2^2\right) = \Phi\left(\bar{\tau}_{imt} - \bar{\delta}_{imt}/\sqrt{\sigma_1^2/\sigma_2^2} - 1/\sigma_1^2\right)$$

$$p\left(\Omega_3 \mid \bar{\tau}_{imt}, \bar{\delta}_{imt}, \psi_i, \sigma_1^2, \sigma_2^2\right) = 1 - p\left(\Omega_2 \mid \bar{\tau}_{imt}, \bar{\delta}_{imt}, \psi_i, \sigma_1^2, \sigma_2^2\right)$$

Finally, to derive the full likelihood of observing a demand $q_{imt}$ for $I$ customers in $M$ categories and $T$ periods, we use the law of total probability:

$$p\left(q_{imt} \mid \theta_i, \sigma_1^2, \sigma_2^2\right) = \prod_{i=1}^{I} \prod_{m=1}^{M} \prod_{t=1}^{T} \sum_{k=1}^{3} p\left(q_{imt} \mid \Omega_k\right) p\left(\Omega_k \mid \theta_i, \sigma_1^2, \sigma_2^2\right) \quad (2.12)$$
Where $\theta_i = (\vec{\beta}_i, \vec{\gamma}_i, \log(\psi_i))$ is the vector of individual parameters. To complete the model, we introduce customer heterogeneity by specifying that the parameters of the models come from a common population distribution, which for simplicity we assume normally distributed:

$$\theta_i = \Lambda \cdot z_i + \nu_i \quad \nu_i \sim N(0, V_\theta)$$  \hspace{1cm} (2.13)

where $z_i$ are customer specific characteristics such as store size, whether it belong to a chain or not among other characteristics. The inclusion of a hierarchical structure plays a central role in the identification of demand parameters with incomplete information. By looking at the demand of other customers with similar demographic profiles we can evaluate how likely is that the customer have reduced his or her consumption and compared against the likelihood of the customer buying from a competitor firm. To complete the model, we specify the prior distribution as follows:

$$V_\theta \sim IW(\nu, V)$$  \hspace{1cm} (2.14)

$$\vec{\nu}(\Lambda) \big| V_\theta \sim N(\vec{\nu}(\bar{\Lambda}), V_\theta \otimes A^{-1})$$  \hspace{1cm} (2.15)

$$\sigma_{1i}^2 \sim \nu_1 ssq_1 / \chi^2_{\nu_1} \quad \sigma_{2i}^2 \sim \nu_2 ssq_2 / \chi^2_{\nu_2}$$  \hspace{1cm} (2.16)

where $\nu, \nu_1, \nu_2, ssq_1, ssq_2, V, \bar{\Lambda}$ and $A$ are chosen to have relatively diffuse priors.

### 3.2. Model Identification

Identification is an important concern due to our incomplete information set. In our model we impose three constraints to identify the structural parameters of the model.

1. The set of regressors used to describe product requirements $\tau_{int}$ cannot have common elements with the set of regressors used to describe price differences $\delta_{int}$. If the same regressor is present in both sets, the regression model can face a multicollinearity problem.
Notice that in the interior solution \((\Omega_3)\) both terms enter additively into the regression equation. In our empirical application we remove the intercept from the matrix of price difference regressors \(XD_{int}\). To facilitate interpretation of posterior estimates we also normalize all variables in \(XD_{int}\) resulting in price differences centered around zero.

2. The variance of price differences \(\sigma^2_{i}\) and the split purchase parameter \(\psi_{i}\) is not jointly identified. The reason is that under the interior solution condition the term \(1/\psi_{i}\) scales the price differences \(\delta_{int}\). Intuitively, if we observe a customer frequently switching from suppliers with no additional information we cannot know if this is because price fluctuations are large or because the customer is very sensitive to price variations. We therefore only estimate the ratio \(\delta_{int}/\psi_{i}\). Alternatively we could use our knowledge of price dispersion to impose tight priors on \(\sigma^2_{i}\). For example, if we assume the price of focal and competitor firm are completely independent and have the same dispersion, then the variance of the price difference is given by \(\nu(\delta_{int} = p_{int} - \hat{p}_{int}) = 2\sigma_{pim}^2\) where \(\sigma_{pim}^2\) is the variation of the focal price which we observe.

