

DOCTORAL DISSERTATION

Essays on Multi-product Pricing

submitted to the
David A. Tepper School of Business
in partial fulfillment for the requirements for the degree of

DOCTOR OF PHILOSOPHY
in Business Administration (Marketing)

By Marcel Goic

Carnegie Mellon University
Tepper School of Business
Pittsburgh, Pennsylvania 15213

DISSERTATION COMMITTEE:

Alan Montgomery (Chair)
Kannan Srinivasan
Kinshuk Jerath
Vishal Singh

November 2010

Acknowledgements

Contents

Chapter 1: Introduction

Chapter 2: Making Better Pricing Decisions with Implied Priors

1.	Introduction.....	9
2.	Illustration of Implied Priors for Univariate Price Optimization.....	13
	Illustration with a Conjugate Prior.....	13
	Illustration with a Non-conjugate, Uniform prior	16
	Discussion	19
3.	Types of Price Information for Implied Priors	19
4.	Sampling Methodology.....	21
5.	Empirical Application.....	22
	Understanding the Prior Distribution	25
	Posterior Estimates	27
6.	Discussion and Conclusions.....	28

Chapter 3: Cross-Market Discounts

1.	Introduction.....	30
2.	Comparison with Related Promotion Strategies.....	37
3.	Model.....	39
4.	Analysis.....	42
4.1.	A simple case: the monopoly-monopoly scenario	42
	Different rates of diminishing marginal utilities	47

Quantity discounts limited to the source market	48
Non-zero marginal cost	50
4.2. The monopoly-duopoly scenario.....	51
5. Extensions	56
5.1. The duopoly-duopoly scenario.....	56
5.2. Discount based on expenditure in the source market.....	60
5.3. Other assumptions	62
6. Conclusions and Discussion.....	63

Chapter 4: Inferring Competitor Pricing with Incomplete Information

1. Introduction.....	67
2. Research on Demand under Unobserved Competition.....	70
2.1. Industrial Buyer Behavior	70
2.2. Consumer Buying Behavior	72
3. An economic based model for supplier selection	75
3.1. Deriving Demand from Cost Minimization	76
3.2. Model Identification.....	81
4. Empirical Application.....	82
4.1. Selection and Description of the Data.....	82
4.2. Econometric specification and Estimation.....	84
4.3. Results.....	86
5. Discussion and Conclusions.....	91

References 94

Technical Appendix 1: Cross-Market Discounts

1. Extensions of the monopoly-monopoly scenario 100

1.1. Analysis with different rates of diminishing marginal utilities 100

1.2. Analysis of quantity discounts limited to the source market 104

1.3. Analysis with nonzero marginal cost..... 107

2. Analysis of the marginal effects of parameters in the monopoly-duopoly scenario..... 109

3. Analysis of the duopoly-duopoly scenario 112

3.1. Without competing cross-market discount program 112

3.2. With competing cross-market discount program..... 116

4. Analysis of cross-market discounts based on expenditure in the source market 119

Technical Appendix 2: Inferring Competitor Pricing with Incomplete Information

1. Derivation of indifference conditions 124

2. Sampler 124

3. Category Selection and Description 126

Chapter 1

Introduction

Managers often make price decisions for several products simultaneously. By doing so, decision makers can control for substitution effects or take advantage of potential synergies between products. My dissertation consists of three essays that investigate novel aspects of this multi-product pricing approach.

My first essay, analyzes the implications in optimal prices and profits of using the information contained in the definition of the feasible set in the estimation of the demand parameters. Traditional approaches to price optimization take a two step approach to setting prices. First the parameters of the demand system are estimated given an observed dataset, and second a profit maximization problem is formulated to decide upon the optimal prices. Often the profit function derived from the estimated demand model is either ill-formed or the answers are non-sensical. Hence, the manager imposes a set of constraints on the prices of all products to identify a more appropriate solution. This process is not consistent with a Bayesian approach, since the manager's constraints on the price solution represent prior information and this information should be incorporated into the prior distribution of the parameter estimates. In this essay we illustrate how statements about optimal prices imply informative prior distributions that can be used in a traditional Bayesian approach. This Bayesian method improves the quality of the pricing decisions

made by managers. It demonstrates how managers can be viewed as experts who have well developed opinions about how prices should be set.

My second essay explores “cross-market discounts” where firms try to attract consumers by offering discounts in other unrelated markets. A prominent example of this promotion strategy is the discounts in gasoline offered by many grocery retailers across North America, Europe and Australia. In this essay, we use an analytical model to investigate the major forces driving the profitability of this novel promotion strategy. We consider a generalized scenario in which purchases in a source market lead to price discounts redeemable in a target market. Our analysis shows that this strategy can be a revenue driver by simultaneously increasing prices as well as sales in the source market, even though sales are negatively elastic in price, *ceteris paribus*. Moreover, it distributes additional consumption (motivated by the discount) in two markets and, under diminishing marginal returns from consumption, this can simultaneously increase firm profits and consumer welfare more effectively than traditional nonlinear pricing strategies. Our study provides many other interesting insights as well, and our key results are in accordance with anecdotal evidence obtained from managers and industry publications.

My third essay studies 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. For mature markets, any increase in sales for an existing customer must come at the expense of the sales of another competing distributor. One would expect these competing distributors to respond either directly by lowering prices for targeted customers or more broadly by lowering prices for non-targeted customers. 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 and we impose first order conditions to characterize regions in the parameter space where consumer will buy from each distributor. We use those conditions to estimate an empirical model of purchase behavior that enables us to identify the likelihood of each consumer buying from the competitor or simply changing his consumption patterns.

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 empirical results shed light on the competitive elements of customer demand that cannot be study with traditional reduced form response models. For example we found that while some costumers 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. 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 expect this result to provide valuable information for vendors to negotiate prices with the customers.

Chapter 2

Making Better Pricing Decisions with Implied Priors

1. Introduction

A basic tenet of Bayesian inference is the construction of a prior probability distribution over the parameter space (Box and Tiao 1971, Zellner 1973). The purpose of the prior is to reflect information known by the analyst. In this research we are especially interested in understanding how optimal price decisions using a demand model are impacted by these prior distributions and conversely how prior expectations about the optimal price can be used to improve the estimates of the demand model and consequently to make better pricing decisions.

Traditionally analysts approach price optimization by estimating a demand model and then making inferences from the model to optimize prices. Quite often analysts are not satisfied with the solutions that are recommended from this approach and place ad hoc constraints to improve the quality of the optimal price recommendations (Reibstein and Gatignon 1984, Montgomery 1997, Khan and Jain 2005). For example, bounds are placed on the price ranges (Deng and Yano 2006). We argue that these constraints reflect prior information, and that treating them in an ad hoc manner yields inferior solutions.

We demonstrate how information used to constrain pricing solutions during the inference stage can be translated into prior information. Our method preserves the integrity of the Bayesian approach, but just as importantly it leads to better decision making. We find that pricing solutions with constraints imposed after the estimation can lead to quite different price elasticity estimates and optimal price solutions than when constraints are treated as prior information about the parameters. Practically this method reverses the use of constraints in current practice. Instead of using managerial knowledge to impose constraints on the pricing decision space, our method constrains the feasible parameter space.

It is clear that information is the key to the successful solution of the price optimization problem. The primary source of this information for many retail price optimization problems is previous retail transactions that record how consumers have reacted to price, promotion, and assortment changes in the retail environment. However, transaction information may not be sufficient—even in transaction rich environments like supermarkets with high turnover and large customer bases. Pricing decisions frequently yield huge decision spaces and traditional, orthogonally designed experiments may be inadequate to supply the information for a purely data based approach. For example, a single store pricing a single category with 20 products and just a high and low price for each product would yield more than a million potential price combinations, while 100 products would yield more than a googol.

Therefore, to solve the pricing problem we must have more information. Information from managers is necessary and in fact desirable for the demand modeling process. Managers can be thought of experts who have important insights into the problem. The combination of expert predictions has a long history in the scientific study of forecasting (Huber 1975, Clemen 1989). Prior information that is typically used includes assumptions about the functional form

(Montgomery and Bradlow 1999) or economic theory that reduces the parameter space (Montgomery and Rossi 1999).

Traditionally analysts do not explicitly recognize this prior information, while the Bayesian approach advocates formalizing this information so that it is explicitly recognized. We argue that explicitly recognizing information through prior distributions is beneficial to the scientific approach since it forces managers to acknowledge their assumptions. Unfortunately, many researchers are uncomfortable with prior elicitation and rely on diffuse priors. Therefore, we believe the elicitation of prior information is a neglected aspect of Bayesian inference and one we seek to improve.

Current price optimization solutions like Montgomery (1997) use managerial input in an ad hoc manner. The manager may place bounds upon the optimal prices that are predicted from the demand model. The reason that managers impose these constraints is that they have well developed opinions about how prices should be set. Logically there is a problem with using these constraints only during the price optimization phase. Consider that the optimal price is a function of the price elasticity, and any prior statements or constraints on the optimal prices implicitly define a prior on the price elasticity. This information should not be applied after the estimation but should be considered prior information, since it is a prior belief that the manager holds even before looking at the results of the model. Therefore, managerial opinions and constraints about optimal price solutions are most appropriately characterized as prior information.

The standard Bayesian approach to prior construction is that the analyst must define a joint distribution over the parameter space of unknown variables. Unfortunately, most analysts have little knowledge about these parameters and subsequently have difficulty setting these parameters directly. Research within the decision support literature (Chakravarti et al 1981, Wierenga et al 1999) has shown that results are mixed when analysts directly assess the parameters, since non-technical users

may not fully understand the implications of their selections. This may be one reason that prior construction is avoided and that diffuse priors are used instead. Our approach may help alleviate problems associated with prior elicitation. Instead of asking analysts to create priors about unobserved parameters, they are instead able to make statements about the outcomes of these parameters for which they may have better insights.

The general approach that we advocate in prior elicitation is to translate simple, direct statements that managers have about their problem into informative prior distributions. The type of information that we elicit might include the range of the optimal prices that would be considered, which price points are thought to be valid, or the relative ordering of product prices. These statements often define marginal properties about inferences made from the model and not to directly make statements about parameters themselves. This approach allows managers to make meaningful statements that populate their prior beliefs and then integrate these priors using a consistent Bayesian method. The idea of constructing priors from marginal properties is not original to this research; Allenby and Rossi (1993) employ such a method in constructing a prior for a multinomial logit model. Sandor and Wedel (2001) also employ a similar idea by eliciting market share information from managers in order to improve the design of conjoint experiments.

In section 2 we provide an example of our approach for a univariate price optimization problem, and demonstrate why ignoring prior information or incorporating it ad hoc leads to inefficient decisions. Section 3 provides a more general taxonomy of prior information and what type of managerial information may be relevant. We define a general methodology using accept-reject sampling techniques (Robert and Casella 2004) in section 4 to estimate the posterior using our implied priors. Section 5 presents an application to a category pricing problem for refrigerated

orange juice and shows large differences in the price solution between the traditional approach and our Bayesian approach. Section 6 concludes with a discussion of our findings.

2. Illustration of Implied Priors for Univariate Price Optimization

Our first goal is to demonstrate through a simple example that a priori information about the parameters—even if it is meant to be diffuse—can be quite informative concerning the parameters, and vice versa. Consider the commonly used double-log model where p is the price and q the corresponding demanded quantity (Wittink et al. (1988), Kalyanam (1996), Kopalle et al. (1999), Montgomery and Bradlow (1999) among others):

$$\ln(q) = \alpha + \beta \ln(p) + \varepsilon, \text{ where } \varepsilon \sim N(0, \sigma^2) \quad (1)$$

In this model the price parameter, β , is also the price elasticity. If we assume that the variable cost associated with a unit is c , then the first order condition on the expected profit implies that the optimal price is given by:

$$p^* = f(\beta) = \frac{c}{1 + 1/\beta} \quad (2)$$

For simplicity we will assume that other model parameters: α , σ and c are fixed and known, and without loss of generality we assume that $c = 1$.

Illustration with a Conjugate Prior

Suppose the analyst makes the assumption that the price coefficient follows a conjugate normal distribution:

$$\beta \sim N(\bar{\beta}, V_\beta^2), \text{ or } p_\beta(\beta) = \frac{1}{\sqrt{2\pi V_\beta^2}} \exp\left\{-\frac{(\beta - \bar{\beta})^2}{2V_\beta^2}\right\} \quad (3)$$

Additionally, we assume that $\bar{\beta}$ is known (e.g., $\bar{\beta} = -3$) and focus on the impact of the variance. It may seem odd that we fix the mean of the prior but not the variance, but for the purpose of this example our emphasis is on the precision of the analyst's beliefs and not the location. Notice that for the trivial example where $\bar{\beta} = -3$ and $V_{\beta} = 0$, or equivalently when β is known with certainty, then the optimal price from equation (2) would be \$1.50 and the markup would be 50% over cost.

The optimal price in equation (2) is a function of β , so we can compute the prior on the optimal price using a standard change of variables approach of equation (3). We refer to this as an *implied* prior. We compute the implied prior for our problem as follows:

$$p_p(p^*) = p_{\beta}\left(f^{-1}(p^*)\right) \cdot \left| \frac{df^{-1}}{dp^*} \right| = p_{\beta}\left(\frac{p^*}{c-p^*}\right) \cdot \frac{c}{(c-p^*)^2}, \quad (4)$$

We can compute the inverse function of equation (2) as: $\beta = f^{-1}(p^*) = p^*/(c-p^*)$.

To fully specify our conjugate prior in equation (3) we must choose a value for V_{β} . Generally, the analyst would directly choose this value perhaps based upon previous experience, meta-analyses, or select large values that reflect diffuse knowledge. Instead of choosing the value directly, we can reparameterize our prior with respect to the probability that the optimal price falls above p_0 is ρ :

$$\Pr(p^* \geq p_0) = \rho \Leftrightarrow \Pr(\beta \geq \beta_0) = \rho, \text{ where } \beta_0 = \frac{p_0}{c-p_0} \quad (5)$$

We can now rewrite this equation to find the corresponding variance parameter as a function of the other parameters:

$$V_{\beta}^2 = \left(\frac{\beta_0 - \bar{\beta}}{\Phi^{-1}(1-\rho)} \right)^2 \quad (6)$$

This prior distribution over the optimal prices is plotted in Figure 1 for various values of the precision parameter. Notice that this implied distribution for optimal price is asymmetric and thick-tailed, and not well approximated by a normal distribution.

One might expect that diffuse priors on the parameters lead to diffuse priors about the optimal prices. The goal of this example is to illustrate that this is not true. Apparently diffuse priors on the parameter space may lead to very informative priors on the optimal price space, and vice versa. Consider a diffuse prior specification on (3) in which V_β tends to ∞ and $\bar{\beta}$ is left unspecified. Under this assumption p^* is \$1 almost surely—at least a priori. Conversely, if the analyst were to conjecture that the optimal price possesses a diffuse distribution this has the perverse effect of pushing V_β to smaller values, and not larger ones. (See Figure 1 which illustrates smaller values of V_β tends to flatten out the implied distribution of optimal prices.) Hence, diffuse beliefs about optimal prices do not imply diffuse beliefs about price sensitivity parameters and vice versa.

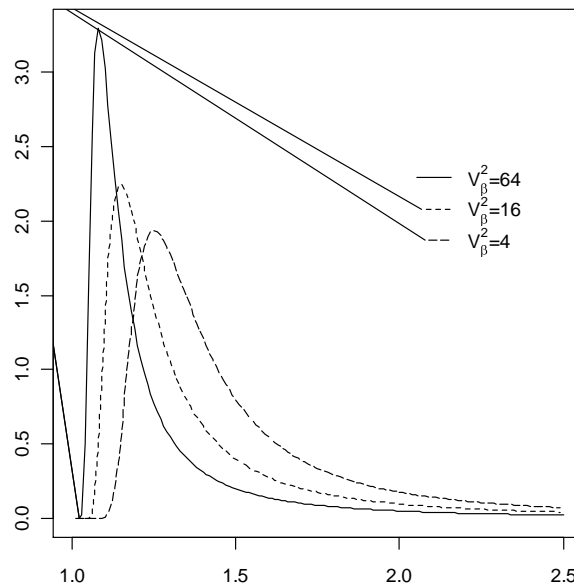


Figure 1. Distribution of optimal prices for several values of the variance of the price sensitivity parameter.

Illustration with a Non-conjugate, Uniform prior

Again assume that the manager believes that demand follows a double-log demand model from equation (1) and wants to employ a flat prior¹ over the price elasticity parameter: $p(\beta) \propto 1$. Additionally, let's assume the manager believes a priori that the optimal price must exceed some minimum price p_0 . For example, suppose that this lower bound is $p_0 = 1.35$ or that the optimal prices are not more than 10% lower than the current price of 1.50. This is consistent with the information in the previous example that $\bar{\beta} = -3$ or the optimal price is \$1.50. Conventionally analysts estimate the model with a diffuse prior and then impose the optimal pricing constraint when using the model to compute the optimal price. We refer to this as the *traditional* approach.

It would be more appropriate to represent this prior information about the optimal price as truncated, flat, improper prior for the price elasticity:

$$p(\beta) \propto \begin{cases} 1 & \text{if } \beta \geq \frac{p_0}{c - p_0} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Notice here we impose the truncation on the price elasticity parameter. Applying the change of variables formula the prior distribution for the price elasticity, the corresponding prior for the optimal price is:

$$p(p^*) = \begin{cases} \frac{c}{(c - p^*)^2} & p^* \geq p_0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

¹ Instead of a diffuse and improper prior, one could consider a proper, uniform prior in which there is some upper limit on the optimal price. As long as the upper limit of the prior distribution is large the substantive results are the same.

Notice that although the prior on β is improper and diffuse, the prior on the optimal price is not flat but asymmetric due to the nonlinear relationship between optimal price and the price elasticity².

Our purpose is to illustrate that apparently diffuse priors on the parameter space can be quite informative about the optimal price. In this case we consider a numerical simulation since a general analytic solution is difficult to obtain³. Table 1 summarizes posterior means, standard errors, and 95% credible intervals for prices and β parameter for all scenarios as well as the percentage change in the posterior means, and Figure 2 illustrates the corresponding prior, likelihood, and posterior distributions. The traditional approach follows a two-step procedure: 1) estimate the model with diffuse parameters, 2) infer the optimal price subject to the price constraint. Notice that for the case where a diffuse prior is used the mean posterior price elasticity is -3.055 and the posterior mean of the optimal price is 1.686. However, when truncating the prior distribution by using the information of the feasible set, the mean of the price elasticity changes to -2.712 which corresponds to a change of 11.2%. Notice the strong truncation induced by the prior in Figure 2 for the implied prior. The change is also noticeable for the optimal price that increases to 1.771 accounting for a change of about 6% of the optimal price. We observe the same qualitative and quantitative results for a log-linear that is available upon request. In conclusion, seemingly non-informative priors on the price elasticity combined with optimal price constraints are in truth quite informative.

² Alternatively, the analyst could start by formulating the prior on the optimal price (perhaps it is normal) and then use a change of variables to find the prior in the parameter space. Although this yields a non-conjugate prior.

³ We simulate a dataset with 100 observations with the same parameter values as the last section, e.g. the likelihood is $N(-3,1)$. The posterior sample are drawn using MCMC simulation where 100,000 draws were obtained after discarding the first 100,000 points and saving one draw every fifth iteration of the chain. We note that the posterior of the parameters can be stated directly, but that an analytic solution to the optimal price is not known (Montgomery and Bradlow 1999).

	TRADITIONAL APPROACH			IMPLIED PRIOR			%CHANGE
β	-3.055	(0.951)	[-4.973, -1.314]	-2.712	(0.702)	[-3.787,-1.264]	11.2%
p^*	1.686	(0.635)	[1.35, 4.067]	1.771	(0.682)	[1.358,4.603]	5.82%

Table 1. Posterior mean, standard error (in parentheses), and 95% credible intervals (in brackets) are given for the price elasticity and optimal price for the traditional approach (diffuse prior and price constraints imposed during inference) and the implied prior (where an equivalent truncated prior on the price elasticity is used).