3. If other nonlinear components are added to the specification of the procurement cost, they are identified only if they can be described as a function of observable covariates that are disjoint from those affecting \(\psi_{int}\). For example, if we want to include price discounts and propose they vary linearly with firm size, then \(\psi_{int}\) cannot be a linear function of firm size. The reason is that their propensity to split demand has the same effect but in opposite directions than quantity discounts which motivates the customer to consolidate purchases from one supplier. In other words we can identify the net effect but not its components unless we have more information to disentangle them.
4. **Empirical Application**

We apply our proposed model to sales transaction data provided by a wholesale food distributor for its customers that is located within a large US metropolitan area. The distributor does not wish to be identified. The distributor’s customers are independent restaurants that are typically small operations with sales of between $1-5m in sales per year. The distributor employs its own salesforce and allows it sales representatives to have a high level of autonomy in negotiating prices with customers. This has led to high variability in prices across customers, as well as a high variability in the service quality provided to these customers. Transactional data correspond to two years of sales spanning from the second semester of 2008 to the first semester of 2010.

4.1. **Selection and Description of the Data**

The data we use in our empirical application comes from two sources: transactional databases with sales and price information and a survey of demographic characteristics of the customers. The vector of transactional information for each product $k$ which is a member of category $m$ that customer $i$ purchases at time $t$ consists of the quantity sold ($q_{ikt}$) and the negotiated price ($p_{ikt}$). Unfortunately, neither prices ($p_{ikt}$) nor demanded quantities ($q_{ikt}$) from the competitors are observed. For our application we aggregate demand to the category level and focus on the weekly level. An important feature of the problem is that prices are negotiated privately and therefore they could be different across customers even for the same period. Figure 2 plots time series of prices actually paid for the same product by a sample of 5 customers.
Figure 2: Time Series of paid price for two representative products and a sample of 5 customers

<table>
<thead>
<tr>
<th>Category</th>
<th>Min Price</th>
<th>Max Price</th>
<th>MAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>y</td>
<td>0.74</td>
<td>1.35</td>
<td>0.13</td>
</tr>
<tr>
<td>2</td>
<td>0.80</td>
<td>1.11</td>
<td>0.08</td>
</tr>
<tr>
<td>3</td>
<td>0.82</td>
<td>1.12</td>
<td>0.07</td>
</tr>
<tr>
<td>5</td>
<td>0.82</td>
<td>1.37</td>
<td>0.11</td>
</tr>
<tr>
<td>7</td>
<td>0.75</td>
<td>1.35</td>
<td>0.15</td>
</tr>
<tr>
<td>Average</td>
<td>0.77</td>
<td>1.30</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 1: Price Dispersion across categories.
Several interesting elements can be derived by inspecting these price series. First, we notice that negotiation processes result in important temporal variations in the prices paid by the same customer, but most importantly there are large differences in prices paid by different customers. Moreover, these differences could be systematically sustained even for products that are frequently purchased. When looking at the whole database we found a similar pattern. Table 1 displays descriptive statistic about the dispersion of the normalized price paid for customers for the same SKU in same week for each of the five categories considered in the analysis. For example if we look at category 6, on average the minimum price is 18% less than the average price while the maximum is 37% more than the average. In this category the price paid by a customer has in average an 11% difference with respect to the average.

The survey dataset includes several customer demographic variables such as the genre of restaurant and the number of employees. A brief description of the demographic information is given in Table 2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A</td>
<td>1 if customer produce product type A</td>
<td>0</td>
<td>1</td>
<td>0.886</td>
</tr>
<tr>
<td>Type B</td>
<td>1 if customer produce product type B</td>
<td>0</td>
<td>1</td>
<td>0.045</td>
</tr>
<tr>
<td>Type C</td>
<td>1 if customer produce product type C</td>
<td>0</td>
<td>1</td>
<td>0.015</td>
</tr>
<tr>
<td>Type D</td>
<td>1 if customer produce product type D</td>
<td>0</td>
<td>1</td>
<td>0.053</td>
</tr>
<tr>
<td>IsLargest</td>
<td>1 if the focal distributor is the main supplier</td>
<td>0</td>
<td>1</td>
<td>0.909</td>
</tr>
<tr>
<td>NEmployee</td>
<td>Number of employees in the store</td>
<td>1</td>
<td>125</td>
<td>28.07</td>
</tr>
</tbody>
</table>

Table 2: Description of customer demographic information.