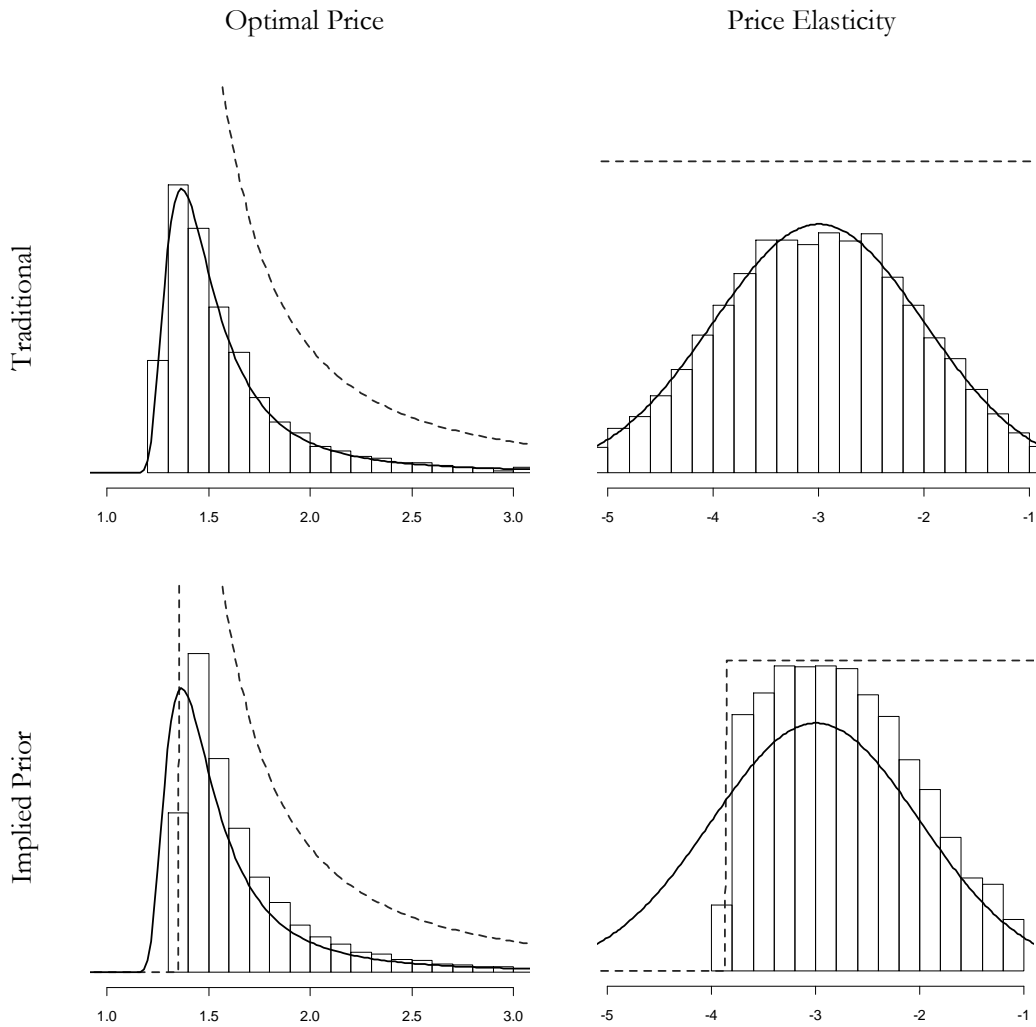


Figure 2. The histograms represent the posterior distribution, the solid lines represent the likelihood, and the dashed lines represent the prior distributions.

Discussion

Our conjecture is that analysts often have difficulty in assessing parameter values directly (Chakravarti et al. 1981), but they do have insights into properties that these parameter values should possess. Unfortunately if this information is ignored by using diffuse parameters on the parameter space, the implied prior on the optimal prices or even more generally the decision space may be quite different. If the manager were then to attempt to use this model for making decisions without the appropriate prior information it is quite likely that he would find that the results are at odds with his beliefs. The purpose of our work on informative priors is to avoid a common reaction that managers can have about decision support systems, namely that the predictions can be strange and totally at odds with their beliefs. A Bayesian method allows us to appropriately capture expert opinion, appropriately combine it with the data, and yield solutions that are logically consistent.

3. Types of Price Information for Implied Priors

In this section we consider a simple taxonomy of constraints to reflect the type of information that might be reflected in a constrained optimization problem and how they might impact our implied prior. Most likely analysts would begin with some knowledge about the parameters themselves, and this would be a starting place for the development of implied priors. Without some knowledge about the parameters implied priors may be improper which may make them difficult to work with. We expect that managers may possess information about price elasticities from prior analyses (Tellis 1988) or economic theory (Montgomery and Rossi 1999). Additionally, constraints on the signs of the price elasticities could be implemented (Boatwright et al. 1999).

In general our approach is that implied priors are to define distributions or characteristics of the distributions over functions of the parameters. For example, expected quantities, expected optimal prices, and any function of them are properly functions of the parameters. Hence, statements such as the quantity is likely to be within a specific range at a given price will yield an implied prior on the parameters. Examples of restricted quantities can be found in Deng and Yano (2006) and Essegai et al. (2002). Additionally, managers might have opinions about the general location of the optimal price solution for the category (e.g., margins should be close to 25%) which could be expressed as constraints on weighted average prices (Montgomery 1997, Khan and Jain 2005). Furthermore, managers typically offer product lines with a variety of price-quality tiers. Retailers may believe that the higher quality products should have higher prices than those of lower quality products. The inequality imposed on the ordering of optimal prices may provide valuable prior information on the parameter space.

However, that does not mean all managerial constraints reflect prior information about the parameter space. For example, a manufacturer may impose a constraint on the number of products that are sold: $q \leq q_n$. If the upper bound is the maximum amount that the manager considers as likely to occur, then the statement is implicitly describing important characteristics of the demand system and therefore should be included a priori. If the inequality is an operational constraint that expresses a capacity limitation then this does not reflect a priori information about demand and the constraint should be imposed only on the optimization stage.

The last important distinction we want to make is between conjugate and nonconjugate priors. The advantage of a conjugate prior specification is that the posterior belongs to the same family as the likelihood and yields more analytically tractable forms. When working with conjugate distributions, Kadane et al (1980) provide interactive methods to elicitate prior information from

experts. Garthwaite et al (2005) give a current review of elicitation techniques. However, conjugate families are not always suitable to accommodate prior information like the truncation induced by price or quantity constraints we analyze here. In general, we would expect that implied priors will not be conjugate.

4. Sampling Methodology

The statistics literature has proposed many methods to sample from the posterior when its distribution is not known analytically, such as the Accept-Reject method, the Metropolis-Hastings algorithm, and the Slice Sampler (Robert and Casella, 2004). Most of these techniques take advantage of the fact that the posterior distribution is proportional to the product of likelihood and the prior distribution and require an explicit expression of the prior distribution. In our setting, the constraints on the decision space require a change of variables transformation and potentially a truncation in the parameter space and therefore in general the prior is not explicitly known. However, we could easily adapt these sampling methods to discard every draw that does not satisfy our proposed managerial constraints. Specifically we can generate draws from the unconstrained distribution and then compute the optimal prices conditional on the demand parameter. Once the conditional optimal prices are obtained we can easily evaluate the feasibility of each managerial constraint and discard every draw that violates any of the constraints.

Our approach is to embed the price constraint checking procedure within a Metropolis-Hastings algorithm. When evaluating whether to accept a new proposed draw or not we need to compute the prior density evaluated at the proposed draw. In such a computation we derive the optimal prices conditional on the proposed parameter and then we check if all constraints are satisfied. If at least one of the constraints is not satisfied we return a zero value for the density at

that point making it impossible to accept that draw. A feasible starting point for the Markov chain is obtained by assuming that all cross elasticities are zero and then solving for the optimal prices⁴.

With the proposed sampling methodology we need to solve an optimization problem for each iteration of the chain. Although analytical solutions can be used in the case of the linear demand model, in general we do not have closed form expressions and we need to rely on numerical methods which can be computationally intensive. In our empirical application that we discuss in the following section we can sample several thousand of draws in a reasonable time. However, when considering more complex demand systems this may require more sophisticated sampling procedures.

5. Empirical Application

In order to better understand how optimal pricing is altered by our methodology we consider an empirical application to the refrigerated orange juice category from Dominick's Finer Foods. Montgomery (1997) provides descriptive statistics and further discussion on the data set as well as a description of the aggregation of the products⁵. Our dataset has 112 weeks of sales and 11 product aggregates which are comprised of individual SKUs which share common pricing and promotional strategies (e.g., home style, natural, or with calcium additives). We focus our analysis on a single store but conceptually the extension to a hierarchical model across all stores in the retail chain is direct. We decided against a hierarchical framework so that we could isolate the

⁴ When all cross elasticities are assumed null the demands for each product in the category are independent and easy to invert. To get a feasible starting point in the parameter space we also need to know a feasible set of prices. Although we could conduct an additional optimization problem to search for prices of minimal unfeasibility, we simply take the current prices and check they are consistent with the set of constraints we want to impose.

⁵ The data is also publicly available in the `bayesm` package for R.

contribution of the manager's knowledge from the price constraints as opposed to the exchangeable prior of the hierarchical model.

We consider weekly price decisions made in a single category as a representative problem that frequently arise in retail planning (Chintagunta et al, 2003). Following previous work in product line pricing (Reibstein and Gatignon, 1984; Montgomery, 1997; Abraham and Lodish, 1993, Mulhern and Leone, 1991; and Kalyanam, 1996), we model the demand of each product i in each week t by a semilog model⁶ ($i \in \{1, \dots, N\}$ and $t \in \{1, \dots, T\}$):

$$\ln(q_{it}) = \alpha_i + \sum_{j=1}^N \eta_{ij} p_{jt} + \xi_i f_{it} + \psi_i d_{it} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, \sigma_i^2) \quad (9)$$

where q_{it} , p_{jt} , f_{it} and d_{it} are weekly sales, price, feature, and display respectively.

We assume that the retailer jointly sets the prices p_i of each of the N brands in the category; by maximizing the expected total profits given the estimates the demand system (Montgomery 1997, Kadiyali et al., 2000 and Chintagunta et al., 2003). Thus, the retailer decides prices to maximize the following objective function:

$$\max_{\{p_1, \dots, p_N\}} \Pi = \sum_{i=1}^N (p_i - c_i) \exp \left\{ \alpha_i + \sum_j \eta_{ij} p_j + \frac{1}{2} \sigma_i^2 \right\} \quad (10)$$

where c_i is the wholesale price paid for product i by the retailer. Notice that we drop the time subscript in (10) since we consider profits for a single week without promotions. The solution of (10) yields a vector of optimal prices, which we denote as \mathbf{p}^* , whose i th element is p_i .

We stack the set of parameters for the i th product, $\{\alpha_i, \eta_{i1}, \dots, \eta_{iN}, \xi_i, \psi_i\}$, into a vector $\boldsymbol{\theta}_i$ and employ the following priors:

$$\sigma_i^2 \sim \chi^{-2}(100, 1) \quad (11)$$

⁶ Another popular model in the literature is the double-log model (Kopalle et al, 1999; Montgomery and Bradlow, 1999). Empirical results for such specification are similar to the ones we present here and are available upon request.

$$\boldsymbol{\theta}_i | \sigma_i^2 \sim N(\mathbf{0}, \sigma_i^2 V_i) \cdot \mathbf{I}[\{p_{0i}^* \leq p_i^* \leq p_{li}^*\}] \cdot \mathbf{I}[\max(\mathbf{p}_A^*) \leq \min(\mathbf{p}_B^*), \max(\mathbf{p}_B^*) \leq \min(\mathbf{p}_C^*)] \quad (12)$$

Furthermore, we parameterize the covariance matrix as $V_i = \kappa I$ where κ is a scalar parameter that controls how informative the prior is.

Notice that the prior for our parameter vector given in (12) would be diffuse if not for the truncation imposed by our two pricing constraints. The first truncation is a direct constraint on the lower and upper bounds of each optimal price: $p_{0i}^* \leq p_i^* \leq p_{li}^*$. Obviously the narrower the feasible set then the more informative our prior will be. The actual values of the lower and upper bounds will depend on manager, but we assume that the optimal price is within 25% of the current price. Second, we assume that the optimal price for lower price-quality tiers must be lower than higher price-quality tiers. In particular we impose the constraint that the prices of the premium national brands (denoted by the vector \mathbf{p}_C^*) have to be greater than the prices of the regular national brands (\mathbf{p}_B^*), and that prices of regular national brands have to be greater than the price of the store brands (\mathbf{p}_A^*). (Potentially, managers could even implement a stronger constraint on the ordering of the optimal prices across individual products instead of groups.) Our contention is that both of these constraints would be readily accepted by a supermarket manager. These constraints would seem to be conservative assumptions, since both allow fairly wide latitude of price movements.

To estimate our model we use a random-walk Metropolis-Hastings algorithm where the parameters are updated in blocks: intercepts α , variance of the error term in the demand function σ^2 , price coefficients η and feature and display parameters ξ and ψ . Moreover, we update own price coefficient parameters and cross price coefficient in different steps. Let Ω be the price set that is considered feasible. Therefore, the distributions we use to compute acceptance probabilities of each proposed draws are given by:

$$p(\boldsymbol{\theta}, \boldsymbol{\sigma} | \text{data}_i) \propto \begin{cases} \prod_i p(\boldsymbol{\theta}_i, \sigma_i^2 | \text{data}_i) & \text{if } p^*(\boldsymbol{\theta}, \boldsymbol{\sigma}) \in \Omega \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

Following Gelman et al (2003) we tuned proposal distributions to get acceptance rates around 0.23. Adaptive jumping strategies showed no significant improvement over the static strategy we used here.

Understanding the Prior Distribution

Typically interest focuses upon the posterior distribution or the estimates from this distribution such as the posterior mean. In our case we are also interested in understanding the prior. Since our prior has potentially highly non-linear dependences on the parameters through the optimal price constraints that need to be found numerically we need to simulate from our prior. However, the probability mass of our prior is highly dispersed in the parameter space, and our rejection sampler has a poor acceptance rate making it difficult to get a sufficiently large sample in a reasonable amount of time. In our estimation problem we do not have this difficulty since we can draw values conditional on the likelihood.

Given our purpose here is to illustrate how the price constraints modify the prior specification, we instead consider a linear demand model since we have closed form solutions for the optimal values. Exploiting the closed form of the optimality conditions, the assumption of optimal prices uniformly distributed in the feasible space, we can create a simple Gibbs-sampler algorithm to draw from the prior. For example, suppose we have to decide simultaneously the optimal price for N products and that the demand function for each product is characterized by K parameters. Then, to move to the next feasible point we sample $N(K-1)$ elements from the prior

and we chose the remaining N components of the vector of parameter to exactly satisfy all N first order conditions derived from the maximization problem that characterize the optimal prices.

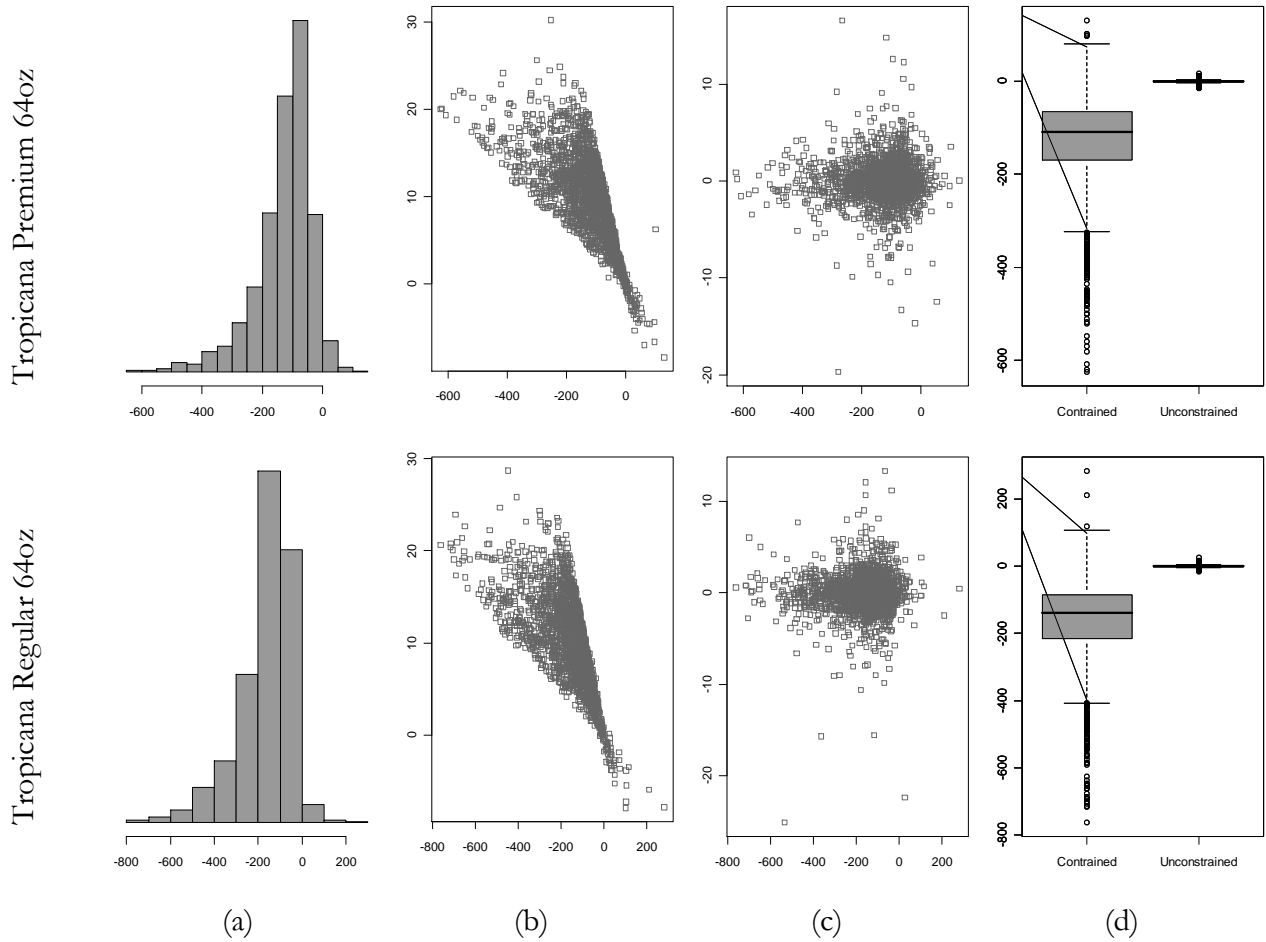


Figure 3. Prior distribution when considering price constraints (a) marginal distribution of own price coefficients, (b) joint distribution between own price coefficient and intercept, (c) joint distribution between own price coefficient and other brand price coefficient and (d) boxplots of own price coefficient with (left) and without (right) considering the information of the bounds on the optimal prices.

Using this sampling approach, we generate 2,000 draws for both cases: proper prior with and without implied information induced by price constraints. Examples of the resulting marginal distribution, joint distributions and mean comparison with and without considering price constraints

in the prior definitions are depicted in Figure 3 for parameters associated with two selected products.

Notice that the resulting prior on the own price coefficient is not normally distributed anymore and it is significantly skewed. When looking at the joint properties of the prior, we observe truncations on the distribution of the intercept and own-price elasticities: for a given intercept, the upper (lower) bound for the optimal price imposes a corresponding upper (lower) bound for the price coefficient. For example, in our implied prior specification the intercept cannot exceed a threshold for a given own price elasticity level because it would imply an optimal price larger than the upper bound we impose. For the cross price elasticities we observe a star shape instead of elliptical ones derived from normal conjugate prior. With respect to the case where price constraints are not considered this shows a significant shift in the own-price coefficients suggesting a more elastic demand curve. The implied prior specification also exhibits higher variability which can be explained by the fact that the truncation occurs close to the mode(s) of the distribution forcing it to concentrate probability mass on the tail of the distribution.

Posterior Estimates

For comparison we estimate⁷ our model using the traditional approach and with our implied prior. For the traditional case without our implied prior we enforce the constraints only during optimization. Tables 2 and 3 provide the posterior mean and standard deviations of the parameters with and without our implied prior, respectively. We can see that several parameters are significantly different. A major effect is observed in terms of the intercept and the own price elasticities. In

⁷ We create 20,000 draws for both cases: 1) when the information contained in the price constraints are translated into the prior specification and 2) when these constrained are not considered in the estimation stage and they are only imposed in the optimization phase. We thin the draws to every 10th iteration due to our highly auto-correlated sequence of draws.

Figure 4 we display boxplots of the parameters for a sub-set of the products. Here we observe that the negative shift on the own price coefficient parameters for most of the brands. Several of these parameters are significant, but at the same time we observe that the intercept estimates are larger when the price constraint information is considered. These results are consistent with our qualitative description of the prior where we cannot have rather price insensitive brands with large intercepts because they would imply an optimal price beyond the bounds the manager would consider likely in his prior assessment. The feature and display coefficients show little difference in their parameter estimates. In terms of the cross price elasticity parameters, we observe minor differences in terms of the location but a substantial reduction in their variances.

Although we observe significant differences in the estimates we do not see substantial reductions in the fit of the model. The average correlation of actual sales and the ones predicted by a linear model evaluated at the mean of the posterior sample yield a mean R^2 of 0.411 (0.142) for the implied case and R^2_{uncon} of 0.435 (0.107) for the traditional, diffuse prior. This suggests that the high dimensional space of the price elasticity problem is relatively flat and yields many solutions that may be good or better estimators if prior information is properly used.

Product	Price Elasticity Matrix											Other Parameters		
	Trop 64	Trop 96	FlNat 64	TropR 64	MM 64	MM 96	CHill 64	TFrsh 64	FGold 64	Dom 64	Dom 128	Const. α	Feat. ξ	Deal ψ
Tropicana Prem. 64 oz.	-1.7 (0.94)	0.16 (0.28)	-0.34 (0.24)	-0.28 (0.21)	-0.44 (0.14)	-0.4 (0.29)	-0.31 (0.17)	0.17 (0.35)	0.028 (0.2)	-0.20 (0.13)	-0.64 (0.18)	13.04 (1.31)	0.44 (1.0)	0.073 (0.69)
Tropicana Prem. 96 oz.	-0.87 (0.36)	-1.2 (1.1)	0.17 (0.23)	-0.5 (0.3)	-0.20 (0.13)	-0.12 (0.2)	0.32 (0.16)	0.22 (0.14)	-0.33 (0.28)	-0.15 (0.19)	0.072 (0.14)	11.22 (1.74)	0.40 (1.2)	0.14 (0.9)
Florida's Natural 64	-0.16 (0.29)	-0.65 (0.21)	-0.83 (1.1)	0.66 (0.27)	-0.52 (0.48)	-0.0037 (0.24)	0.012 (0.17)	-0.15 (0.21)	0.021 (0.15)	-0.06 (0.15)	0.11 (0.22)	9.17 (0.98)	0.76 (1.2)	0.23 (0.77)
Tropicana 64 oz.	0.19 (0.29)	-0.7 (0.34)	-0.28 (0.25)	-1.6 (0.97)	-0.12 (0.22)	-0.16 (0.21)	-0.38 (0.24)	-0.41 (0.27)	-0.28 (0.20)	-0.32 (0.18)	0.3 (0.13)	12.09 (0.77)	1.1 (0.95)	0.17 (0.92)
Minute Maid 64 oz.	-0.39 (0.24)	0.14 (0.37)	0.49 (0.20)	-0.31 (0.14)	-1.1 (0.91)	-0.44 (0.19)	-0.30 (0.15)	-0.28 (0.34)	0.37 (0.22)	-0.3 (0.11)	0.33 (0.19)	10.37 (1.06)	0.95 (0.92)	0.47 (0.8)
Minute Maid 96 oz.	0.41 (0.24)	-0.005 (0.33)	-0.37 (0.35)	0.0027 (0.16)	0.57 (0.23)	-1.5 (0.88)	-0.15 (0.25)	-0.22 (0.18)	0.13 (0.19)	-0.19 (0.20)	-0.43 (0.16)	9.73 (1.19)	0.23 (1.0)	0.073 (0.74)
Citrus-Hill 64 oz.	-0.7 (0.34)	0.32 (0.32)	-0.88 (0.49)	0.091 (0.18)	-0.71 (0.2)	-0.9 (0.52)	-0.031 (0.93)	-0.23 (0.22)	0.4 (0.33)	0.006 (0.19)	0.14 (0.19)	9.98 (1.83)	1.1 (1.1)	0.38 (0.77)
Tree Fresh 64 oz.	-0.005 (0.27)	-0.47 (0.24)	-0.68 (0.52)	0.075 (0.13)	-0.025 (0.20)	-0.026 (0.38)	-0.25 (0.22)	-0.61 (0.86)	-0.18 (0.26)	-0.15 (0.31)	-0.64 (0.41)	9.99 (2.11)	0.78 (1.1)	-0.017 (0.73)
Florida Gold 64 oz.	0.079 (0.46)	-0.034 (0.46)	-0.47 (0.28)	-0.21 (0.20)	-0.74 (0.27)	-0.36 (0.33)	0.17 (0.21)	0.44 (0.27)	-1.4 (0.92)	-0.13 (0.31)	-0.10 (0.16)	9.62 (1.66)	1.1 (1.0)	0.43 (0.76)
Dominicks 64 oz.	0.37 (0.20)	-1.1 (0.83)	0.011 (0.19)	0.12 (0.24)	-0.21 (0.22)	-0.095 (0.14)	-0.4 (0.34)	-0.58 (0.42)	-0.14 (0.20)	-1.2 (0.57)	-0.07 (0.14)	12.57 (2.27)	0.64 (0.91)	0.29 (0.81)
Dominicks 128 oz.	-0.66 (0.46)	-0.73 (0.39)	-0.18 (0.26)	-0.096 (0.22)	-0.017 (0.23)	0.05 (0.20)	-0.27 (0.20)	0.33 (0.31)	0.049 (0.19)	-0.42 (0.32)	-1.1 (0.92)	11.05 (1.87)	0.30 (1.1)	0.26 (0.88)

Table 2. Parameter estimates when considering implied priors with price constraints in the prior specification

Product	Price Elasticity Matrix											Other Parameters		
	Trop 64	Trop 96	FlNat 64	TropR 64	MM 64	MM 96	CHill 64	TFrsh 64	FGold 64	Dom 64	Dom 128	Const. α	Feat. ξ	Deal ψ
Tropicana Prem. 64 oz.	-0.3 (0.72)	0.011 (0.81)	-0.001 (0.73)	0.030 (0.57)	-0.008 (0.57)	0.0097 (0.71)	0.01 (0.61)	-0.003 (0.58)	-0.02 (0.53)	0.021 (0.44)	-0.01 (0.49)	9.11 (2.06)	0.65 (1.0)	0.46 (0.8)
Tropicana Prem. 96 oz.	0.015 (0.7)	-0.17 (0.8)	-0.001 (0.73)	-0.018 (0.58)	-0.015 (0.55)	-0.025 (0.7)	-0.03 (0.6)	-0.02 (0.57)	-0.02 (0.53)	0.01 (0.43)	-0.02 (0.49)	9.103 (1.99)	0.33 (1.2)	0.23 (0.89)
Florida's Natural 64	0.067 (0.72)	0.021 (0.83)	-0.18 (0.74)	0.03 (0.56)	-0.005 (0.57)	-0.017 (0.69)	-0.04 (0.61)	0.016 (0.57)	0.001 (0.54)	-0.01 (0.44)	-0.02 (0.49)	7.76 (2.00)	0.83 (1.2)	0.23 (0.82)
Tropicana 64 oz.	0.12 (0.72)	0.086 (0.81)	-0.004 (0.73)	-0.28 (0.6)	0.041 (0.56)	-0.009 (0.71)	-0.01 (0.6)	-0.02 (0.58)	-0.02 (0.53)	0.005 (0.43)	-0.01 (0.5)	8.466 (2.03)	0.9 (1.0)	0.52 (0.98)
Minute Maid 64 oz.	-0.010 (0.7)	-0.019 (0.83)	0.034 (0.71)	0.022 (0.57)	-0.18 (0.59)	-0.037 (0.72)	-0.04 (0.6)	-0.03 (0.56)	0.006 (0.52)	0.044 (0.43)	-0.01 (0.49)	8.735 (2.03)	0.96 (0.99)	0.61 (0.86)
Minute Maid 96 oz.	-0.003 (0.7)	0.038 (0.8)	0.0058 (0.74)	-0.018 (0.57)	-0.04 (0.58)	-0.15 (0.7)	0.004 (0.6)	-0.04 (0.56)	-0.003 (0.53)	-0.01 (0.44)	0.002 (0.49)	8.031 (1.96)	0.28 (1.1)	0.19 (0.83)
Citrus-Hill 64 oz.	0.054 (0.68)	-0.03 (0.83)	0.011 (0.72)	-0.001 (0.57)	0.037 (0.57)	-0.002 (0.72)	-0.15 (0.62)	-0.004 (0.57)	-0.02 (0.52)	-0.01 (0.44)	-0.01 (0.5)	7.641 (1.996)	1.1 (1.2)	0.46 (0.83)
Tree Fresh 64 oz.	0.0008 (0.7)	0.032 (0.8)	-0.029 (0.73)	0.0055 (0.57)	0.0081 (0.57)	-0.012 (0.71)	0.018 (0.59)	-0.13 (0.57)	0.0022 (0.53)	0.041 (0.43)	0.012 (0.5)	7.31 (2.03)	0.6 (1.3)	0.15 (0.8)
Florida Gold 64 oz.	0.024 (0.7)	-0.045 (0.82)	0.013 (0.73)	-0.059 (0.58)	0.014 (0.56)	-0.064 (0.72)	-0.02 (0.59)	0.00 (0.57)	-0.22 (0.54)	0.063 (0.44)	0.007 (0.49)	6.885 (2.04)	1.2 (1.1)	0.66 (0.82)
Dominicks 64 oz.	0.025 (0.7)	-0.086 (0.82)	0.023 (0.74)	0.10 (0.59)	0.055 (0.57)	-0.005 (0.72)	-0.06 (0.6)	-0.007 (0.57)	0.0037 (0.52)	-0.2 (0.44)	0.002 (0.5)	8.5 (2.07)	0.64 (1)	0.86 (0.88)
Dominicks 128 oz.	-0.041 (0.72)	-0.038 (0.8)	-0.013 (0.74)	-0.024 (0.57)	-0.024 (0.57)	-0.028 (0.7)	-0.02 (0.61)	-0.03 (0.59)	-0.019 (0.54)	0.033 (0.43)	-0.1 (0.5)	8.5 (2.04)	0.38 (1.2)	0.32 (0.95)

Table 3. Parameter estimates from the traditional approach without considering price constraints in the prior specification

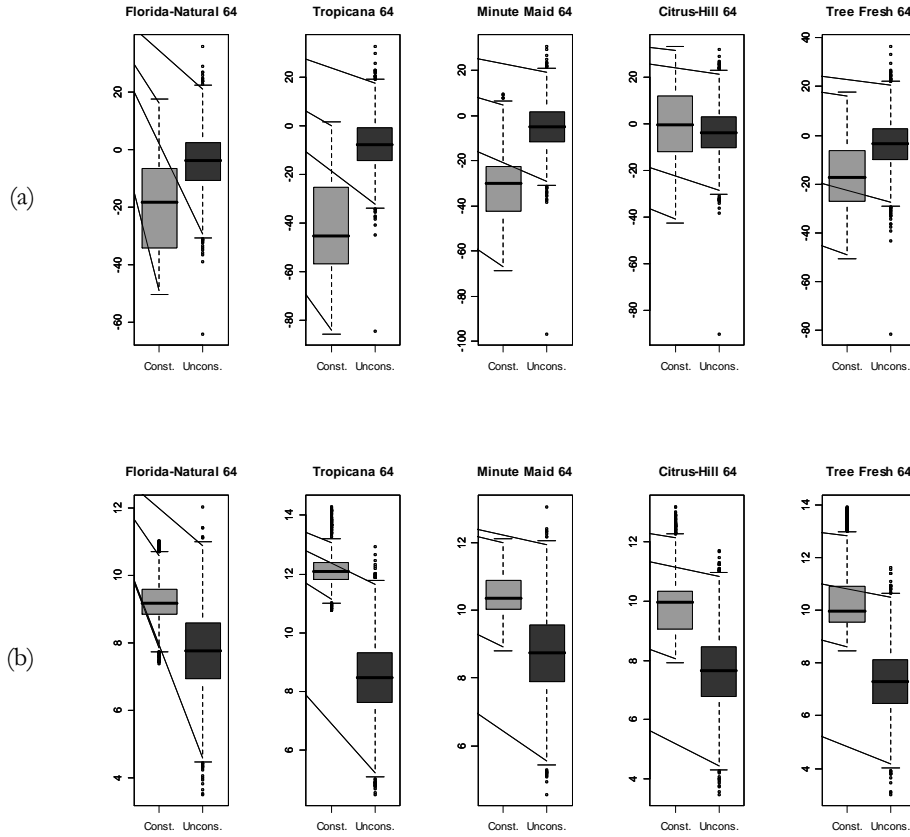


Figure 4. Boxplots of parameter estimates with and without considering price constraints on the definition of prior specification (a) Own price coefficient. (b) Intercept

From a managerial perspective, the estimation of the parameters of the demand model is only an intermediate step, and therefore we are ultimately interested in the impact on optimal prices and revenues. In Table 4 we report the posterior mean, standard errors and 95% credible set for all products derived for both the traditional and the implied prior approaches. We also report the percentage difference in posterior mean for these two approaches. To illustrate this table consider Minute Maid 64 oz. The price bounds are (0.0272, 0.0453) per ounce or (\$1.74, \$2.90) per carton with the original price being midway between these bounds at 0.03625 per ounce or \$2.32. The

mean of the optimal price for the traditional approach is 0.0409 per ounce or \$2.62 per carton versus the result from our implied prior of 0.0360 per ounce or \$2.30 per carton, or a 12% increase in the point estimate. Notice that the 95% credible interval is truncated for the traditional method at .0453 per ounce, but not for our implied prior approach.

TIER	PRODUCT	TRADITIONAL		IMPLIED PRIOR		BOUNDS (Lower, Upper)	% CHANGE	
		Mean (Std.Err.)	95% CI	Mean (Std.Err.)	95% CI			
Premium	Tropicana Prem. 64 oz.	0.0526 (0.00330)	[0.0474, 0.0579]	0.0515 (0.00463)	[0.0418, 0.0577]	(0.0347, 0.0579)	-2.1%	
	Tropicana Prem. 96 oz.	0.0573 (0.00348)	[0.0520, 0.0639]	0.0559 (0.00516)	[0.0467, 0.0636]	(0.0384, 0.0639)	-2.4%	
National	Florida's Natural 64	0.0482 (0.00243)	[0.0448, 0.0542]	0.0466 (0.00357)	[0.0406, 0.0542]	(0.0343, 0.0571)	-3.3%	
	Tropicana 64 oz.	0.0408 (0.00275)	[0.0364, 0.0457]	0.0379 (0.00330)	[0.0320, 0.0446]	(0.0274, 0.0457)	-7.1%	
	Minute Maid 64 oz.	0.0409 (0.00282)	[0.0362, 0.0453]	0.0360 (0.00292)	[0.0304, 0.0432]	(0.0272, 0.0453)	-12.0%	
	Minute Maid 96 oz.	0.0462 (0.00218)	[0.0417, 0.0508]	0.0387 (0.00353)	[0.0334, 0.0457]	(0.0332, 0.0554)	-16.2%	
	Citrus-Hill 64 oz.	0.0398 (0.00243)	[0.0354, 0.0456]	0.0376 (0.00345)	[0.0307, 0.0449]	(0.0286, 0.0476)	-5.5%	
	Tree Fresh 64 oz.	0.0370 (0.00226)	[0.0334, 0.0427]	0.0375 (0.00321)	[0.0313, 0.0436]	(0.0266, 0.0443)	1.4%	
	Florida Gold 64 oz.	0.0350 (0.00204)	[0.0320, 0.0400]	0.0366 (0.00349)	[0.0305, 0.0416]	(0.0251, 0.0418)	4.6%	
	Store	Dominicks 64 oz.	0.0321 (0.00222)	[0.0278, 0.0350]	0.0276 (0.00288)	[0.0219, 0.0333]	(0.0210, 0.0350)	-14.0%
		Dominicks 128 oz.	0.0335 (0.00193)	[0.0294, 0.0374]	0.0309 (0.00262)	[0.0253, 0.0358]	(0.0232, 0.0386)	-7.8%
		TOTAL PROFITS	11074 (2372)	[7468, 16580]	7073 (1135)	[5277, 9909]		-36.1%

Table 4. Optimal prices for each product (mean, standard errors and 95% credible intervals) for the Traditional Approach and the Implied Prior approach. The upper and lower price bounds for the optimal price constraints are given in the *Bounds* column. The last column of the table reports the percentage differences in the means between the Traditional and Implied Prior Approaches. The final row reports the Total Category Profits.

More generally, the price constraints are binding (e.g., the credible interval is truncated by the upper bounds) for five products (Tropicana Premium 64 oz, Tropicana Premium 96 oz,

Tropicana 64 oz, Minute Maid 64 oz, and Dominicks 64 oz), while the price constraints are close for several of the products using the Implied Prior, none truncate the 95% credible interval. For nine of the products our implied prior approach suggests lower optimal prices than the traditional approach, and for several of these products the optimal price differences are quite considerable. The presence of some negative differences illustrates the complexity of the implied prior when several constraints are imposed and the nontrivial implications for pricing decisions.

Figure 5 compares the distribution of optimal prices for an arbitrary sub-set of 4 products. The upper right cells plot the optimal prices derived from the standard approach (dark) and the lower left cells plot the optimal prices derived from our proposed methodology (light). Along the diagonal are the histograms of the posterior distribution of optimal prices derived from both methods. Visual inspection of the plot provides us with interesting insights. First, we ascertain that the distributions of the optimal prices are indeed quite different. Under the traditional methodology our estimates suggest that demanded quantities are rather inelastic with respect to prices and recommend increasing the price to higher levels. However, when implied priors are used we see that demand is more elastic which implies lower prices. The optimal prices derived from our proposed prior seems to be more consistent with what is observed in practice, but still shows that the retailer is not at the optimal price but closer to it.

It is also important to observe that while in the standard approach the price constraints for optimal profits are frequently binding, in our proposed Bayesian method most of the probability mass is allocated to the interior of the feasible optimal price set. The traditional approach tends to result in optimal pricing solutions that lie along the boundary. Therefore, it is easy for the analyst to directly manipulate the pricing solutions in an ad hoc manner. A property of the Bayesian solution is that optimal prices tend to occur within the interior. Therefore it is more difficult for the analyst

to manipulate the final pricing solution through binding constraints since the data will still exert an influence along with the implied to yield an appropriately weighted posterior. In effect, the optimal prices are less sensitive to the constraints. We believe this is an attractive property of the Bayesian procedure.

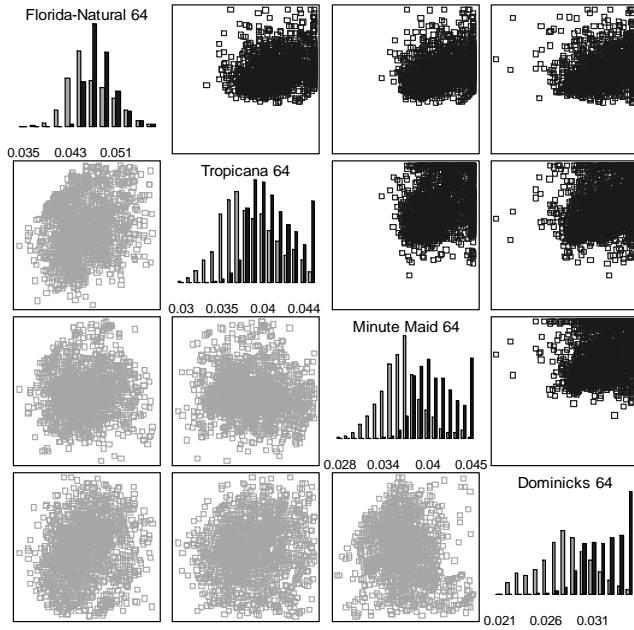


Figure 5. Comparison of optimal prices

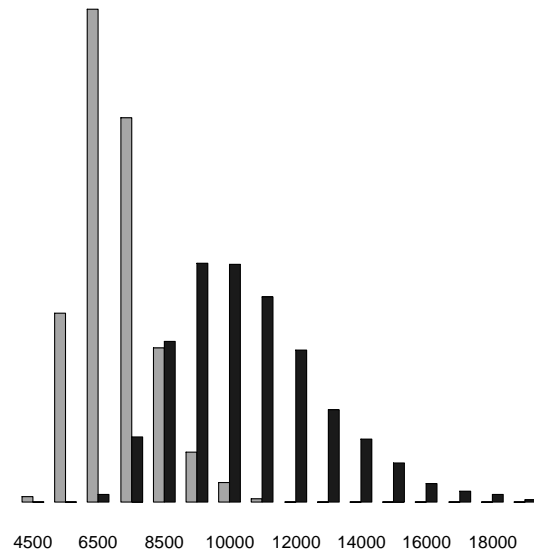


Figure 6. Comparison of optimal revenues.

In Figure 6 we compare the distribution of the optimal profits derived from the standard approach (dark) and those derived from our proposed methodology (light). As a direct consequence of lower prices we end up with lower expected profits. In some sense, the introduction of managerial knowledge provides a reality correction of the benefits that can be achieved by conducting regression analysis and determining prices based on those results.

6. Discussion and Conclusions

This paper explores the impact on the parameter estimates of a demand model by introducing the idea that price constraints reflect prior information on the part of the manager about price response. We argue that this information is useful in estimating the demand model, and can have an important impact about our posterior inferences of the price elasticities and subsequently the optimal prices. In the traditional demand modeling task, constraints on optimal prices are considered in an ad hoc manner, which means that information is not properly weighted within a decision theoretic framework. Practically our approach means that instead of constraining the decision space as is done in the traditional approach, one should instead truncate the parameter space in the prior distribution to yield better predictions that do not need to be constrained. We believe that this is a simple idea, but a practical one that can dramatically improve the pricing solutions derived.

The improvement in the price decision making can be clearly observed in the difference between the prices of the traditional and the Bayesian method. If one accepts the information in the prior distribution then it is well known that Bayes rules that correspond with proper prior distributions are admissible. If one were to follow the recommendations from the traditional method, our “optimal” prices would actually be suboptimal. We point out that in general our estimators are statistically biased, although this bias reflects the information interjected by the analyst and improves our optimal price estimators.

The implied information from optimal price constraints creates potentially complex, non-linear relationships amongst the parameters. This imposes a difficult and complex change-of-variables problem. We propose a simple method to include this information in a rejection sampling

approach that could potentially been used to a wide range of managerial decision processes. Although this structure is flexible it does add a computational burden to the estimation process.

Under our methodology, optimal prices are lower and more consistent with a manager's actual decisions. Little and Shapiro (1980) realized that price elasticity parameters were not consistent with the prices observed in the retail environment. They pointed out that stores act as if the customer were considerably more sensitive to prices than purchase data seem to indicate. Montgomery and Bradlow (1999) suggested uncertainty about functional form and Fox et al (2009) suggest endogenous store traffic can also explain this phenomenon. Our results provide another explanation to this old problem of why it appears that managers are overpricing. Namely, that price elasticities appear to be inelastic if constraints on optimal prices are imposed ad hoc instead of being imposed a priori.

We consider these results promising and hope it motivates future researchers to consider the information induced by managers who impose price constraints in the optimization phase. There are two future directions that we think are quite important. First we think more efficient sampling techniques should be developed. Second, it is possible that managerial constraints may be inconsistent with one another or even at odds with the data. We have not said how managers create their beliefs or even whether they are correct. Potentially, analysts may want to attach probabilities that the constraints are "true". Implicitly our implied priors dogmatically impose that the constraints are true. One could imagine a case where the analyst wishes to learn whether the data is consistent with constraints. In conclusion, we hope these results will encourage practitioners and researchers to more carefully articulate their prior beliefs, and express them a priori and not ad hoc.