For our empirical application, we restrict our attention to a subset of product categories that concentrate almost 80% of the transactions and more than 90% of the monetary volume (see §3 in
Appendix 2). In the resulting data set we select a random sample of customers with a minimum number of transactions in each of the selected categories. The final data set we use in our estimation consists of 132 customers making purchases in 5 categories for 104 weeks.

4.2. Econometric specification and Estimation

To estimate the model we need to choose the functional form of the models that describe the required quantities and price differences given in equations (2.9). As a first order approximation we use linear models for both regressions:

\[
\tau_{int} = X_{int} \beta_i + \epsilon_{1int} \tag{2.17}
\]

\[
\delta_{int} = X_{D_{int}} \gamma_i + \epsilon_{2int} \tag{2.18}
\]

Our set of regressors in \(X_{int}\) includes an intercept for each product category to reflect that customer’s have different needs in each category and a price coefficient which is assumed constant across categories. Notice that given we are describing product requirements and not actual sales we expect that price has only a moderate effect accounting for limited adjustments in both the mix of materials used to produce final products and short term inventories of ingredients.

For the matrices \(X_{D_{int}}\) we consider the sales in the product category from the previous period and the total volume of sales across all categories in the previous periods. The inclusion of the autoregressive components is motivated by the belief that the actual price customers pay may be influenced by the intensity of the relationship from previous periods. For example, a very loyal customer might pay more since they have demonstrated a preference for the services provided by the supplier. Alternatively infrequent or non-loyal customers may receive aggressive price offers so that the distributor may increase sales.

The hierarchical structure of the model makes it natural to employ a Bayesian approach and to estimate the model using Monte Carlo Markov Chain methods. The model derived from
imposing first order conditions destroys the conjugacy of the hierarchical linear model and therefore we use a Metropolis-Hastings step to update customer level parameters. The hyperdistributions which describe the population level parameters are computed using the standard conjugate updating. For a formal discussion of the sampler and the values of priors we use in the estimation see §2 in Appendix 2.

Given the Metropolis-Hastings step is conducted at the individual level parameter we need to find tuning parameters for each customer to get optimal mixing behavior of the chain. We therefore apply the adaptive tuning scheme by Haario et al. (2001) which is applied in the first half of the iterations of the burn-in period (for a formal discussion of sampling properties of and adaptive scheme, see Andrieu and Thoms (2008)). The chain is run for 20,000 iterations and we save every fifth iteration for making inferences about the posterior distribution. The first 15,000 iterations are discarded to allow for convergence while the last 5,000 are used for inference.

Before applying our proposed model to actual data, we tested our estimation approach on simulated data based on known parameter values and evaluate its ability to recover the “true” response parameters. After imposing all the identification constraints discussed in section 3.1 we found that all individual levels are recovered with high accuracy and our burn-in period is adequate.

4.3. Results

Results of our empirical application are organized as follows. First we start by discussing posterior estimates of the customer-level parameters. Given that prices are negotiated individually for each customer and potentially on a weekly basis, the analysis of individual level parameters forms the basis of any price recommendation. Figure 3 displays boxplots of price differences for all five categories and a sample of customers. Price differences are centered around zero, but there are important variations across categories. Thus, while the focal distributor might be offering very
competitive prices in one category it might be too expensive in others. Please notice that these differences are not constant across customers. For example some customers may be supplied by a very cheap local provider that is not available for other customers in other regions of our metropolitan area.

**Figure 3:** Individual level parameters for a sample of three customers.

The right panels of Figure 3 display the histogram of price coefficients. The values reported here are representative of others customers. For the majority of the customers, price coefficients are significant and negative. However there are several customers where the posterior distribution of the price distribution is massed around zero. Given that price coefficients measure the effect in the requirements of raw materials and not the actual demanded quantity, we consider it plausible that short-term price variations may have no effect in the mix of raw materials used to produce their final
products. The moderate values of most price coefficients suggest that customers effectively have relatively fixed material requirements. Mild negative values can be explained by minor adjustments in the production process and the use of operational inventories.