Chapter 3

Cross-Market Discounts

1. Introduction

Giant Eagle is a dominant grocery chain in the Pittsburgh, PA metropolitan area. In contrast, the gasoline market in Pittsburgh has many players and is highly competitive⁸. However, one of the gas station chains, GetGo, is owned by Giant Eagle. A few years ago, Giant Eagle started the “fuelperks!” program under which a consumer, upon purchasing \$50 of groceries from Giant Eagle, earns 10¢ off per gallon of gasoline purchased from GetGo at the next purchase occasion. This per-gallon discount increases in a (stepwise) linear fashion: \$50 gets 10¢ off, \$100 gets 20¢ off and so on, and if consumers spend enough money on groceries, they can even purchase fuel from GetGo for free.

fuelperks! has been a tremendously successful promotional program for the Giant Eagle-GetGo combine, leading to a significant increase in sales and profits (PG 2006). The other gas station operators in Pittsburgh, on the other hand, have been hit hard, to the extent that they jointly filed (and lost) a lawsuit against Giant Eagle, in which they accused it of employing unfair sales practices (PG 2005).

⁸ Giant Eagle has more than 50% share of the grocery market, with the rest distributed between Walmart, Shop n' Save and some other smaller retailers (Linderman 2007). In contrast, in the gasoline market, many firms such as Exxon, Shell, Sunoco, British Petroleum, Gulf, GetGo, CitGo and CoGo have a sizeable presence and no firm is dominant.

Attracted by the success of fuelperks! in Pittsburgh, other grocery retailers and gas station operators across USA are quickly adopting this promotion strategy. In fact, Giant Eagle has started a new company, Excentus Corporation, through which it is implementing the fuelperks! program for retailers across USA: Giant Eagle in Pennsylvania, Ohio and Maryland, BI-LO in Georgia, North Carolina, South Carolina and Tennessee, Roundy's in Minnesota and Wisconsin, Winn-Dixie in Florida, Alabama, Louisiana, Georgia and Mississippi, and Ukrop in Virginia. Although most of these grocery retailers do not own their own fuel pumps, they have set up their programs jointly with local fuel pumps, e.g., BI-LO with Sunoco, Roundy's and Ukrop with British Petroleum, etc. Similarly, Safeway, which has stores located throughout western and central USA and western Canada, Genuardi's, which is run by Safeway in Pennsylvania, New Jersey and Delaware, and Tom Thumb, which is run by Safeway in Texas, have recently started their own independent "Power Pump" rewards programs, which are very similar to the fuelperks! program, in partnership with British Petroleum⁹. Shaw's has introduced a fuelperks!-type scheme in its stores throughout Maine, Massachusetts, New Hampshire, Rhode Island, Connecticut and Vermont, and Kroger has also introduced a similar program throughout its stores in the Midwestern and Southern United States¹⁰. Clearly, the idea of grocery retailers offering discounts on gasoline has quickly caught on all over North America. Moreover, supermarket chains in the U.K., France and Australia have also recently started similar promotional programs (Gans and King 2004).

⁹ Interestingly, Excentus, which has patented the fuelperks! idea, filed and won a patent-infringement lawsuit against Safeway in January 2010 (*Excentus* 2010). At the time of writing this manuscript, Safeway had temporarily halted its Power Pump program while it worked out a solution with Excentus.

¹⁰ Some details regarding these programs are available at the following web sites:

Fuelperks (<http://www.fuelperks.com>)

Power Pump (<http://www.safeway.com/IFL/Grocery/PowerPump-Locator>)

Shaw's promotion: (<http://www.shaws.com/pages/promotionsIrving.php>)

Kroger promotion: (<http://www.kroger.com/in\ store/fuel/pages/default.aspx>)

At its core, fuelperks! is a "cross-market discount" program --- purchases in one market lead to a discount in an otherwise unrelated market. This is a new and intriguing addition to the vast array of promotional tools that marketing managers have at their disposal, and its widespread success in the grocery-fuel combine raises a number of interesting questions regarding such cross-market discounts in general. Why are retailers adopting them *en masse* at such a rapid rate? How do they influence prices and consumption in both markets and where do the profits come from? What is the impact of competition on the rate of the cross-market discount offered, and what are the incentives of competitors to introduce similar programs? Which market pairs are ideal for implementing such a scheme? Are such discounts good or bad for consumers? Finally, how does this compare with other similar-looking promotion strategies such as bundling, loss leadership and quantity discounts?

Suppose one "parent firm" owns a grocery store and a fuel pump¹¹, and consumers derive higher utility from higher consumption in both markets, but have diminishing marginal returns from consumption. We start with a simple scenario in which both the grocery store and the fuel pump are monopolies in their respective markets. Upon purchasing groceries, consumers get lower per-unit prices for fuel and the size of this discount depends on the quantity purchased of groceries. In this scenario, we find that price decreases and sales increase in the fuel market, as expected. However, the price in the grocery market rises above even the monopoly price that was already being charged before the discount was introduced. Moreover, grocery sales also increase. In other words, due to a "cross-market leverage" effect, both price and sales for groceries increase *together* (in spite of the fact that, in the demand function, demand is assumed to be reducing in price). Overall, the total profits of the parent firm increase as a larger cross-market discount is offered and, under certain conditions,

¹¹ For ease of exposition, we use grocery and fuel as the two markets to discuss the insights in the introduction, but our modeling studies the general case. We also provide some other examples in the conclusion section.

the parent firm might even sell fuel at a loss to achieve this. On the other hand, we find that consumers are worse off --- since there is no competition, the firm can set the cross-market discount to extract the full consumer surplus across the two markets.

Effectively, a cross-market discount implements a nonlinear pricing schedule across market boundaries by allowing consumers to purchase in one market and avail a discount in a different market. This motivates a comparison to a nonlinear pricing schedule within a market, i.e., a quantity discount-type scheme in which the unit price of a product decreases as more of it is purchased. Our analysis shows that under diminishing marginal returns from consumption, firms will typically find a cross-market discount more profitable than a similar quantity discount in the same market. A cross-market discount distributes the additional consumption (motivated by the price discount) across both the markets rather than motivating more consumption in the same market. This delays the point at which reduced marginal utility from additional consumption sets in for both markets, which therefore leads to increased total consumption. In other words, besides the posted prices in the two markets, this pricing strategy introduces a "third lever to pull," in the form of the cross-market discount rate, which allows the firm to exploit the less price sensitive portions of the consumers' utility functions.

Two features of our model are important for our results --- first, price discounts motivate greater purchasing and consumption, and second, marginal utility from consumption decreases as consumption increases. Both assumptions are common in the Economics and Marketing literature, and we believe that they are also appropriate in our setting. First, it is well established that price discounts lead to greater purchasing and, moreover, greater purchasing also endogenously leads to

greater consumption (Ailawadi and Neslin 1998, Chandon and Wansink 2002, Sun 2005, Wansink 1996). Moreover, to further induce greater purchasing on the grocery side, grocery retailers have come up with the novel idea of heavily promoting gift cards of *other* non-grocery retailers. For instance, the advertisements and the web sites for the fuelperks! and Power Pump programs prominently display gift cards for other retailers such as Best Buy, Macy's, Toys"R"Us, The Home Depot, the iTunes store, etc., on them. Second, it is well accepted that marginal utility from consumption decreases as consumption increases, and this assumption is often invoked in analytical models (Mas-Colell, Whinston and Green 1995, Singh and Vives 1987, Varian 1992). A primary reason is that consumers typically get satiated from increased consumption of a good (Brickman and Campbell 1971, Coombs and Avrunin 1977). They also face increasing holding and storage costs upon purchasing larger quantities of any good (Bell, Iyer and Padmanabhan 2002, Sun 2005, Wansink 1996). For instance, it can be very effort intensive and costly for consumers to store excess gasoline after they fill up the fuel tank of their car.

Next, we bring competition into the picture by analyzing a scenario in which the grocery market is a monopoly but the fuel market is a competitive duopoly, and the parent firm owns the grocery store and one fuel pump. (We call the parent firm's fuel pump as "fuel pump 1" and the competitor's fuel pump as "fuel pump 2") This is a closer representation of our motivating example from the Pittsburgh market. Suppose a fuelperks! scheme is introduced under which consumers can avail a discount at fuel pump 1 upon purchasing groceries. In this case, if the competitive intensity in the fuel market is low, the insights derived without competition largely carry over, while if this intensity is high, there are significant new dynamics at play. Most notably, due to a strong pricing reaction from fuel pump 2, the total profit of the parent firm has an inverted-U shape in the rate of cross-

market discount, unlike in the previous scenario. And what is the impact of the intensity of competition in the fuel market on the optimal rate of the cross-market discount? Will this rate be higher (to attract more consumers to the joint offering) or lower (to avoid a strong reaction from fuel pump 2)? From our model, we find that the second effect dominates --- to avoid dissipating its profits by inviting a stronger pricing reaction from the competitor, the parent firm reduces the rate of the cross-market discount as the degree of competition in the fuel market increases.

We further extend our analysis to a scenario in which both the grocery and fuel markets are competitive. This raises an interesting possibility --- the second grocery store can form a strategic alliance with the second fuel pump and also offer a competing cross-market discount. We find that these two players will indeed always make this joint offering, but the implications for profits and consumer surplus vary significantly with the intensity of competition. If the intensity of competition in both markets is low, then both partnerships see increased profits from cross-market discounts, while consumers are worse off. If the intensity of competition is high, both partnerships are in a “prisoners' dilemma” situation, i.e., both lose profit from cross-market discounts but both offer them, and consumer surplus increases. However, when the intensity of competition is in a medium range, then both partnerships see increased profits from this scheme *and* consumers are also better off. Therefore, cross-market discounts can increase social welfare.

This result above is in sharp contrast with the conclusions in Gans and King (2006), the only other study on cross-market discounts that we know of. They find that competing partnerships will see no increase in profit from cross-market discounts and consumer surplus will always decrease. However, unlike us, they assume demand to be price inelastic, which is a main difference in assumptions

(among others) driving this difference in conclusions. Anecdotal evidence indicates that cross-market discounts have visibly increased household-level consumption of both grocery and fuel. This, coupled with the fact that fuelperks!-type programs are being adopted increasingly, makes a strong case supporting the results we obtain.

A natural question that arises is whether there is something special about the grocery-fuel combine that makes this strategy successful, or could it work for any two unrelated markets (or even two categories in a market, or two products in a category)? From our analysis, we find that the total profit increase from a cross-market discount program will be small if the "importance of consumption" (defined more precisely later) in either of the two markets is small, and this increase in profit is larger as the importance of consumption increases. Therefore, we can expect to see cross-market discounts in pairs of markets in which consumption is high enough (as measured by, say, frequency of purchasing and average dollar amount spent per purchase). Both grocery and fuel constitute a large fraction of a typical household's budget and therefore will typically satisfy this criterion, while any two arbitrary market combinations might not. Since programs such as fuelperks! and Power Pump have provided a "proof of concept" in the grocery-fuel combine, we might soon observe cross-market discounts in other settings that satisfy this criterion. These possibilities further motivate development of a deeper understanding of this promotion strategy.

The rest of the paper is organized as follows. In the next section, we compare cross-market discounts with related promotion strategies. In Section 2, we describe our model, and in Section 3, we present the results from our analysis. In Section 4, we consider extensions to the basic model. In Section 5.3, we summarize our results and conclude with a discussion.

2. Comparison with Related Promotion Strategies

Pricing and promotion strategies have been studied in great depth in Economics and Marketing literature (Neslin 2002, Winer 2006), and here we discuss existing research that shares some common features with cross-market discounts.

A first related stream of research is on *product bundling* --- the strategy of selling two or more products in a package at a discount (Stremersch and Tellis 2002). Bundling is similar to cross-market discounting in that both strategies provide a lower total price when two goods are jointly purchased. However, they differ in several regards. The primary arguments for bundling rest on negative correlation among valuations of the products in the bundle (Adams and Yellen 1976, Venkatesh and Kamakura 2003) and complementarity among utilities from bundled products (Matutes and Regibeau 1992). We do not make any of these assumptions and show that cross-market discounts can be profitable even when implemented in completely independent markets. Furthermore, following Adams and Yellen (1976), much of the academic research on bundling rests on a price discrimination argument and uses a reservation-price paradigm with inelastic demand. In our model, elasticity of demand is an important driver of the results.

Complementary pricing strategies also share similarities with cross-market discounts. From Tellis' taxonomy of pricing strategies (Tellis 1986), two variants are closely related: captive pricing and loss leadership. In a typical *captive pricing* program, one product is sold at a low margin to penetrate the market and lock customers into purchasing a more profitable “tied” product (e.g., razors are sold cheap while the profit is made on blades that need to be used with that razor)¹². Several theories

¹² In the Economics literature, this strategy is also referred as *requirements tying* to emphasize that consumers have to purchase both the *tying* and *tied* good from the same seller.

have been proposed to describe captive pricing (Slade 1998, Mathewson and Winter 1997), price discrimination once again being the prominent one.

In our case, however, tying requirements are not necessary and, modulo the cross-market discount itself, demand in the two markets is completely independent.

Complementary pricing can also take the form of *loss leadership*, a slightly different arrangement in which a retailer offers certain products at low prices to lure customers into the store with the hope that they will purchase other more profitable products (Hess and Gerstner 1987). A successful loss-leader program typically requires economies of scale in shopping for consumers (Lal and Matutes 1994), a feature that our results do not depend on.

Using *cross-ruff coupons* (discount coupons which are obtained upon purchasing one brand, and are redeemable on another brand at a later time) is another complementary pricing strategy which works in a similar fashion to cross-market discounts, and can be modeled similarly. Dhar and Raju (1998) study this but focus on a very different question, namely the time-lagged impact on sales of various types of offerings of these coupons. Moreover, they only consider a monopolist making these decisions, and prices and the amount of discount are treated as exogenous, while these are all critical elements of our study.

An important distinction between our study and the above complementary pricing strategies is that there is no explicit relationship between the prices paid for the various goods in any of these complementary pricing strategies. In the fuelperks!-type of cross-market discounts that grocery

retailers are implementing nowadays, increased purchasing in the source market leads to a larger per-unit discount in the target market, which leads to very different incentives from the above.

Previous work has also shown how *nonlinear pricing* strategies can be used to increase firm profitability in a monopoly (Spence 1977) and in competitive environments (Oren, Smith and Wilson 1983). However, unlike our research, this literature focuses on a single market. While there are various kinds of nonlinear pricing schedules, quantity discounts are the most relevant to our paper. This is similar to our analysis comparing cross-market discounts to similarly specified quantity discounts redeemable in the source market itself, and we characterize conditions under which the firm prefers the former over the latter.

3. Model

We assume that a firm, called firm 1, operates in two distinct and independent markets, s and t , and sells a product in each market at prices p_{s1} and p_{t1} , respectively¹³. Another firm, called firm 2, operates in market t and sells one product in this market at price p_{t2} . Therefore, market s is a monopoly and market t is a duopoly. We assume that the consumption utility function of a representative consumer is given by

$$\mathcal{U}(q_{s1}, q_{t1}, q_{t2}) = \alpha_s \left(q_{s1} - \frac{q_{s1}^2}{2} \right) + \alpha_t \left(q_{t1} - \frac{q_{t1}^2}{2} + q_{t2} - \frac{q_{t2}^2}{2} - \theta_t q_{t1} q_{t2} \right)$$

¹³ We have assumed that firm 1 operates in both markets and, therefore, optimizes jointly in the two markets. If the entities in the two markets are owned separately, our formulation is equivalent to assuming that the two entities cooperatively bargain, which is equivalent to joint optimization in the two markets. Sharing of total profits can be modeled through a lump sum transfer payment decided based on the relative bargaining powers of the two entities, but this will not impact the variables that we focus on here.

where q_{s1} and q_{t1} denote the quantities consumed of firm 1's products in the markets s and t , respectively, and q_{t2} denotes the quantity consumed of firm 2's product in market t . $q_{s1}-q_{s1}^2/2$ and $q_{s1}-q_{s1}^2/2+q_{t1}-q_{t1}^2/2-\theta_t q_{t1}q_{t2}$ denote the consumption utilities in markets s and t , respectively, from consuming the amounts q_{s1} , q_{t1} and q_{t2} , where θ_t denotes the degree of substitutability between the products of the two firms in market t ($0 \leq \theta_t < 1$). Therefore, α_s and α_t denote the “importance” of consumption utility from the markets s and t , respectively, in the aggregated consumption utility function. Note that higher importance of consumption in a market implies lower price sensitivity in that market. Since θ_t denotes the degree of substitutability between the products of the two firms in market t , it serves as a measure of the intensity of competition in market t , with a larger value of θ_t denoting a greater intensity of competition. Note that we have assumed the consumption utility function to be concave in all quantities consumed. This captures diminishing marginal returns from consumption, i.e., as a consumer consumes more of a product, the marginal utility from consuming an extra unit of this product decreases. Using such a quadratic consumption utility function is a standard practice in Economics and Marketing literature (e.g., Arya et al. 2008, Jerath and Zhang 2009, Singh and Vives 1984).

Suppose that firm 1 offers a cross-market discount, $\delta \geq 0$, from s , the source market, to t , the target market. Specifically, if the consumer purchases quantity q_{s1} in the source market, she pays a unit price of $p_{t1}-\delta q_{s1}$ in the target market¹⁴. In this case, the expenditure function associated with consuming quantities q_{s1} , q_{t1} and q_{t2} is given by

¹⁴ We assume that the discount on the per-unit price of firm 1's product in market t grows linearly with the amount of firm 1's product purchased in market s . This is an approximation of the actual fuelperks! discount that grows in a stepwise-linear manner on the per-gallon price of fuel, as discussed in the introduction.

$$\mathcal{E}(q_{s1}, q_{t1}, q_{t2} | p_{s1}, p_{t1}, p_{t2}, \delta) = p_{s1}q_{s1} + (p_{t1} - \delta q_{s1})q_{t1} + p_{t2}q_{t2}$$

An important point to note here is that in the above formulation the price discount in the target market depends on the quantity purchased in the source market, while in our motivating fuelperks! example the price discount depends on the total expenditure in the source market. We start with this simpler formulation to obtain basic insights into the working of cross-market discounts while keeping the model analytically tractable. In Section 5.1, we analyze the case of expenditure-based cross-market discounts. There we have to resort to a partial numerical analysis, and find that, while there are some new insights, the simpler model captures the key insights related to cross-market discounts very well. The net consumer surplus is obtained by subtracting the expenditure from the consumption utility, and is given by

$$\mathcal{CS}(q_{s1}, q_{t1}, q_{t2} | p_{s1}, p_{t1}, p_{t2}, \delta) = \mathcal{U}(q_{s1}, q_{t1}, q_{t2}) - \mathcal{E}(q_{s1}, q_{t1}, q_{t2} | p_{s1}, p_{t1}, p_{t2}, \delta)$$

The profit of firm 1 is given by $\Pi_{st,1}(p_{s1}, p_{t1}, \delta) = \Pi_{s1} + \Pi_{t1} = p_{s1}q_{s1} + (p_{t1} - \delta q_{s1})q_{t1}$, where Π_{s1} and Π_{t1} denote the profits of firm 1 in market s and market t , respectively, and $\Pi_{st,1}$ denotes the joint profit of firm 1 from the two markets. The profit of firm 2 is given by $\Pi_{t2}(p_{t2}) = p_{t2}q_{t2}$ ¹⁵.

We model the game in three stages. In Stage 1, firm 1 decides the cross-market discount δ . In Stage 2, the two firms simultaneously decide the posted prices p_{s1} , p_{t1} and p_{t2} in the two markets, given δ . In Stage 3, the consumer decides how much to purchase of each product in each market, given the

¹⁵ Note that while we are designating one market as the source market and the other as the target market, this distinction is primarily for ease of exposition and ease of accounting of profits --- the expenditure reduction for the consumer and the profit reduction for the firm have the symmetric multiplicative form of $\delta q_{s1}q_{t1}$, whatever the “direction” of the discount. In Section 5.2 we consider expenditure-based cross-market discounts and find that they are not symmetric.

posted prices and the cross-market discount. Our assumption that δ is decided before the prices are set is meant to reflect the fact that the cross-market discount is a long-term decision, while prices can be changed more often. For instance, in the Pittsburgh market, Giant Eagle has not changed the rate of fuelperks! in the last several years. This might also be important to clearly convey the characteristics of the fuelperks! scheme to consumers. Prices, on the other hand, change almost on a daily basis in both the grocery and the fuel markets.

We solve for the subgame-perfect equilibrium of the game using backward induction. In Stage 3, given the posted prices and the cross-market discount, the consumer solves $\max_{\{q_{s1}, q_{t1}, q_{t2}\}} \mathcal{CS}(q_{s1}, q_{t1}, q_{t2} | p_{s1}, p_{t1}, p_{t2}, \delta)$. In Stage 2, given δ , the two firms simultaneously solve $\max_{\{p_{s1}, p_{t1}\}} \Pi_{st,1}(p_{s1}, p_{t1} | \delta)$ and $\max_{\{p_{t2}\}} \Pi_{t2}(p_{t2} | \delta)$. In Stage 1, firm 1 solves $\max_{\delta} \Pi_{st,1}(\delta)$. In each stage we also need to impose the following conditions to guarantee that the model is well defined: we require nonnegative prices and quantities, and in the consumer utility maximization problem we require a nonnegative net consumer surplus (implicitly assuming that the consumer's outside option, corresponding to the case in which she does not purchase anything, is zero).

4. Analysis

We first analyze a monopoly-monopoly scenario. Understanding this simplified scenario helps us to obtain some basic insights into the dynamics of cross-market discounts before we proceed to more complicated scenarios.

4.1. A simple case: the monopoly-monopoly scenario

In this scenario we assume that both market s and market t are monopolies (i.e., $\theta_t = 0$). Since our focus is on firm 1 and firm 2's actions have no effect on firm 1's actions when $\theta_t = 0$, we can simply ignore firm 2.

In Stage 3, given the posted prices and the cross-market discount, the consumer decides the optimal quantities to purchase by maximizing her surplus. This gives the following functions for quantities consumed in each market:

$$q_{s1} = \frac{\alpha_t(\alpha_s + \delta)}{\alpha_s\alpha_t - \delta^2} - \frac{\alpha_t}{\alpha_s\alpha_t - \delta^2} p_{s1} - \frac{\delta}{\alpha_s\alpha_t - \delta^2} p_{t1}$$

$$q_{t1} = \frac{\alpha_s(\alpha_t + \delta)}{\alpha_s\alpha_t - \delta^2} - \frac{\alpha_s}{\alpha_s\alpha_t - \delta^2} p_{t1} - \frac{\delta}{\alpha_s\alpha_t - \delta^2} p_{s1}$$

It is insightful to discuss some salient properties of these demand functions. First, if $\delta = 0$, i.e., no cross-market discount is offered, the demand functions are given by $q_{s1} = 1 - (1/\alpha_s) p_{s1}$ and $q_{t1} = 1 - (1/\alpha_t) p_{t1}$, and the quantities demanded in the two markets are completely independent. In this case, the price sensitivity in market s is given by $1/\alpha_s$, which implies that, as the importance of the source-market product increases, consumers become less price sensitive for this product. α_t plays the same role in market t . Second, the quantities consumed are linear in prices but nonlinear in the cross-market discount. Third, the cross-market discount leads to a “coupling” between the two markets, and a price increase in *either* market leads to consumers purchasing lesser in *both* markets even though they were originally independent. Fourth, as the cross-market discount (δ) increases, the base demand, the own-price sensitivity and the cross-price sensitivity of the consumers increase in both markets, all else equal.

In Stage 2, the firm foresees the above response by the consumers and sets the optimal posted prices, p_{s1} and p_{t1} , by solving its profit maximization problem. The prices and the corresponding quantities consumed are given in Table 1.

PRICES	
p_{s1}	$\frac{\alpha_s \alpha_t (2\alpha_s + \delta)}{4\alpha_s \alpha_t - \delta^2}$
p_{t1}	$\frac{\alpha_s \alpha_t (2\alpha_t + \delta)}{4\alpha_s \alpha_t - \delta^2}$
$p_{s1} - \delta q_{s1}$	$\frac{\alpha_t (2\alpha_s \alpha_t - \alpha_s \delta - \delta^2)}{4\alpha_s \alpha_t - \delta^2}$
QUANTITIES	
q_{s1}	$\frac{\alpha_t (2\alpha_s + \delta)}{4\alpha_s \alpha_t - \delta^2}$
q_{t1}	$\frac{\alpha_s (2\alpha_t + \delta)}{4\alpha_s \alpha_t - \delta^2}$

Table 1: Prices and quantities in Stage 2 in the monopoly-monopoly scenario.

These expressions reveal an interesting “cross-market leverage” effect that the firm uses to its advantage. For a fixed δ , the effective per-unit price in the target market decreases with quantity consumed in the source market. Therefore, consumers are incited by a cross-market discount to increase consumption in the source market, even at a higher price in this market. Foreseeing this, the firm *raises* the price in the source market. In other words, in spite of the fact that consumers have a negative price elasticity of demand, they increase the quantity they purchase in the source market at a higher price when a cross-market discount is in place. (Specifically, in our model, both price and quantity simultaneously increase because the base demand, i.e., the intercept of demand, increases with δ .) Therefore, there is increased consumption in both markets, at a higher price in the source market and a lower effective price in the target market. We state the above result in the following proposition.

Proposition 1. *After the introduction of a cross-market discount, both price and quantity sold in the source market increase simultaneously.*

The above proposition holds whenever the model is well defined, the condition for which is $\delta < \sqrt{\alpha_s \alpha_t}$. The firm makes a larger profit in the source market (since both price and consumption increase in this market) but a lower profit in the target market. Overall, as δ increases, the combined profit from the two markets increases due to a larger cross-market leverage effect. From the consumers' point of view, their surplus *decreases* as a result of the cross-market discount, because they are purchasing more in the target market at a lower price and more in the source market, but at a higher price. The decrease in surplus due to the price increase in the source market dominates any increases in surplus due to higher consumption in both markets or lower price in the target market. Therefore, when a cross-market discount is being offered, the prices charged are such that it is optimal for consumers to avail it, but they would have been better off if it were not offered in the first place.

Given the results of Stage 2 above, in Stage 1 the firm decides the optimal level of the cross-market discount rate (δ) to maximize its joint profit from the two markets. The expression for the joint profit is given by $\Pi_{s,t} (p_{s1}, p_{t1}, \delta) = \alpha_s \alpha_t (\alpha_s + \alpha_t + \delta) / (4\alpha_s \alpha_t - \delta^2)$. From the expressions and the discussion above, we can see that this profit monotonically increases in δ , while the consumer surplus and the effective price in the target market monotonically decrease in δ . Therefore, the firm will keep increasing δ up to the point at which either the consumer surplus is zero or the effective price in the target market is zero, and this gives the optimal level of δ . At this point, the total profit

of the firm from the two markets is higher than the sum of the monopoly profits that it was making in the two markets without using a cross-market discount.

If consumption in the target market is not important (i.e., α_t is considerably smaller than α_s), prices in the target market will be low because consumers are not willing to spend a lot in this market. In this case, the firm increases δ up to the point at which the effective price in the target market is zero. On the other hand, if consumption in the target market is important enough, the firm can charge a higher price in this market and, in this case, the constraint that consumer surplus should be nonnegative is binding, i.e., the firm can extract the full consumer surplus by employing a cross-market discount and increasing price in the source market. Note that the firm can achieve this even when the price elasticity of demand is negative in both markets, and this was not possible without the cross-market discount. We state the above result as a proposition.

Proposition 2. *If neither market is competitive, the firm can always use a cross-market discount to extract greater consumer surplus across the two markets. Furthermore, if the target market is less price sensitive than the source market (i.e., $\alpha_t > \alpha_s$), then the firm can set the cross-market discount rate (δ) to extract the full consumer surplus across the two markets.*

Furthermore, the optimal cross-market discount rate increases with both α_s and α_t . This is because a larger value for either of these parameters increases the importance of these categories in the consumption utility function because of which consumers purchase more of both. Therefore, they attain more consumption utility, all of which the firm extracts by setting a higher value of δ , and leaves them with no surplus. We now provide some generalizations to the above model and compare them with quantity discounts restricted to the source market.

Different rates of diminishing marginal utilities

In the above analysis, for analytical simplicity, we assume that the rates of diminishing marginal utility from consumption are equal in the source and the target markets. We now consider a more general case in which the consumption utility function is given by $\mathcal{U}(q_{s1}, q_{t1}) = \alpha_s (q_{s1} - \psi_s q_{s1}^2/2) + \alpha_t (q_{t1} - \psi_t q_{t1}^2/2)$, where $\psi_s > 0$ and $\psi_t > 0$. By increasing or decreasing the values of ψ_s and ψ_t , we can increase or decrease the rates of decrease of marginal utility from consumption (or, the rates of “satiation” from consumption) in the source and target markets, respectively. (In our basic model, $\psi_s = \psi_t = 1$.) When no cross-market discount is applied ($\delta = 0$), the demand functions are given by $q_{s1} = 1/\psi_s - 1/(\alpha_s \psi_s) p_{s1}$ and $q_{t1} = 1/\psi_t - 1/(\alpha_t \psi_t) p_{t1}$. Therefore, if ψ_s and ψ_t increase, the base demand in the market in question decreases and the demand in this market becomes less sensitive to price.

As before, the equilibrium in this scenario is derived using backward induction (details are provided in Section 1 in the Technical Appendix 1). The basic insights of Propositions 1 and 2 remain unchanged because both prices and quantities are increasing in δ . However, larger values of ψ_s and ψ_t imply that demand is less price sensitive (as mentioned above), which reduces the rates at which prices and quantities, and therefore profit, increase with an increasing cross-market discount rate. The other key impact of the rates of diminishing marginal utilities is on the binding constraints in Stage 1. It turns out that both consumer surplus and effective price in the target market decrease at a slower pace for larger values of rates of diminishing marginal utilities, enlarging the region of feasible values of δ .

Therefore, when consumer satiate faster, we have two forces moving in opposite directions. Although profit increases at a slower rate with δ , the firm can impose a larger rate of the discount. In this scenario, the former effect dominates, providing the firm smaller gains in profits when consumers have larger rates of satiation. The ability of the firm to extract consumer surplus also depends on the rates of diminishing marginal utilities, and the effect of the rate in each market is different. We find that the set of values of α_s and α_t where the firm can fully extract consumer surplus is enlarged by larger values of ψ_s and smaller values of ψ_t .

Finally, it is worthwhile to note that if marginal utility from consumption is constant rather than decreasing (i.e., $\psi_s = \psi_t = 0$), then cross-market discounts will be futile. This is because this is a degenerate case in which if the price in a market is lower than marginal utility, then the consumer will consume the maximum amount available. In other words, with constant marginal utility, if the consumer has the incentive to consume one unit, then she has the incentive to consume all the units available. Therefore, there is no reason to offer any discount to induce more consumption.

Quantity discounts limited to the source market

How do cross-market discounts compare to similar discounts in the source market itself? Such a “self-market discount” scheme could be considered similar to the more standard quantity discounts or loyalty programs in which consumers are offered more grocery rebates as they purchase more¹⁶.

In this section, we show that due to the diminishing marginal returns from consumption, a firm may prefer cross-market discounts to self-market discounts. We consider only the monopoly-monopoly scenario; the same basic insights hold for the other scenarios (that we discuss subsequently) as well.

¹⁶ In a typical quantity discount schedule, the unit price of a good decreases at a decreasing rate as more quantity is purchased. However, to study the self-market discount analog of the fuelperks! scheme, we model a “supercharged” quantity discount in which the unit price of the good decreases linearly as more quantity is purchased.

To model the impact of diminishing marginal returns, we assume that the consumption utility function is given, as above, by $\mathcal{U}(q_{s1}, q_{t1}) = \alpha_s (q_{s1} - \psi_s q_{s1}^2/2) + \alpha_t (q_{t1} - \psi_t q_{t1}^2/2)$ where $\psi_s > 0$, $\psi_t > 0$. In the case of a cross-market discount, denoted by δ , the expenditure function is given as before. In the case of a self-market discount, denoted by δ_s , the effective per-unit price of the product in market s is given by $p_{s1} - \delta_s q_{s1}$, and the expenditure function is given by $\mathcal{E}(q_{s1}, q_{t1} | p_{s1}, p_{t1}, \delta) = (p_{s1} - \delta_s q_{s1}) q_{s1} + p_{t1} q_{t1}$. We conduct the analysis in the same way as in Section

Error! Reference source not found.; the details are in Section 1.1 in the Technical Appendix 1.

Consider the simplified case in which $\psi_s = \psi_t = \psi$, i.e., the rate of diminishing marginal returns is the same in both the source and the target markets. A self-market discount induces more consumption only in the source market, while a cross-market discount induces more consumption partly in the source market and partly in the target market, as shown earlier. This implies that a self-market discount leads to consumers reaching the point of reduced consumption utility earlier than a cross-market discount, because the latter slows down “satiation” in both markets by distributing additional consumption (motivated by the discount) across the two markets. Consequently, the optimal cross-market discount is larger than, and induces more total consumption than, the optimal self-market discount. Therefore, as ψ increases, the firm finds it optimal to employ a cross-market discount scheme rather than an analogous self-market discount scheme, because cross-market discounts allow the firm to exploit the less price sensitive parts of the consumers' utility functions in the two markets.

The above insights carry over to the cases in which $\psi_t \neq \psi_s$. Notably, if $\psi_s > \psi_t$, then the rate at which marginal utility decreases with additional consumption is faster in the source market than in the target market and, therefore, cross-market discounts are even more likely to be offered as compared to self-market discounts. On the other hand, if ψ_s is significantly smaller than ψ_t , then self-market discounts are preferred. The exact condition under which the firm will prefer a cross-market discount strategy over a self-market discount strategy is $2\alpha_s\psi_s - 3\alpha_t\psi_t + \sqrt{8\psi_s\psi_t\alpha_s\alpha_t + \alpha_s\alpha_t} > 0$. We summarize the discussion above in the following proposition.

Proposition 3. *The firm prefers a cross-market discount over a self-market discount, except when the marginal utility of consumption in the target market diminishes significantly faster than it does in the source market.*

Non-zero marginal cost

In our analysis above, we assume that the marginal cost for the firm is zero in both markets, and impose the condition that the effective price in the target market should be nonnegative. In Section 1.2 in the Technical Appendix 1, we allow a positive marginal cost in the target market and find that the firm may offer a cross-market discount large enough to price below marginal cost in this market. In our motivating example in the introduction (in which a discount of 10¢ per gallon is offered for every \$50 spent on grocery), this indeed seems to be the case since fuel pumps across USA are known to make a profit of only 3¢ to 15¢ per gallon (Robbins 2008). Besides the above, there is no qualitative difference in the results. (Note that a zero marginal cost in the source market is not a restrictive assumption since prices in the source market increase after the introduction of a cross-market discount.)

4.2. The monopoly-duopoly scenario

In the monopoly-duopoly scenario, we allow for competition in the target market by assuming $0 \leq \theta_t < 1$. This scenario is a closer representation of our motivating example. We build on the insights obtained from the analysis of the monopoly-monopoly scenario, some of which carry over qualitatively. The new effects that arise can directly be attributed to the strategic interaction between competitors in the target market.

As before, we start with Stage 3, in which the consumer decides the optimal quantities to purchase by maximizing her surplus, given the posted prices and the cross-market discount. This gives the following demand functions:

$$q_{s1} = \frac{1}{\alpha_s \alpha_t (1 - \theta_t^2) - \delta^2} \left(\alpha_t (1 - \theta_t) (\alpha_s (1 + \theta_t) + \delta) - \alpha_t (1 - \theta_t) p_{s1} - \delta p_{t1} + \delta \theta_t p_{t2} \right)$$

$$q_{t1} = \frac{1}{\alpha_s \alpha_t (1 - \theta_t^2) - \delta^2} \left((\alpha_s \alpha_t (1 - \theta_t) + \alpha_s \delta) - \delta p_{s1} - \alpha_s p_{t1} + \alpha_s \theta_t p_{t2} \right)$$

$$q_{t2} = \frac{1}{\alpha_s \alpha_t (1 - \theta_t^2) - \delta^2} \left((\alpha_s \alpha_t (1 - \theta_t) - \alpha_s \theta_t \delta - \delta^2) + \delta \theta_t p_{s1} + \alpha_s \theta_t p_{t1} - (\alpha_s \alpha_t - \delta^2) p_{t2} \right)$$

First, note that if $\theta_t = 0$, we obtain exactly the demand functions for firm 1 in the monopoly-monopoly scenario. When $\theta_t > 0$, i.e., there is competition in market t , the quantities consumed are linear in prices but nonlinear in the cross-market discount. As before, if $\delta = 0$, the two markets are completely independent, and if $\delta > 0$, the two markets are “coupled”. If firm 1 increases price in either market, all else equal, the quantities demanded for its products in both markets reduce and the quantity demanded of firm 2's product in market t increases. If firm 2 increases its price in market t , its own demand decreases, while the quantities demanded of firm 1's products in both markets increase. If firm 1 increases δ , the base demands for both of firm 1's products increase and the base

demand for firm 2's product in market t decreases, and the above discussed reactions of consumers to prices become stronger.

The firms foresee the above response by the consumers and, in Stage 2, they simultaneously solve $\max_{\{p_{s1}, p_{t1}\}} \Pi_{st,1}(p_{s1}, p_{t1} | \delta)$ and $\max_{\{p_{t2}\}} \Pi_{t2}(p_{t2} | \delta)$. The prices and the corresponding quantities are given in Table 2. Figure 1 shows equilibrium prices, quantities demanded and profits as functions of the discount rate δ for a representative case. As in the monopoly-monopoly scenario, this analysis also reveals the “cross-market leverage” effect when $\delta > 0$ - for firm 1, posted prices are higher in both markets (while price after discount is lower in market t), quantities consumed are higher in both markets and total profit is higher (decomposing this, profit is higher in market s but lower in market t).

PRICES	
p_{s1}	$\frac{\alpha_s \alpha_t (\alpha_s^2 \alpha_t (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_t \delta (\alpha_t (2 - \theta_t^2 - \theta_t) - \delta (4 - 3\theta_t^2))) + \delta^3 (2 - \theta_t)}{2\alpha_s^2 \alpha_t^2 (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \alpha_t \delta^2 (7\theta_t^2 - 10) + 2\delta^4}$
p_{t1}	$\frac{\alpha_s \alpha_t (1 - \theta_t^2) (\alpha_s \alpha_t (2\alpha_t (2 - \theta_t^2 - \theta_t) + \delta (2 - \theta_t^2)) - 2\delta^2 (\delta + \alpha_t (2 - \theta_t)))}{2\alpha_s^2 \alpha_t^2 (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \alpha_t \delta^2 (7\theta_t^2 - 10) + 2\delta^4}$
p_{t2}	$\frac{\alpha_t (\delta^2 - \alpha_s \alpha_t (1 - \theta_t^2)) (\alpha_s (\delta \theta_t - 2\alpha_t (2 - \theta_t^2 - \theta_t)) + \delta^2)}{2\alpha_s^2 \alpha_t^2 (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \alpha_t \delta^2 (7\theta_t^2 - 10) + 2\delta^4}$
$p_{s1} - \delta q_{s1}$	$\frac{\alpha_t (2\alpha_s^2 \alpha_t (1 - \theta_t^2) (\alpha_t (2 - \theta_t^2 - \theta_t) - \delta) + \alpha_s \delta^2 (\alpha_t (-2\theta_t^3 + 5\theta_t^3 + 3\theta_t - 6) + \delta (2 - \theta_t^2))) + \delta^4 (2 - \theta_t)}{2\alpha_s^2 \alpha_t^2 (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \alpha_t \delta^2 (7\theta_t^2 - 10) + 2\delta^4}$
QUANTITIES	
q_{s1}	$\frac{\alpha_t (\alpha_s^2 \alpha_t (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \delta (\alpha_t (2 - \theta_t^2 - \theta_t) - \delta (4 - 3\theta_t^2))) + \delta^3 (2 - \theta_t)}{2\alpha_s^2 \alpha_t^2 (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \alpha_t \delta^2 (7\theta_t^2 - 10) + 2\delta^4}$
q_{t1}	$\frac{\alpha_s (\alpha_s \alpha_t (2\alpha_t (2 - \theta_t^2 - \theta_t) + \delta (2 - \theta_t^2)) - 2\delta^2 (\delta - \alpha_t (\theta_t - 2)))}{2\alpha_s^2 \alpha_t^2 (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \alpha_t \delta^2 (7\theta_t^2 - 10) + 2\delta^4}$
q_{t2}	$\frac{\alpha_s^2 \alpha_t (2\alpha_t (2 - \theta_t^2 + \theta_t) - \delta \theta_t) + \alpha_s \delta^2 (\alpha_t (2\theta_t^2 + 2\theta_t - 5) + \delta \theta_t) + \delta^4}{2\alpha_s^2 \alpha_t^2 (\theta_t^4 - 5\theta_t^2 + 4) + \alpha_s \alpha_t \delta^2 (7\theta_t^2 - 10) + 2\delta^4}$

Table 2: Prices and quantities in Stage 2 in the monopoly-duopoly scenario.

But the story is richer in this scenario because of the strategic reaction of the competitor in market t . If $\delta > 0$, firm 1's effective price in market t decreases, in response to which firm 2 lowers its price in market t as well, and as δ increases, this reaction from firm 2 becomes stronger.

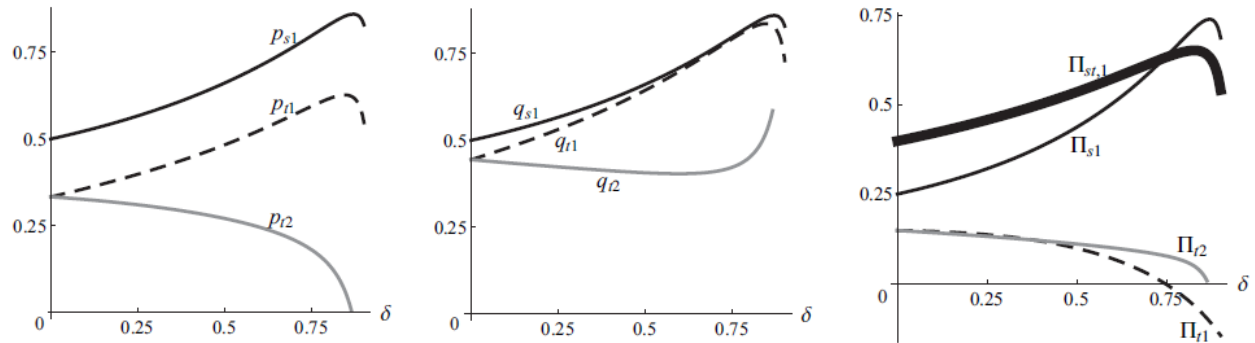


Figure 1: Prices, quantities and profits as functions of the cross-market discount rate (δ) in Stage 2 in the monopoly-duopoly scenario. For these plots, we use $\alpha_s = \alpha_t = 1$ and $\theta_t = 1/2$.