Figure 4: Time series of posterior means of the probabilities of being in each region of the parameter space for a sample of customers

As we discussed in section 3 each customer could be in three different states: buying from the competitor only (Ω₁), buying from the focal supplier only (Ω₂) or splitting his demand between the two (Ω₃). From individual level parameters we can compute the probabilities of being each state as is depicted in Figure 4. Please recall that these states are calculated on a weekly, category basis. The model predicts that observed customer demand can be rationalized in different ways. For example panel (a) shows it is very likely that the customer 8 is heavily buying from the competitors in category 4, while panel (b) shows that customer 53 is most likely splitting his demand between focal and competitor supplier in category 5. For other cases, the model is not very informative
suggesting similar probabilities of being in each of the regions as is illustrated in panel (c). Finally, by looking at the time series we could potentially identify customers who are exhibiting a change in purchase behavior. This is the case of customer 105 in category 5 in panel (d) where the focal firm is progressively gaining share of wallet from competitors.

Figure 5 compares actual sales vs. the posterior mean of the ones predicted by our model for a representative sample of customers and product categories. These plots suggest that even with a limited set of explanatory variables, the model can capture the basic components of the observed demand such as base levels and changes of regimes. However, there is a significant fraction of the variation that is not being explained by the model. Potentially the addition of more covariates could successfully improve the fit of the model.

![Figure 5: Time series of actual (solid black lines) and predicted (dotted gray lines) sales](image)

We now discuss the posterior estimates at the population level parameters $\mathbf{A}$ describing the relationship between individual level parameters and customer demographics. Population level
parameters can guide marketing efforts at a more strategic level and can be used to predict the behavior of new customers. Table 3 display posterior means, standard deviations and 90% credible intervals for all parameters in $\Lambda$.

Parameters in the upper right cells (rows $\beta_1$-$\beta_5$ and columns Type A-Type D) simply account for different category requirements per customer type. For example we could say that on average a customer producing a final product of Type B requires 178 more units of raw materials of category 2 than customers producing other types. Surprisingly, the number of employees plays no significant role in the amount of raw materials required by the customers.