For δ large enough, firm 2 charges a very small price to increase its own sales, with the result that the cross-market leverage effect for firm 1 starts to decrease because fewer units are purchased from it in market t . Due to reduced cross-market leverage, any further increase in δ leads to a decrease in posted prices, sales and profits for firm 1 in both markets. In summary, for firm 1, the joint profit from the two markets has an inverted-U shape in δ , unlike in the monopoly-monopoly scenario in which it always monotonically increases with δ . This is due to the strategic reaction by the competitor in market t ¹⁷.

¹⁷ At extreme values of θ_t , however, the trends are still monotonic. At one extreme, if θ_t is close to 0, the model resembles the monopoly-monopoly scenario and therefore prices and quantities demanded are monotonically increasing in δ . At the other extreme, if θ_t is close to 1, prices are very low because of intense competition. In this case, the constraint that the effective price in the target market should be nonnegative is binding before the decreasing part of firm 1's profit curve is reached.

Given the results of Stage 2, in Stage 1, firm 1 solves $\max_{\{\delta\}} \Pi_{st,1}(\delta)$ to decide the optimal level of the cross-market discount rate δ . From the analysis above, we can see that the optimum solution for δ may be interior. At this optimum, firm 2 makes lower profits than when no cross-market discount was being offered, while Firm 1 makes higher profits because of the cross-market leverage¹⁸. Note that the consumer surplus at the optimum level of the cross-market discount may be positive and larger than without a cross-market discount, unlike in the monopoly-monopoly case. In the monopoly-duopoly case, after cross-market discounts have been introduced, consumers obtain smaller surplus in the grocery market (due to higher prices in this market through the cross-market leverage effect) but a significantly larger surplus in the target market due to competition in this market. Overall, if the competitive intensity in the target market is high, total consumer surplus increases after the introduction of a cross-market discount, while if it is low, the total consumer surplus decreases (as in the monopoly-monopoly case). Therefore, both firm profit and consumer surplus can simultaneously increase in this case after cross-market discounts are introduced.

We now study the characteristics of the optimal cross-market discount rate. First, the optimal value of δ increases with both α_s and α_t , keeping other parameters fixed. This is because a larger value for either of these parameters increases the importance of these categories in the consumption utility function because of which consumers purchase more of both. Therefore, they attain more consumption utility, but the firm also sets δ at a larger value and can extract more of the consumer surplus (even if it is unable to extract all of it).

¹⁸ Profit of firm 2 might also go to zero at the boundaries of the parameter space: when target market has no attractiveness to the consumer ($\alpha_t \rightarrow 0$), asymptotically when the importance of the source market dominates consumer's decision ($\alpha_s \rightarrow \infty$), or when the degree of competition approaches its maximum ($\theta_t \rightarrow 1$).

The impact of the degree of competition in the target market, θ_t , on the optimal cross-market discount rate, δ , is more interesting. Arguably, an increase in θ_t can lead to an increase or a decrease in δ . On the one hand, one can argue that if the target market is more competitive, the discount rate should be higher to induce more consumers to purchase from firm 1 in market t . On the other hand, a larger cross-market discount in a more competitive market could lead to a stronger reaction by firm 2, which can lead to overall lowering of prices and reduce everybody's profits.

We find that, in our model, the second effect dominates the first, i.e., the cross-market discount rate is smaller in a more competitive market. As θ_t increases from zero, i.e., intensity of competition in the target market increases, the profit that can be made in this market, and the extra profit from cross-market leverage, both decrease. Since there is lesser profit to extract, the optimal value of the cross-market discount rate, δ , decreases. This is shown in Figure 2 for a representative case. Furthermore, as long as θ_t is small, firm 2 does not react very strongly to a cross-market discount. So, while the cross-market discount rate decreases with increasing θ_t , it decreases slowly. However, as θ_t becomes larger, the reaction by the competitor becomes progressively stronger and optimal δ starts decreasing in a concave fashion, i.e., decreasing at an increasing rate. As θ_t gets close to 1, price sensitivities in the target market become very large and prices in the target market approach zero (equal to marginal cost), and there is no profit in the target market to extract. Therefore, offering a cross-market discount is of no use to firm 1, so that the optimal value of δ is zero. We confirm all of the results above through a comprehensive analysis allowing for different values of the exogenous parameters of the model (θ_t , α_s , and α_t ; details are provided in Section 1.3 in the Technical Appendix 1), and summarize them in the following proposition.

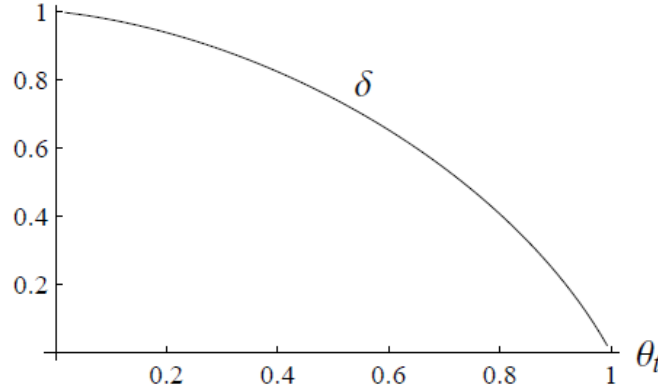


Figure 2: Plot of the optimal rate of the cross-market discount (δ) with competitive intensity in the target market (θ_t), when $\alpha_s = \alpha_t = 1$.

Proposition 4. *If the target market is competitive, the optimal value of the cross-market discount rate (δ):*

- *Decreases in the competitive intensity in the target market (θ_t).*
- *Increases in the importance of the source-market consumption utility (α_s) and in the importance of the target-market consumption utility (α_t).*

5. Extensions

5.1. The duopoly-duopoly scenario

While it is reasonable to assume that the source market is a monopoly for our motivating example of the Pittsburgh, PA market, grocery retailers that are adopting the fuelperks! and Power Pump schemes in other areas might be facing significant competition in the grocery market as well. Therefore, we extend our analysis to a duopoly-duopoly scenario by assuming that the focal firm faces competition in the source market as well.

In this scenario, when only the focal firm offers a cross-market discount and the competitors in the two markets act independently, the focal firm is better off and the competitors in both markets are worse off. Furthermore, competition in the source market has similar effects as competition in the

target market and the basic insights from the model remain largely unchanged. More details are available in Section 2 in the Technical Appendix 1.

However, if competing firms are present in both markets, then these firms (i.e., the focal firm's competitor in the source market and its competitor in the target market) might have the incentive to join hands and introduce their own cross-market discount program in competition with the focal firm's cross-market discount program. We assume that two firms (1 and 2) compete simultaneously in both source (s) and target (t) markets. (Note that we have clubbed the competing entities in the two markets into “firm 2”, and firms 1 and 2 are assumed to be completely symmetric.) We keep all notation used in previous scenarios except for the introduction of an index for the rate of cross-market discount offered by each firm (δ_1 and δ_2). Then, in the symmetric duopoly-duopoly scenario the consumer utility and expenditure are given by

$$\mathcal{U}(q_{s1}, q_{s2}, q_{t1}, q_{t2}) = \alpha_s \left(q_{s1} - \frac{q_{s1}^2}{2} + q_{s2} - \frac{q_{s2}^2}{2} - \theta_s q_{s1} q_{s2} \right) + \alpha_t \left(q_{t1} - \frac{q_{t1}^2}{2} + q_{t2} - \frac{q_{t2}^2}{2} - \theta_t q_{t1} q_{t2} \right)$$

$$\mathcal{E}(q_{s1}, q_{s2}, q_{t1}, q_{t2} | p_{s1}, p_{s2}, p_{t1}, p_{t2}, \delta_1, \delta_2) = p_{s1} q_{s1} + p_{s2} q_{s2} + (p_{t1} - \delta_1 q_{s1}) q_{t1} + (p_{t2} - \delta_2 q_{s2}) q_{t2}$$

We again model the game in three stages. In Stage 3, the consumer decides how much she will consume of each product in each market. In Stage 2, the two firms simultaneously decide all retail prices in both markets. In Stage 1, both firms simultaneously decide their optimal cross-market discount rates¹⁹. Below we discuss the insights obtained from our analysis; more details are available in Section 3 in the technical appendix.

¹⁹ As we described in a footnote in Section XX, if the competing entities in the two markets are owned separately, our formulation is equivalent to assuming that the two entities cooperatively bargain with lump sum transfer payments (i.e., s_1 and t_1 cooperatively bargain, and s_2 and t_2 cooperatively bargain), which is equivalent to joint optimization by these entities in the two markets (i.e., joint optimization by s_1 and t_1 , and joint optimization by s_2 and t_2).

We are interested in analyzing how firm profits and consumer surplus change in comparison to the case in which cross-market discounts are not implemented. To simplify the analysis and focus on the key insights, we assume that both markets are equally important for the consumers and equally competitive for the firms, i.e., $\alpha_s = \alpha_t = \alpha$ and $\theta_s = \theta_t = \theta$. Our analysis shows that both firms always offer cross-market discounts. In Figure 3, we plot for a representative case the optimal rate of the cross-market discount (which, in equilibrium, is equal for both firms due to symmetry, i.e., $\delta_1 = \delta_2 = \delta$), the change in profit for the firms and the change in surplus for the consumers due to the cross-market discounts.

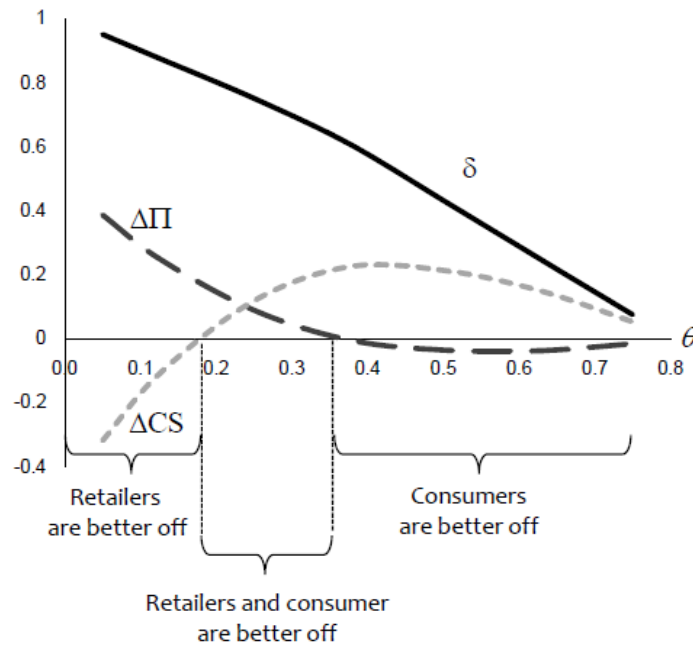


Figure 3: Impact of the competitive intensity (θ) on the optimal rate of the cross-market discount (δ) offered by both firms, the change in profit for each firm ($\Delta\Pi$) and the change in consumer surplus (ΔCS) due to the cross-market discounts. For this figure, we use $\alpha=1$.

As expected, when the competitive intensity (θ) is small, this scenario resembles the monopoly-monopoly scenario and the retailers are better off and consumers are worse off with cross-market discounts. On the other extreme, when the competitive intensity is high, cross-market discounts

actually hurt firms' profits because they invite stronger reactions from the competing firm. However, each firm still offers them because of a classic “prisoners' dilemma” situation. In this case, consumer surplus is higher after cross-market discounts are introduced.

The more interesting result, however, occurs when the intensity of competition is medium. In this case, *both* the consumers and the firms are better off through cross-market discounts. This is because consumption is higher and, because competitive intensity is medium, prices neither rise too much (which would have hurt the consumers) nor decrease too much (which would have hurt the firms). Therefore, the benefit from a cross-market discount strategy is positive even when a competitor decides to imitate the strategy and, moreover, this benefit is not at the expense of the consumers. We confirm that these results are replicated for different values of the exogenous parameters (θ and α) by conducting a comprehensive numerical analysis (details available in Section 3.1 in the Technical Appendix 1), and summarize them in the proposition below.

Proposition 5. *If both the source market and the target market are competitive, both firms always offer cross-market discounts. Under moderately intense competition, profit of each firm increases, and consumer surplus also increases.*

The above result is in sharp contrast with the results in Gans and King (2006), the only other study on cross-market discounts that we know of. They model cross-market discounts as a bundling problem and obtain the conclusion that nobody benefits from them under any conditions - competing partnerships see no increase in profit and consumer surplus also decreases. However, they consider cross-market discounts primarily as a way to poach a competitor's consumers, and do

not allow consumer demand to be price elastic. This is a main difference between the two models (among others), which leads to these very different conclusions. In actual fuelperks! implementations, there is evidence that consumption of both grocery and fuel has increased. Indeed, when explaining why fuelperks! work, a Giant Eagle executive stated that “*essentially fuelperks! is funded by the amount of additional store sales that we get inside our multiple store formats - both GetGo and Giant Eagle*” (McTaggart 2006). We believe that this evidence offers strong support for our formulation and insights.

5.2. Discount based on expenditure in the source market

In the basic model, we assume that a cross-market discount is tied to quantity purchased in the source market. In this extension, we consider a cross-market discount tied to the total amount spent by a consumer in the source market, as we see in several implementations of this strategy. We find that most of our key results from the previous sections are reproduced, and we discuss some new insights as well.

Suppose that a cross-market discount, δ_r , is offered from market s to market t such that, if a consumer purchases an amount q_{s1} at the price p_{s1} in market s , she obtains a unit price of $p_{t1} - \delta_r p_{s1}$ in market t . The only change from the original monopoly-duopoly model is in the expenditure function, which is now given by $\mathcal{E}(q_{s1}, q_{t1}, q_{t2} | p_{s1}, p_{t1}, p_{t2}, \delta_r) = p_{s1} q_{s1} + (p_{t1} - \delta_r p_{s1}) q_{t1} + p_{t2} q_{t2}$. The analysis in this case is algebraically intractable in Stage 1, so we resort to a numerical study. The details are provided in Section 3.2 in the Technical Appendix 1.

First, we find that the qualitative impact of the discount rate δ_r on prices and quantities is the same as in our basic model. Profits also follow exactly the same patterns as before: total profit for firm 1

increases in δ_r if competitive intensity in the target market is low, and has an inverted-U shape in δ_r if this competitive intensity is high. Consumer surplus also follows the same trends as before: it decreases in δ_r if competitive intensity is low, and increases in δ_r if competitive intensity is high enough. Consequently, the qualitative effect of each parameter of the model on the optimal level of the cross-market discount also remains unchanged. We conclude that, compared to this formulation, the basic model is significantly more parsimonious and captures the relevant insights of cross-market discounts.

However, a new and interesting feature of this extension is that a discount from market s to market t may not be symmetric to a discount from market t to market s , because the prices in the two markets may be different. Therefore, we can explore which direction of the discount is more profitable for the focal firm. Our numerical analysis shows two interesting results. First, when market s is relatively more important to the consumer than market t , the focal firm makes larger profits by allowing consumers to accumulate discounts in market s to redeem them in market t . When market t is relatively more important, the relation reverses. Second, competitive structure of a market also plays a role. We find that when both markets have the same importance, the focal firm will make larger profits by offering a discount from the monopoly to the duopoly market. In other words, it is more profitable if cross-market discounts are redeemable in the market that faces competition.

Both of the above results are in line with what we observe in the actual implementations of cross-market discounts from grocery to gasoline. First, according to the Consumer Expenditure Survey conducted by the Bureau of Labor Statistics, consumers spend on average more than double on

grocery than on gasoline (CES 2009)²⁰. Second, gasoline is a highly commoditized market, making it more competitive than grocery. These indicate that, in actual implementations, the source market (grocery) is indeed the more important market and the less competitive market than the target market (gasoline).

5.3. Other assumptions

We now briefly discuss some simplifying assumptions that we have made in our model. In this paper, for both cross-market discounts and self-market discounts, we have only considered price discounts that are linear in quantity consumed. We find that such a discount motivates higher consumption in both cases, and typically motivates higher consumption in the cross-market discount case. Other “more accelerated” forms of discount, such as one in which the price discount increases faster than linear with quantity consumed, can motivate even higher total consumption. However, even for such an accelerated discount schedule, we can expect that cross-market discounts may be more beneficial because they distribute additional consumption across two markets, thus delaying satiation in both. Future research can explore this question in more detail.

Another simplifying assumption we make is that consumption utility functions are concave. However, these functions might have different forms, such as an S-shaped form. In this case, a cross-market discount will not have a significant impact at low levels of consumption (i.e., the initial flat part of the consumption utility function). However, we expect our results to be applicable at higher levels of consumption. Therefore, we expect our results and insights to hold, although their impact will be “delayed”. Furthermore, this might, based on the specific characteristics of the S-

²⁰ We compare the total expenditure in the categories “Food at Home” and “Gasoline and Motor Oil” as defined in the Consumer Expenditure Survey. The average ratio from 1997 to 2007 is 2.22. This is a conservative estimate because “grocery” typically contains many more categories than those included under “Food at Home” in the CES.

shaped utility functions in the two markets, also have an impact on the direction of the cross-market discount.

On a similar note, studies have shown that there can be significant uncertainty about the functional form of demand outside observed consumption/purchasing ranges (Montgomery and Bradlow 1999). Our analysis, which assumes demand to be linear in posted prices, cannot directly speak to these situations. While we believe that even in these cases managers will still have the incentive to increase prices using cross-market discounts, it is important to bring other factors into the mix while making this decision, such as stronger consumer reactions to higher prices in the source market.

6. Conclusions and Discussion

Promotional programs such as fuelperks! and Power Pump have become very popular with grocery retailers in the last few years. In these cross-market discount programs, consumers can accumulate discounts by purchasing groceries at a particular grocery store (grocery being the “source market”), and redeem these discounts when purchasing fuel at partnering gas stations (fuel being the “target market”). In this paper, we use an analytical model to obtain a deeper understanding of the working of such cross-market discounts.

We find that, across different competitive structures, profits in this scheme accrue from a simultaneous increase in both prices and sales in the source market, while the effective price that consumers pay in the target market reduces. In fact, firms can have an incentive to sell the product in the target market below marginal cost to increase total profits across the two markets. We also provide a characterization of the attributes of market pairs that favor a profitable cross-market

discount program. Our model predicts that cross-market discounts will be preferred from a higher-importance-of-consumption market to a lower-importance-of-consumption market, and from a less competitive to a more competitive market. Both predictions are consistent with the fuelperks!-type promotions observed in reality. We also find that cross-market discounts can be more profitable than comparable nonlinear pricing strategies that focus on only one market, because they exploit less price sensitive portions of consumers' utility functions (when there is diminishing marginal utility).

Another trend is that, in most markets, new grocery-gasoline combines are introducing fuelperks!-type schemes in response to such programs started by competitors. In accordance with this, our model predicts that when both source and target markets are competitive, competing partnerships will simultaneously offer cross-market discounts. Furthermore, when the competitive intensity is low, these partnerships will benefit at the cost of consumers. When the competitive intensity is high, both will see reduced profits from cross-market discounts but still offer them because they are stuck in a “prisoners' dilemma” situation, and consumers will see increased surplus. Interestingly, when competitive intensity is medium, both the competing partnerships and the consumers will benefit simultaneously.

While we focus on the important grocery-gasoline combine throughout the paper, our results might also be applicable in certain other similar settings. For instance, airlines typically offer deals under which miles accumulated on one flight route can be redeemed for a discount on a different flight route, or for a hotel stay or car rental. Another such interesting program is the “aeroplan” rewards program from Air Canada in which, upon purchasing air tickets, consumers obtain points that can be used for subsidized fuel at Esso gas stations in Canada.

We now discuss some avenues for future work on cross-market discounts. We assume for simplicity that the focal firm owns stores in both the source market and the target market. If this is not the case, the incentives of firms in different markets to participate in such a promotional program need to be understood. For example, while a cooperative bargaining solution will still lead to the same pricing and discounting decisions that we find, the issue of dividing profits in these strategic alliances will arise. Furthermore, we assume that partnering firms in different markets sign exclusive partnership contracts. Future work can allow non-exclusive partnerships, which can lead to interesting incentives, and resulting configurations, for cross-market alliances.

We have compared cross-market discounts with self-market discounts only in the source market. However, self-market discounts could be used in both the source market and the target market. We have conducted some basic exploratory analysis which shows that, under certain conditions, a cross-market discount in addition to two self-market discounts can further increase firm profits by motivating more consumption. A full analysis of this strategy is outside the scope of this paper, and future work can study it in more detail.

We also assume independence between demands in the two markets. This assumption helps to show that cross-market discounts are distinct from bundling, as they can be useful even when the usual incentives for bundling are absent. Analyzing cases in which market demands and valuations are correlated can further enhance our understanding of cross-market discounts. It is also possible that the importance of consumption parameters for the two markets could be impacted by a cross-market discount between the markets. This could lead to interesting implications for these discounts.