<table>
<thead>
<tr>
<th>$(\beta_1, \gamma_1)$</th>
<th>Type A Mean (s.d)</th>
<th>Type B Mean (s.d)</th>
<th>Type C Mean (s.d)</th>
<th>Type D Mean (s.d)</th>
<th>IsLargest Mean (s.d)</th>
<th>NEmployee Mean (s.d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$ Mean (s.d)</td>
<td>65.2 (48.2)</td>
<td>164 (61.2)</td>
<td>12.6 (114)</td>
<td>222 (70.9)</td>
<td>0.767 (0.561)</td>
<td>53.2 (45.7)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(-28.7, 156)</td>
<td>(42.4, 289)</td>
<td>(-204, 239)</td>
<td>(79.4, 361)</td>
<td>(-0.306, 1.82)</td>
<td>(-31.8, 142)</td>
</tr>
<tr>
<td>$\beta_2$ Mean (s.d)</td>
<td>158 (44.4)</td>
<td>178 (56.5)</td>
<td>108 (106)</td>
<td>133 (63.9)</td>
<td>0.581 (0.506)</td>
<td>-5.46 (42.3)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(69.3, 248)</td>
<td>(69.3, 296)</td>
<td>(-105, 319)</td>
<td>(7.31, 256)</td>
<td>(-0.448, 1.59)</td>
<td>(-88.4, 81.6)</td>
</tr>
<tr>
<td>$\beta_3$ Mean (s.d)</td>
<td>42.9 (53.9)</td>
<td>96.7 (67.1)</td>
<td>-9.7 (117)</td>
<td>19.4 (78.4)</td>
<td>2.17 (0.631)</td>
<td>51.7 (52.4)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(-61.9, 150)</td>
<td>(-38.6, 223)</td>
<td>(-232, 217)</td>
<td>(-129, 168)</td>
<td>(0.912, 3.46)</td>
<td>(-50.1, 155)</td>
</tr>
<tr>
<td>$\beta_4$ Mean (s.d)</td>
<td>85.8 (30.5)</td>
<td>138 (38.8)</td>
<td>26.5 (72.2)</td>
<td>71.5 (44.4)</td>
<td>0.724 (0.354)</td>
<td>31.4 (29.0)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(26.9, 145)</td>
<td>(62.7, 216)</td>
<td>(-119, 168)</td>
<td>(-14.4, 157)</td>
<td>(0.001, 1.41)</td>
<td>(-24.8, 86.4)</td>
</tr>
<tr>
<td>$\beta_5$ Mean (s.d)</td>
<td>101 (29.5)</td>
<td>150 (37.4)</td>
<td>51.6 (70.1)</td>
<td>82 (42.7)</td>
<td>0.555 (0.341)</td>
<td>23.1 (28)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(44.7, 158)</td>
<td>(75.2, 222)</td>
<td>(-85.4, 188)</td>
<td>(-0.144, 164)</td>
<td>(-0.157, 1.21)</td>
<td>(-33.1, 75.8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\beta_6, \gamma_2)$</th>
<th>Type A Mean (s.d)</th>
<th>Type B Mean (s.d)</th>
<th>Type C Mean (s.d)</th>
<th>Type D Mean (s.d)</th>
<th>IsLargest Mean (s.d)</th>
<th>NEmployee Mean (s.d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_6$ Mean (s.d)</td>
<td>0.0797 (0.69)</td>
<td>-2.49 (0.85)</td>
<td>0.561 (1.59)</td>
<td>1.06 (1.03)</td>
<td>-0.011 (0.00776)</td>
<td>-0.978 (0.667)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(-1.26, 1.5)</td>
<td>(-4.11, -0.867)</td>
<td>(-2.7, 3.8)</td>
<td>(-0.913, 3.09)</td>
<td>(-0.0272, 0.0034)</td>
<td>(-2.33, 0.254)</td>
</tr>
<tr>
<td>$\gamma_1$ Mean (s.d)</td>
<td>-31.9 (37.1)</td>
<td>-2.97 (48.3)</td>
<td>-15.0 (87.3)</td>
<td>-83.2 (55.2)</td>
<td>-1.12 (0.436)</td>
<td>-23.6 (35.4)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(-106, 41.1)</td>
<td>(-102, 91.5)</td>
<td>(-186, 154)</td>
<td>(-194, 22.5)</td>
<td>(-1.93, -0.231)</td>
<td>(-92.3, 44.5)</td>
</tr>
<tr>
<td>$\gamma_2$ Mean (s.d)</td>
<td>11.5 (15.1)</td>
<td>11.0 (18.8)</td>
<td>18.3 (33.7)</td>
<td>-27.8 (21.4)</td>
<td>-0.494 (0.169)</td>
<td>-2.44 (14.5)</td>
</tr>
<tr>
<td>90% C.I</td>
<td>(-18.7, 41.7)</td>
<td>(-28.6, 47.1)</td>
<td>(-52.2, 78.6)</td>
<td>(-69.3, 12.0)</td>
<td>(-0.82, -0.152)</td>
<td>(-31.9, 26.2)</td>
</tr>
</tbody>
</table>

Table 3: Posterior of population level parameters

Interestingly, being the largest supplier has a negative effect on both gamma components that describe price differences. Given that the covariates used in $\delta_{mt}$ equations are functions of past sales, the negative estimates we found provide evidence that prices of more loyal customers are less
sensitive to purchases of the previous periods. The effect of having the focal distributor as the main supplier on the category requirements should be interpreted carefully. A positive estimate is interpreted as the propensity of the focal firm of being the main supplier for a customer with larger product requirements. From our results we conclude that it is more likely that customers with larger requirements in categories 3 and 4 are mainly supplied by the focal distributor.

Finally, we found that the only customer demographic variable which has a significant effect on the price coefficient is the Type B dummy variable. Recall that coefficient does not measure how the quantity demanded is going to change with price, but instead how the requirement of the product will change. Therefore, our results indicate that customers producing output Type B have more flexibility to adjust the volume of raw materials as a reaction to price changes.

5. Discussion and Conclusions

In this article we have proposed a new econometric model to describe how industrial buyers made their procurement decisions from their suppliers. The model is tailored for the situations where distributors primarily use their salesforce to visit customers for the purpose of taking orders and negotiating prices. Given that prices are individually negotiated, most of the information about customer activity with the competitor is not observed posing a great challenge for the study of the customer demand. Our model assumes that at every purchase occasion customers minimize their procurement costs subject to having enough raw materials to satisfy their own demand. First order conditions enable us to express the likelihood of observing sales as a function of price differences with respect to their competitors.

We apply our proposed model to a wholesale food distributor and we find widespread heterogeneity in purchase patterns. As expected some customers are loyal, while others are not, and
the remainder fall in between. The model fits the data well and appears to capture the main components of demand. More importantly it can shed light on the competitive elements of demand that cannot be studied with traditional reduced form response models. For example our approach provides a novel mechanism to infer share of wallet with incomplete information. In our empirical application we found that while some customers satisfy most of their requirement from one of their distributors, other consistently split their demands among them.

Our proposed methodology could also help to guide strategic decision making. Posterior estimates of population level parameters can provide managers with valuable information about the relative attractiveness of prices for different segments of customers. In our empirical application we found that price sensitivity of customers making most of their purchases with the focal supplier are less affected by the volume of purchases in previous periods.

We recognize many extensions of our current framework that we wish to study in the future. First, we could couple our description of procurement decisions with an inventory model that explicitly takes into account the possibility of stockpiling or delaying purchases. In our problem many products like fresh, refrigerated meat are perishable, so we felt that inventorying is less likely in our problem. However, inventory may be important in storable categories or in other industries. Second, we could consider strategic buying behavior by firms. If a buyer recognizes that future prices depend on current behavior, then the buyer may change their current behavior.

Second, our model specification can be improved. As we discussed in section 4.2, more covariates may add to the explanatory power of our model. Also, we could enrich the description of the interaction of purchases from different categories by explicitly introducing cross-effect terms. This may be especially helpful in deriving cross-selling strategies. Stochastic frontier is another modeling approach that could be used to enrich our description of the product requirement.
regression (Luo and Donthu 2005; Kim and Kim 2000). Third, in our specification we have assumed that propensities to split purchases for a given customer are constant across categories. A more general model could relax this assumption and study how those differences are correlated between categories. Also, if fixed cost were present, they could introduce interesting dynamics that are not capture by the current version of the model.

Fifth we have not formally compared the results of our model to any benchmark. Standard response models constitute natural comparisons, but they do not address the most relevant issues we study here such as competitor prices. We expect that response models would fit the data well and have good forecasting power, but remain silent about the unobservables. Finally and most immediately, we wish to focus upon optimal pricing in our current framework. Now that a firm can make inferences about competitor prices, how should the firm design an optimal pricing strategy?
References


Appendix

1. Derivation of indifference conditions

Case 1: The customer is indifferent between buying from focal firm only or buying from the competitor. For the boundary (2.6) to hold and also \( q_{\text{int}} = \tau_{\text{int}} \), then

\[
q_{\text{int}} = \frac{1}{2} \left( \tau_{\text{int}} + \frac{p_{\text{int}}^c - p_{\text{int}}}{\gamma + \eta} \right) \Rightarrow \gamma + \eta = \frac{1}{\tau_{\text{int}}} \left( p_{\text{int}}^c - p_{\text{int}} \right)
\]

Case 2: The customer is indifferent between buying from competitors only or also buying from the focal firm.

\[
q_{\text{int}}^c = \frac{1}{2} \left( \tau_{\text{int}} + \frac{p_{\text{int}}^c - p_{\text{int}}^c}{\gamma + \eta} \right) \Rightarrow \gamma + \eta = \frac{1}{\tau_{\text{int}}} \left( p_{\text{int}} - p_{\text{int}}^c \right)
\]

2. Sampler

In our empirical application, we draw from the posterior distribution following the following sequence:

1) Start with initial values of \( \theta_{\text{int}} \), \( \sigma_{\text{i1}}^2 \) and \( \sigma_{\text{2i}}^2 \)