We also limit our analysis to a static setting which allows us to obtain some key insights using a simplified model. However, in a dynamic model, in which consumers can purchase in the source market and redeem discounts in the target market at different times, interesting dynamics such as discount accumulation and dynamically optimizing consumption jointly in the two markets will arise. A dynamic model can also explicitly build a “loyalty program” or a “reward program” component into the analysis of cross-market discounts. This can capture dynamics through which current purchasing from a firm promotes future purchasing from this firm or its partner in a different market, which can help to examine the time-lagged impact of cross-market discounts on prices, demand and intensity of competition (Kim, Shi and Srinivasan 2001).

Finally, in 2009, Giant Eagle and GetGo started a “foodperks!” program simultaneously with the “fuelperks!” program. Consumers can now not only purchase groceries to earn discounts redeemable on fuel, but the money spent on fuel further earns discounts redeemable on groceries. This, of course, gives consumers an added incentive to increase consumption in both markets, and is likely to further increase total profits. However, under diminishing marginal returns from consumption, if the foodperks! program operates in addition to the fuelperks! program, its marginal impact on profits might not be as great as the impact of the original fuelperks! program. While in this paper we limit ourselves to fuelperks! only, a joint analysis of fuelperks! and foodperks! is a fruitful opportunity for future work.

Chapter 4

Inferring Competitor Pricing with Incomplete Information

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 its maturity, high level of competition and a high level

of customer service. Distributors tend to employ their own salesforce which typically contacts customers every week, collects orders, arranges deliveries, and provides 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 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 this is an important industry, we are unaware of any academic studies of it.

The wholesale food distributor does not directly observe competitors' prices nor 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 informational is atypical in marketing research and it is helpful to contrast our problem against competitive inference problems in data rich

environments. 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 short (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 (see for example Webster, 1978). 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. In our empirical setting,

small adjustments in quantities required are needed for each purchase occasion to satisfy own demand and balance short term inventories. Also, variations in the product mix 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 (Puto, Patton and King 1985) 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 correspond to the practice where a company split the production between different firms (Lyon 2006), and it is 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, the most commonly being 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 aims to contribute to the relatively scarce empirical literature on industrial buyer

behavior, where most of the research is either based on survey data or mainly descriptive (see for example Noordewier et al, 1990 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). While the purchase of one product might motivate a final consumer to buy a complementary product in a shopping trip, this is less likely to occur for industrial buyers with well defined material requirements. Another difference comes when studying purchase incidence. While final customers 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 fact, in both cases 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. 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. Gladys, and Croux (2009) propose an 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 strategically... 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 a minimum level of raw materials are available for production (Puto, Patton and King 1985).

$$\begin{aligned} \min \quad & TC(\{q_{imt}, q_{imt}^c\}_{m=1}^M) \\ \text{s.t.} \quad & q_{imt} + q_{imt}^c \geq \tau_{imt} \quad t = 1 \dots T, m = 1 \dots M \end{aligned} \quad (1.1)$$

Where q_{imt} and q_{imt}^c are the quantity purchased by customer i in category m in period t from the focal and the competitor companies respectively and τ_{imt} 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:

$$TC(\{q_{imt}, q_{imt}^c\}_{m=1}^M) = \sum_{m=1}^M (p_{imt}q_{imt} + p_{imt}^c q_{imt}^c - \psi_{imt} q_{imt} q_{imt}^c) \quad (1.2)$$

where parameters ψ_{imt} capture customer propensity to have multiple suppliers. We impose the condition: $\psi_{imt} > 0$, 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 in which all purchases are made at one supplier or another than this interaction term vanishes and we are left with the usual cost minimization problem.

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 section 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 (1.1). Given our assumptions about the cost function, the problem is separable between categories. Let λ_{imt} be the shadow price associated to the minimum

quantity constraints. Then, the optimal solution is characterized by the following system of equations.

$$p_{imt} - \psi_{imt} q_{imt}^c - \lambda_{imt} \leq 0, \quad \text{with equality if } q_{imt} > 0 \quad (1.3)$$

$$p_{imt}^c - \psi_{imt} q_{imt} - \lambda_{imt} \leq 0, \quad \text{with equality if } q_{imt}^c > 0 \quad (1.4)$$

$$q_{imt} + q_{imt}^c = \tau_{imt} \quad (1.5)$$

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 ($q_{imt}^c = \tau_{imt}$) and (c) the interior solution where customer buys from the focal and competitor firms in which case demands are given by:

$$q_{imt} = \frac{1}{2} \left(\tau_{imt} - \frac{\delta_{imt}}{\psi_{imt}} \right) \quad q_{imt}^c = \frac{1}{2} \left(\tau_{imt} + \frac{\delta_{imt}}{\psi_{imt}} \right) \quad (1.6)$$

where $\delta_{imt} = p_{imt} - p_{imt}^c$.

We point out that our solution has a similarity with Kim et al (2002) in that we both 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 ψ_{imt} . 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 Ω_1 and Ω_2 as the regions in the parameter space where the customer buy from the

competitor and focal firm respectively while the set Ω_3 correspond to the case where the customer buys from both.

$$\begin{aligned}
 \Omega_1 &= \left\{ (\tau_{imt}, \delta_{imt}, \psi_{imt}) : \psi_{imt} \leq \delta_{imt} / \tau_{imt} \right\} \\
 \Omega_2 &= \left\{ (\tau_{imt}, \delta_{imt}, \psi_{imt}) : \psi_{imt} \leq -\delta_{imt} / \tau_{imt} \right\} \\
 \Omega_3 &= \left\{ (\tau_{imt}, \delta_{imt}, \psi_{imt}) : \psi_{imt} \geq |\delta_{imt} / \tau_{imt}| \right\}
 \end{aligned} \tag{1.7}$$

A graphical depiction of the regions where each alternative is chosen in a plane $\delta_{imt} - \psi_{imt}$ is shown in Figure 1. Remember we impose a positivity condition on ψ_{imt} 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 ψ_{imt} 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 $\psi_{imt} \rightarrow 0$, 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.

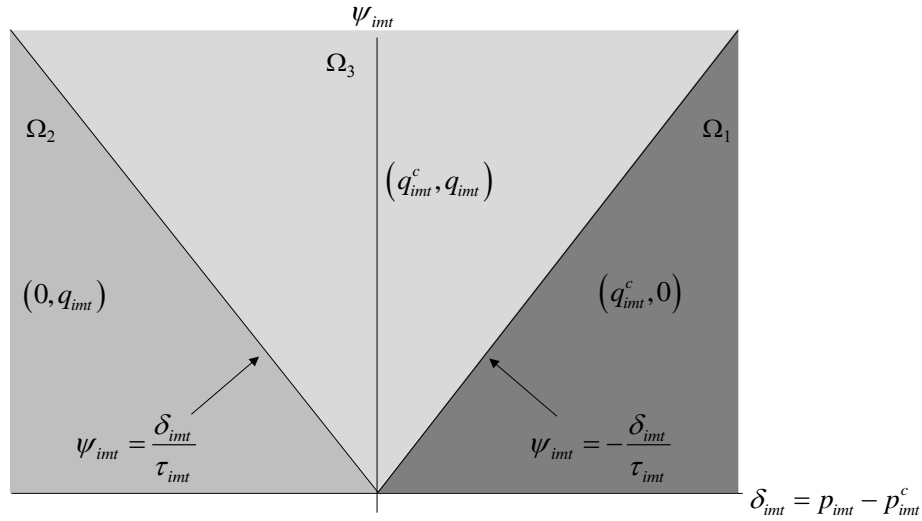


Figure 1: Graphical description of first order condition cases

Therefore, the conditional likelihood of the observed demand q_{imt} is defined by parts as follows:

$$q_{imt} = \begin{cases} \tau_{imt} & \text{if } (\tau_{imt}, \delta_{imt}, \psi_{imt}) \in \Omega_2 \\ \frac{1}{2} \left(\tau_{imt} - \frac{\delta_{imt}}{\psi_{imt}} \right) & \text{if } (\tau_{imt}, \delta_{imt}, \psi_{imt}) \in \Omega_3 \\ 0 & \text{if } (\tau_{imt}, \delta_{imt}, \psi_{imt}) \in \Omega_1 \end{cases} \quad (1.8)$$

The probability of being in each region of the parameter space (Ω_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 $\psi_{imt} = \psi_i$ is constant in time and across categories. On the other hand, we assume that both the required quantity (τ_{imt}) and the price difference (δ_{imt}) are described by the following regression equations.

$$\tau_{imt} = \bar{\tau}_{imt}(\beta_i, XT_{imt}) + \varepsilon_{1imt} \quad (1.9)$$

$$\delta_{imt} = \bar{\delta}_{imt}(\gamma_i, XD_{imt}) + \varepsilon_{2imt} \quad (1.10)$$

Where XT_{imt} and XD_{imt} are matrices of observable covariates, β_i and γ_i are parameters to be estimated and ε_{1imt} and ε_{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 τ_{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 ε_{1imt} and ε_{2imt} are normally distributed with 0 mean and variances σ_{1i}^2 and σ_{2i}^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):

$$\begin{aligned}
p\left(\Omega_1|\bar{\tau}_{imt},\bar{\delta}_{imt},\psi_i,\sigma_{1i}^2,\sigma_{2i}^2\right) &= 1 - \Phi\left(\bar{\tau}_{imt}\psi_i - \bar{\delta}_{imt}/\sigma_{1i}\sigma_{2i}\sqrt{\psi_i^2/\sigma_{2i}^2 - 1/\sigma_{1i}^2}\right) \\
p\left(\Omega_2|\bar{\tau}_{imt},\bar{\delta}_{imt},\psi_i,\sigma_{1i}^2,\sigma_{2i}^2\right) &= \Phi\left(-\bar{\tau}_{imt}\psi_i - \bar{\delta}_{imt}/\sigma_{1i}\sigma_{2i}\sqrt{\psi_i^2/\sigma_{2i}^2 - 1/\sigma_{1i}^2}\right) \\
p\left(\Omega_3|\bar{\tau}_{imt},\bar{\delta}_{imt},\psi_i,\sigma_{1i}^2,\sigma_{2i}^2\right) &= 1 - p\left(\Omega_2|\bar{\tau}_{imt},\bar{\delta}_{imt},\psi_i,\sigma_{1i}^2,\sigma_{2i}^2\right) - p\left(\Omega_3|\bar{\tau}_{imt},\bar{\delta}_{imt},\psi_i,\sigma_{1i}^2,\sigma_{2i}^2\right)
\end{aligned} \tag{1.11}$$

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

$$p\left(\{q_{imt}\}_{i,m,t}|\{\theta_i,\sigma_{1i}^2,\sigma_{2i}^2\}_i\right) = \prod_{i=1}^I \prod_{m=1}^M \prod_{t=1}^T \sum_{k=1}^3 p\left(q_{imt}|\Omega_k\right) p\left(\Omega_k|\theta_i,\sigma_{1i}^2,\sigma_{2i}^2\right) \tag{1.12}$$

To complete the model, we introduce customer heterogeneity into the model by specifying that the parameters of the models come from a common population distribution. Let $\theta_i = (\text{vec}(\beta_i), \text{vec}(\gamma_i), \log(\psi_i))$ be the vector of individual parameters.

$$\theta_i = \Lambda \cdot z_i + v_i \quad v_i \sim N(0, V_\theta) \tag{1.13}$$

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:

$$V_\theta \sim IW(\nu, V) \tag{1.14}$$

$$\text{vec}(\Lambda)|V_\theta \sim N(\text{vec}(\bar{\Lambda}), V_\theta \otimes A^{-1}) \tag{1.15}$$

$$\sigma_{1i}^2 \sim \nu_1 \text{ssq}_1 / \chi_{\nu_1}^2 \quad \sigma_{2i}^2 \sim \nu_2 \text{ssq}_2 / \chi_{\nu_2}^2 \tag{1.16}$$

where $\nu, \nu_1, \nu_2, \text{ssq}_1, \text{ssq}_2, V, \bar{\Lambda}$ 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 τ_{imt} cannot have common elements with the set of regressors used to describe price differences δ_{imt} . If the same regressor is present in both sets, the regression model can face a multicollinearity problem. Notice that in the interior solution (Ω_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_{imt} . To facilitate interpretation of posterior estimates we also normalize all variables in XD_{imt} resulting in price differences centered around zero.
2. The variance of price differences σ_{2i}^2 and the split purchase parameter ψ_i is not jointly identified. The reason is that under the interior solution condition the term $1/\psi_i$ scales the price differences δ_{imt} . 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 δ_{imt}/ψ_i . Alternatively we could use our knowledge of price dispersion to impose tight priors on σ_{2i}^2 . 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 $V(\delta_{imt} = p_{imt} - p_{imt}^c) = 2\sigma_{pim}^2$ where σ_{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 ψ_{imt} . For example, if we want to include price discounts and propose they vary linearly with firm size, then ψ_{imt} 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 its 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_{imkt}) and the negotiated price (p_{imkt}). Unfortunately, neither prices (p_{imkt}^c) nor demanded quantities (q_{imkt}^c) 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 costumers even for the same period. Figure 2 plots time series of prices actually paid for the same product by a sample of 5 customers. 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.

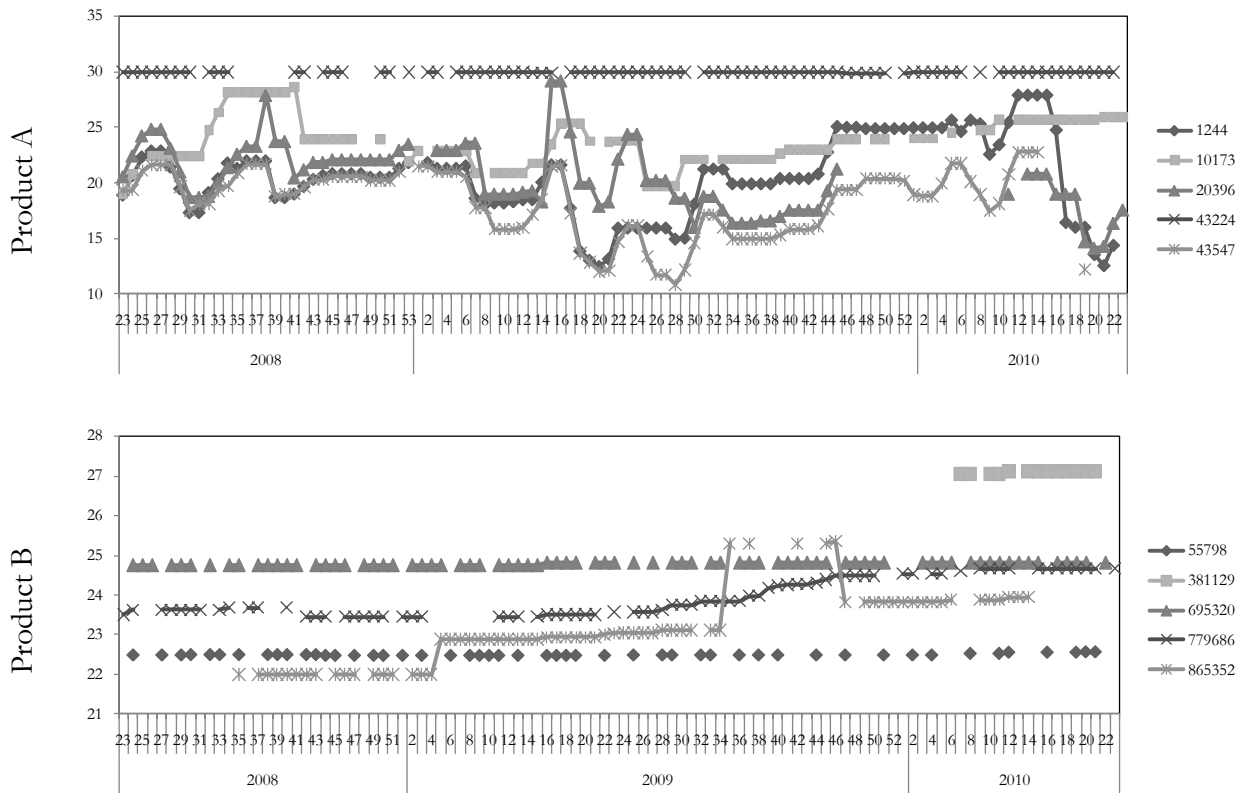


Figure 2: Time Series of paid price for two representative products and a sample of 5 customers

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 1.

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

Table 1: 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 section 3 in Technical Appendix 2). In the resulting data set we select a random sample of costumers 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 (1.9) and **Error! Reference source not found..** For simplicity we use linear models for both regressions:

$$\tau_{imt} = XT'_{imt}\beta_i + \varepsilon_{1imt} \quad (1.17)$$

$$\delta_{imt} = XD'_{imt}\gamma_i + \varepsilon_{2imt} \quad (1.18)$$

Our set of regressors in XT_{imt} 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 XD_{imt} 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 description of the sampler and the values of priors we use in the estimation see section 2 in Technical 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 costumers. 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. 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.

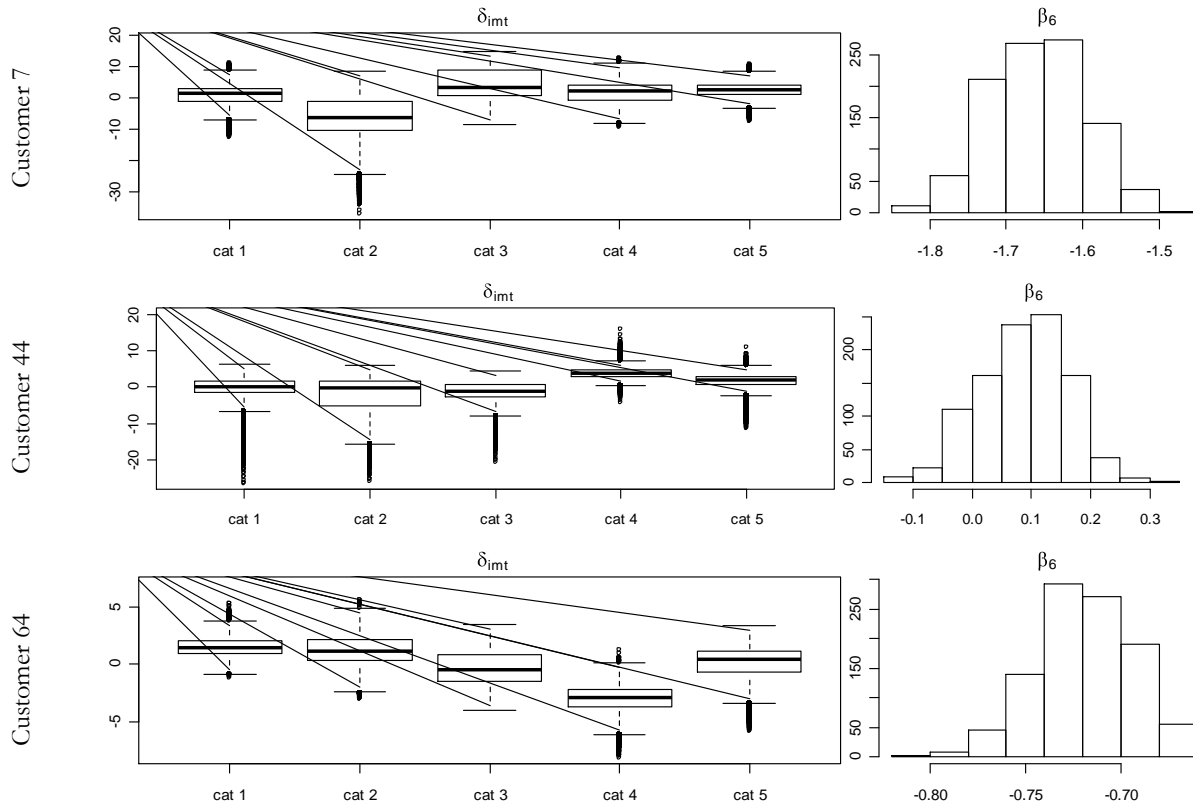


Figure 3: Individual level parameters for a sample of three customers.

As we discussed in section 3 each customer could be in three different states: buying from the competitor only (Ω_1), buying from the focal supplier only (Ω_2) or splitting his demand between the two (Ω_3). 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 4: Time series of posterior means of the probabilities of being in each region of the parameter space for a sample of customers

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.