2) For each customer

   a) Propose a new value \( \theta_{\text{int}} \) according to a random walk process. And accept it with probability

\[
\min \left\{ 1, \frac{p\left( \{q_{\text{int}}\} | \theta_{\text{old}}, \sigma_{\text{i1}}, \sigma_{\text{2i}} \) p\left( \theta_{\text{new}} \right) \right)}{p\left( \{q_{\text{int}}\} | \theta_{\text{old}}, \sigma_{\text{i1}}, \sigma_{\text{2i}} \) p\left( \theta_{\text{old}} \right) \right) \}
\]
b) Propose a new value \( \sigma_i^2 \) according to a random walk process. And accept it with probability

\[
\min \left\{ 1, \frac{p\left(\{q_{im}\} | \theta_i^{\text{new}}, \sigma_i^{2\text{new}}, \sigma_{2i}^{2\text{new}}\right)p\left(\sigma_{2i}^{2\text{new}}\right)}{p\left(\{q_{im}\} | \theta_i, \sigma_i^2, \sigma_{2i}^2\right)p\left(\sigma_{2i}^2\right)} \right\}
\]

c) Propose a new value \( \sigma_{2i}^2 \) according to a random walk process. And accept it with

probability \[
\min \left\{ 1, \frac{p\left(\{q_{im}\} | \theta_i^{\text{new}}, \sigma_i^{2\text{new}}, \sigma_{2i}^{2\text{new}}\right)p\left(\sigma_{2i}^{2\text{new}}\right)}{p\left(\{q_{im}\} | \theta_i, \sigma_i^2, \sigma_{2i}^2\right)p\left(\sigma_{2i}^2\right)} \right\}
\]

3) Update upper level regression parameters \( \Lambda \) and \( V_\theta \) following conjugate multivariate regression model. The posterior of the conjugate multivariate linear regression

\[
V_\theta | \theta_m, Z \sim IW(v + T, V + \tilde{S}) \\
\Lambda | \theta_m, Z, V_\theta \sim N\left(\tilde{\lambda}_{m}, V_\theta \otimes (Z'Z + A)^{-1}\right)
\]

Where

\[
\tilde{\lambda}_m = \text{vec}(\tilde{\Lambda}), \quad \tilde{\Lambda} = (Z'Z + A)^{-1}(Z'Z\hat{\Lambda} + A\tilde{\Lambda})
\]

\[
\tilde{S} = (\theta_m - Z\hat{\Lambda})' (\theta_m - Z\hat{\Lambda}) + (\hat{\Lambda} - \tilde{\Lambda})' A (\hat{\Lambda} - \tilde{\Lambda})
\]

4) Repeat as necessary.

In our empirical application we use the following hyper priors parameters: \( v = v_1 = v_2 = k + 3, \)

\[
ssq_1 = ssq_2 = 100, \quad V = v \cdot I_k, \quad A = 0.001 \cdot I_{n_z} \text{ and } \tilde{\Lambda} = \theta, \text{ the } n_z \times k \text{ zero matrix.}
\]

Our sampler is very general because makes no use of the special structure of the regression functions. However, there are several opportunities to adapt our proposed method to gain in computational efficiency. For example, conditional on the value of the regime \( \Omega_k (k \in \{1, 2, 3\}) \) the rest of the parameters could be estimated using the standard hierarchical linear regression model.
Here, the conditional posterior distribution is known exactly and therefore we can directly sample from it without needing to discard draws from burning initial iterations or thinning the chain. To implement this sampling approach we would also need to efficiently sample from $\Omega_k$.

Unfortunately, our computational experiments in sampling from those regime probabilities show that navigating the augmented parameter space is no faster than the other sampler we propose. We consider there are interesting opportunities to improve the efficiency of the sampler and we leave this issue for future research.

3. Category Selection and Description

In the transactional database we observe purchases made in 9 categories. For our empirical application we select only the 5 categories with larger volumes. The distribution of quantities and expenditures per categories are displayed in Figure TA2.1.

Figure TA2.1: Distribution of transaction per product category

Figure TA2.1 displays the aggregated time series of sales for the 5 selected categories. From a visual inspection of the plots we found no strong evidence of seasonal or trends effects.
Figure TA2.2: Aggregated time series for the selected product categories.