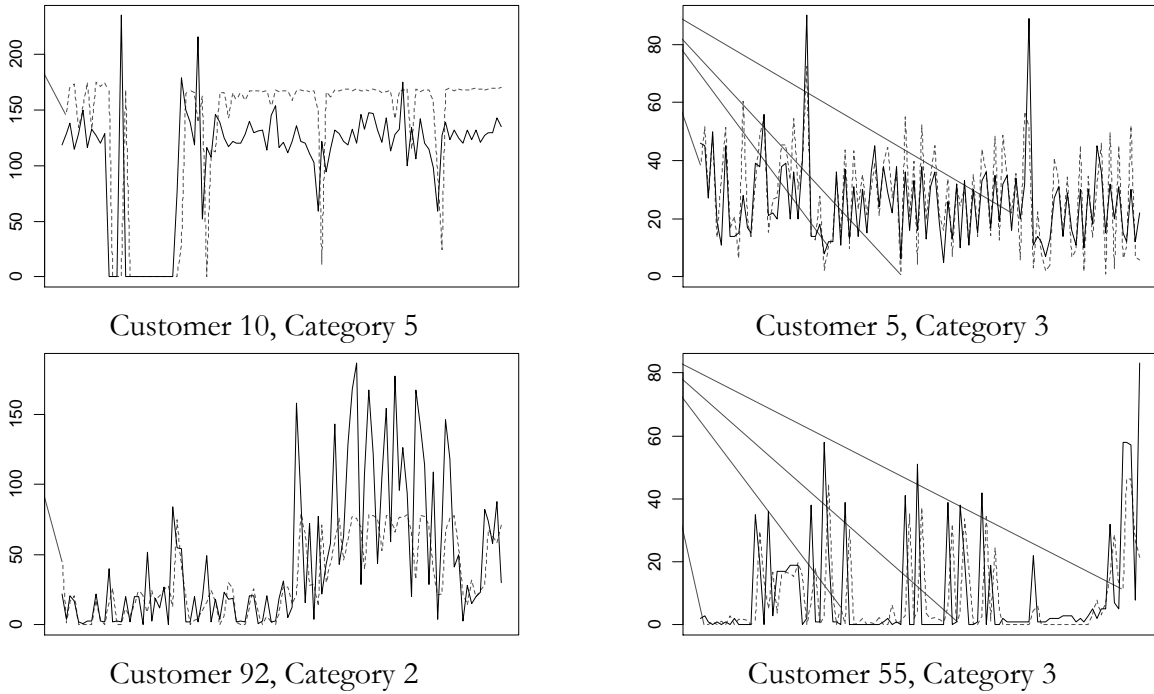


Figure 5: Time series of actual (solid black lines) and predicted (dotted graylines) sales

We now discuss the posterior estimates at the population level parameters Λ 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 2 display posterior means, standard deviations and 90% credible intervals for all parameters in Λ .

Parameters in the upper right cells (rows β_1 - β_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.

		Type A	Type B	Type C	Type D	IsLargest	NEmployee
β_1	Mean (s.d)	65.2 (48.2)	164 (61.2)	12.6 (114)	222 (70.9)	0.767 (0.561)	53.2 (45.7)
	90% C.I	(-28.7, 156)	(42.4, 289)	(-204, 239)	(79.4, 361)	(-0.306, 1.82)	(-31.8, 142)
β_2	Mean (s.d)	158 (44.4)	178 (56.5)	108 (106)	133 (63.9)	0.581 (0.506)	-5.46 (42.3)
	90% C.I	(69.3, 248)	(69.3, 296)	(-105, 319)	(7.31, 256)	(-0.448, 1.59)	(-88.4, 81.6)
β_3	Mean (s.d)	42.9 (53.9)	96.7 (67.1)	-9.7 (117)	19.4 (78.4)	2.17 (0.631)	51.7 (52.4)
	90% C.I	(-61.9, 150)	(-38.6, 223)	(-232, 217)	(-129, 168)	(0.912, 3.46)	(-50.1, 155)
β_4	Mean (s.d)	85.8 (30.5)	138 (38.8)	26.5 (72.2)	71.5 (44.4)	0.724 (0.354)	31.4 (29.0)
	90% C.I	(26.9, 145)	(62.7, 216)	(-119, 168)	(-14.4, 157)	(0.00101, 1.41)	(-24.8, 86.4)
β_5	Mean (s.d)	101 (29.5)	150 (37.4)	51.6 (70.1)	82 (42.7)	0.555 (0.341)	23.1 (28)
	90% C.I	(44.7, 158)	(75.2, 222)	(-85.4, 188)	(-0.144, 164)	(-0.157, 1.21)	(-33.1, 75.8)
β_6	Mean (s.d)	0.0797 (0.69)	-2.49 (0.85)	0.561 (1.59)	1.06 (1.03)	-0.011 (0.00776)	-0.978 (0.667)
	90% C.I	(-1.26, 1.5)	(-4.11, -0.867)	(-2.7, 3.8)	(-0.913, 3.09)	(-0.0272, 0.0034)	(-2.33, 0.254)
γ_1	Mean (s.d)	-31.9 (37.1)	-2.97 (48.3)	-15.0 (87.3)	-83.2 (55.2)	-1.12 (0.436)	-23.6 (35.4)
	90% C.I	(-106, 41.1)	(-102, 91.5)	(-186, 154)	(-194, 22.5)	(-1.93, -0.231)	(-92.3, 44.5)
γ_2	Mean (s.d)	11.5 (15.1)	11.0 (18.8)	18.3 (33.7)	-27.8 (21.4)	-0.494 (0.169)	-2.44 (14.5)
	90% C.I	(-18.7, 41.7)	(-28.6, 47.1)	(-52.2, 78.6)	(-69.3, 12.0)	(-0.82, -0.152)	(-31.9, 26.2)

Table 2: 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 δ_{imt} 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 shade light on the competitive elements of demand that cannot be study 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 costumers 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 segment 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.

Third, 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. Fourth, in our specification we have assumed that price differences 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

- Abraham, Magid and Leonard Lodish (1993) "An Implemented System for Improving Promotion Productivity Using Store Scanner Data", *Marketing Science*, 12 (3), 248-269.
- Allenby, Greg M. and Peter E. Rossi (1993), "A Marginal-Predictive Approach to Identifying Household Parameters", *Marketing Letters*, 4, 227-239.
- Boatwright, Peter, Peter Rossi and Robert McCulloch (1999) "Account-Level Modeling for Trade Promotion: An Application of a Constrained Parameter Hierarchical Model" *Journal of the American Statistical Association*, 94, 1063-1073.
- Box, George E. P. and George C. Tiao (1973), *Bayesian Inference in Statistical Analysis*, Wiley-Interscience.
- Chakravarti, Dipankar, Andrew Mitchell and Richard Staelin (1981), "Judgment Based Marketing Decision Models: Problems and Possible Solutions." *Journal of Marketing*, 45 (4), Fall, 13-21.
- Chintagunta, Pradeep, Jean-Pierre Dube and Vishal Singh (2003) "Balancing Profitability and Customer Welfare in a Supermarket Chain", *Quantitative Marketing and Economics*, 1 (1), 111-147.
- Clemen, 1989, "Combining Forecasts – a Review and Annotated-Bibliography", *International Journal of Forecasting*, 5 (4), 559-583
- Deng, Shiming and Candace A. Yano (2006), "Joint Production and Pricing Decisions with Setup Costs and Capacity Constraints", *Management Science*, 52 (5), 741-756.
- Essegai, Skander, Sunil Gupta, and Z. John Zhang (2002), "Pricing Access Services", *Marketing Science*, 21 (2).
- Fox, Edward J., Steven Postrel, and John H. Semple (2009), "Optimal Category Pricing with Endogenous Store Traffic", *Marketing Science*, 28 (4), 709-722.
- Garthwaite, Paul H., Joseph B. Kadane and Anthony O'hagan (2005) "Statistical Methods for Eliciting Probability Distributions" *Journal of the American Statistical Association*, 100, 470.
- Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin (2003), *Bayesian Data Analysis*, London: Chapman and Hall.
- Huber, Joel. 1975. Bootstrapping of Data and Decisions. *The Journal of Consumer Research* 2, no. 3 (December): 229-234.
- Kadane, Joseph B., James M. Dickey, Robert L. Winkler, Wayne S. Smith and Stephen C. Peters (1980) "Interactive Elicitation of Opinion for a Normal Linear Model" *Journal of the American Statistical Association*, 75, 372.
- Kadiyali, Vrinda, Pradeep Chintagunta and Naufel Vilcassim (2000) "Manufacturer-Retailer Channel Interactions and Implications for Channel Power: An Empirical Investigation of Pricing in a Local Market" *Marketing Science*, 19 (2), 127-148
- Kalyanam, Kirthi (1996), "Pricing Decisions Under Demand Uncertainty: A Bayesian Mixture Model Approach", *Marketing Science*, 15 (3), 207-221.

- Khan, Romana J. and Dipak C. Jain (2005), "An Empirical Analysis of Price Discrimination Mechanisms and Retailer Profitability", *Journal of Marketing Research*, 42 (4), 516-.
- Kopalle, Praveen, Carl Mela and Lawrence Marsh (1999) "The Dynamic Effect of Discounting on Sales: Empirical Analysis and Normative". *Marketing Science*, 18 (3), 317-332.
- Little, John D. C. and Jeremy F. Shapiro (1980), "A Theory for Pricing Nonfeatured Products in Supermarkets" *The Journal of Business*, 53 (3), part 2.
- Montgomery, Alan L. (1997), "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data", *Marketing Science*, 16 (4), 315-337.
- Montgomery, Alan L. and Eric T. Bradlow (1999), "Why Analyst Overconfidence About the Functional Form of Demand Models Can Lead to Overpricing". *Marketing Science*, 18 (4), 569-583.
- Montgomery, Alan L. and Peter E. Rossi (1999), "Estimating Price Elasticities with Theory-based Priors", *Journal of Marketing Research*, 36 (4), 413-423.
- Mulhern Francis J and Robert Leone (1991), "Implicit Price Bundling of Retail Products: A Multiproduct Approach to Maximizing Store Profitability" *Journal of Marketing*, 55 (October), 63-76
- Reibstein, David and Hubert Gatignon (1984), "Optimal Product Line Pricing: The Influence of Elasticities and Cross Elasticities". *Journal of Marketing Research*, 21 (August), 259-67.
- Robert, Christian P. and George Casella (2004), *Monte Carlo Statistical Methods*, Springer, Second Edition.
- Sandor, Zsolt and Michel Wedel (2001), "Designing conjoint choice experiments using managers' prior beliefs", *Journal of Marketing Research*, 38 (4), 430-444.
- Tellis, Gerald J. (1988), "The price elasticity of selective demand: A meta-analysis of econometric models of sales", *Journal of Marketing Research*, 25, 331-341.
- Wierenga, Berend, Gerrit H. Van Bruggen, and Richard Staelin (1999), "The Success of Marketing Management Support Systems", *Marketing Science*, 18 (3), 196-207.
- Wittink, Dick R., M. J. Addona, W.J. Hawkes, and J. C. Porter (1988), "SCAN*PRO: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data", Working Paper, Cornell University.
- Zellner, Arnold (1971), *An Introduction to Bayesian Inference in Economics*, Wiley-Interscience.
-
- Adams, William James and Janet L. Yellen (1976), "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 90(3), 475--498.
- Ailawadi, Kusum and Scott A. Neslin (1998), "The Effect of Promotion on Consumption," *Journal of Marketing Research*, XXXV(August), 390--398.
- Arya, Anil, Brian Mittendorf and Dae-Hee Yoon (2008), "Friction in Related-Party Trade When a Rival Is Also a Customer," *Management Science*, 54(11), 1850--1860.

- Bell, David R., Ganesh Iyer and V. Padmanabhan (2002), "Price Competition Under Stockpiling and Flexible Consumption," *Journal of Marketing Research*, XXXIX(August), 292--303.
- Brickman, Philip and Donald T. Campbell (1971), "Hedonic Relativism and Planning the Good Society," in *Adaptation-Level Theory*, ed. M. H. Appley, New York: Academic Press.
- CES (2009), "Consumer Expenditures in 2007," *U.S. Bureau of Labor Statistics*, Report 1016, April 2009.
- Chandon, Pierre and Brian Wansink (2002), "When Are Stockpiled Products Consumed Faster? A Convenience-Salience Framework of Postpurchase Consumption Incidence and Quantity," *Journal of Marketing Research*, XXXIX(August), 321--335.
- Coombs, Clyde H. and George S. Avrunin (1977), "Single-Peaked Functions and the Theory of Preference," *Psychological Review*, 84(2), 216--230.
- Dhar, Sanjay and Jagmohan S. Raju (1998), "The Effects of Cross-Ruff Coupons on Sales and Profits," *Management Science*, 44(11), Part 1 of 2, 1501--1516.
- Excentus (2010), "Excentus Corporation and Safeway Inc. Settle Lawsuit and Enter Into Patent Licensing Agreement," <http://www.excentus.com/index.php?mod=content\&id=138\&p=1>. [Last accessed: May 2, 2010.]
- Gans, Joshua S. and Stephen P. King (2004), "Supermarkets and Shopper Dockets: The Australian Experience," *Australian Economic Review*, 37(3), 311--316.
- Gans, Joshua S. and Stephen P. King (2006), "Paying for Loyalty: Product Bundling in Oligopoly," *Journal of Industrial Economics*, 54(1), 43--62.
- Hess, James D. and Eitan Gerstner (1987), "Loss Leader Pricing and Rain Check Policy," *Marketing Science*, 6(4), 358--374.
- Jerath, Kinshuk and Z. John Zhang (2009), "Store-Within-a-Store," *Journal of Marketing Research*, forthcoming.
- Lal, Rajiv and Carmen Matutes (1994), "Retail Pricing and Advertising Strategies," *Journal of Business*, 67(3), 345--370.
- Lindeman, Teresa F. (2007), "Giant Eagle Bags a Bigger Share of Market," *Pittsburgh Post Gazette*, Feb 14, 2007. <http://www.post-gazette.com/pg/07045/761852-28.stm>.
- Kim, Byung-Do, Mengze Shi and Kannan Srinivasan (2001), "Reward Programs and Tacit Collusion," *Marketing Science*, 20(2), 99--120.
- Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green (1995), "Microeconomic Theory," Oxford University Press, Inc., New York, USA.
- Mathewson, Frank and Ralph Winter (1997), "Tying as a Response to Demand Uncertainty," *RAND Journal of Economics*, 28(3), 566--583.
- Matutes, Carmen and Pierre Regibeau (1992), "Compatibility and Bundling of Complementary Goods in a Duopoly," *Journal of Industrial Economics*, 40(1), 37--54.

- McTaggart, Jenny (2006), "Pump Up the Volume," *Progressive Grocer*, September 1, 2006.
- Montgomery, Alan L. and Eric T. Bradlow (1999), "Why Analyst Overconfidence about the Functional Form of Demand Can Lead to Overpricing," *Marketing Science*, 18(4), 569--583.
- Neslin, Scott A. (2002), "Sales Promotion," Marketing Science Institute, Cambridge, MA.
- Oren, Shmuel S., Stephen A. Smith and Robert B. Wilson (1983), "Competitive Nonlinear Tariffs," *Journal of Economic Theory*, 29(1), 49--71.
- PG (2005), "P.A. Attorney General OKs Giant Eagle's fuel perks! Discounts," *Progressive Grocer*, September 7, 2005.
- PG (2006), "Taking the Lead," *Progressive Grocer*, May 1, 2006.
- PG (2009), "Safeway Sued Over Fuel Rewards Program," *Progressive Grocer*, June 18, 2009.
- Robbins, Ted (2008), "Higher Prices Mean Slimmer Profits for Gas Stations," *NPR Morning Edition*, May 15, 2008. <http://www.npr.org/templates/story/story.php?storyId=90453142>.
- Singh, Nirvikar and Xavier Vives (1984), "Price and Quantity Competition in a Differentiated Duopoly," *RAND Journal of Economics*, 15(4), 546--554.
- Slade, Margaret E. (1998), "The Leverage Theory of Tying Revisited: Evidence from Newspaper Advertising," *Southern Economic Journal*, 65(2), 204--222.
- Spence, Michael (1977), "Nonlinear Prices and Welfare," *Journal of Public Economics*, 8(1), 1--18.
- Stremersch, Stefan and Gerard J. Tellis (2002), "Strategic Bundling of Products and Prices: A New Synthesis for Marketing," *Journal of Marketing*, 66(1), 55--72.
- Sun, Baohong (2005), "Promotion Effect on Endogenous Consumption," *Marketing Science*, 24(3), 430--443.
- Tellis, Gerard J. (1986), "Strategic Bundling of Products and Prices: A New Synthesis for Marketing," *Journal of Marketing*, 50(4), 146--160.
- Varian, Hal R. (1992), "Microeconomic Analysis, Third Edition," W. W. Norton & Company, Inc., New York, USA.
- Venkatesh, R. and Wagner Kamakura (2003), "Optimal Bundling and Pricing Under a Monopoly: Contrasting Complements and Substitutes from Independent Valued Products," *Journal of Business*, 76(2), 211--231.
- Wansink, Brian (1996), "Does Package Size Accelerate Usage Volume?" *Journal of Marketing*, 60(3), 1-14.
- Winer, Russell S. (2006), "Pricing," Marketing Science Institute, Cambridge, MA.
-
- Allenby, Greg M. (1989) "A Unified Approach to Identifying, Estimating and Testing Demand Structures with Aggregate Scanner Data," *Marketing Science*, Vol 8, No 3, pp 265-280.
- Anderson, James C. and James A. Narus (2003) "Selectively Pursuing More of Your Customer's Business," *MIT Sloan Management Review*, Vol 44, No 3, pp 42-49

- Andrieu, Christophe and Johannes Thoms (2008) "A tutorial on adaptive MCMC," *Journal Statistics and Computing*, Vol 18, No 4.
- Banker Rajiv D. and Robert J. Kauffman (2004) "The Evolution of Research on Information Systems: A Fiftieth-Year Survey of the Literature in Management Science," *Management Science*, Vol. 50, No 3, pp-281-298.
- Blattberg Robert C. and Kenneth J. Wisniewski (1989) "Price-Induced Patterns of Competition," *Marketing Science*, Vol 8, No 4, pp 291-309.
- Du, Yuxing Rex, Wagner A. Kamakura and Carl F. Mela (2007) "Size and Share of Customer Wallet," *Journal of Marketing*, Vol 71, No 2
- Fox, Edward J. and Jacquelyn S. Thomas (2006) "A Hierarchical Bayesian Approach to Predicting Retail Customers' Share-of-Wallet Loyalty," SMU Cox School of Business Research Paper Series No. 07-003
- First Research (2010) "Food Distributors – Quartely Update 7/26/2010", *First Research Industry profiles*.
- Glady, Nicolas and Christopher Croux (2009) "Predicting Customer Wallet Without Survey Data," *Journal of Service Research*, Vol 11, No 3.
- Haario, Heikki, Eero Saksman and Johanna Tamminen (2001) "An Adaptive Metropolis Algorithm" *Bernoulli*, Vol. 7, No. 2, pp. 223-242
- Hinkley, D. V. (1969) "On the Ratio of Two Correlated Normal Random Variables", *Biometrika*, No 56, Vol 3, pp 635-639
- Keiningham, Timothy L, Tiffany Perkins-Munn and Heather Evans (2003) "The Impact of Customer Satisfaction on Share-of-Wallet in a Business-to-Business Environment," *Journal of Service Research*, Vol 6, No 1.
- Kim, Jaehwan, Greg M. Allenby and Peter E. Rossi (2002), "Modeling Consumer Demand for Variety", *Marketing Science*, Vol 21, No. 3, pp 229-250.
- Klotz, Dorothy E. and Kalyan Chatterjee (1995) "Dual Sourcing in Repeated Procurement Competitions," *Management Science*, Vol 41, No 8, pp. 1317-1327
- Lyon, Thomas (2006) "Does Dual Sourcing Lower Procurement Costs?," *The Journal of Industrial Economics*, Vol. 54, No 2, pp 223–252
- Montgomery, Alan L. (1997), "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data", *Marketing Science*, 16 (4), 315-337.
- Moon, Sangkil, Wagner A. Kamakura and Johannes Ledolter (2007) "Estimating Promotion Response When Competitive Promotions Are Unobservable," *Journal of Marketing Research*, Vol 44, No 3, pp 503-515.
- Narasimhan, Chakravarthi (1988) "Competitive Promotional Strategies", *The Journal of Business*, Vol. 61, No. 4, pp. 427-449

- Noordewier, Thomas G., George John and John R. Nevin (1990) "Performance Outcomes of Purchasing Arrangements in Industrial Buyer-Vendor Relationships," *The Journal of Marketing*, Vol. 54, No. 4, pp. 80-93
- Rauyruena, Papassapa and Kenneth E. Miller (2007) "Relationship quality as a predictor of B2B customer loyalty," *Journal of Business Research*, Vol 60, No 1, pp 21-31.
- Seetharaman, P.B, Siddartha Chib, Andrew Ainslie, Peter Boatwright, Tat Chan, Sachin Gupta, Nitin Mehta, Vithala Rao, Andrei Strijnev (2005) "Models of Multi-Category Choice Behavior," *Marketing Letters*, Vol 16, No 3-4, pp 239-254.
- Shugan, Steven (2004) "Endogeneity in Marketing Decision Models," *Marketing Science*, Vol 23, No 1, pp 1-3.
- Silver, Edward A. (1981), "Operations Research in Inventory Management: A Review and Critique" *Operations Research*, Vol. 29, No. 4, Operations Management (Jul. - Aug), pp. 628-645
- Whitin, T. M. (1955) "Inventory Control and Price Theory", *Management Science*, Vol. 2, No. 1, pp. 61-68

Technical Appendix 1:

Cross-Market Discounts

1. Extensions of the monopoly-monopoly scenario

1.1. Analysis with different rates of diminishing marginal utilities

We assume that firm 1 is the monopolist in two markets s and t , selling at prices p_{s1} and p_{t1} respectively²¹. We introduce more general rates of diminishing marginal utilities by assuming that the utility function of a representative consumer is given by

$$\mathcal{U}(q_{s1}, q_{t1}, q_{t2}) = \alpha_s \left(q_{s1} - \psi_s \frac{q_{s1}^2}{2} \right) + \alpha_t \left(q_{t1} - \psi_t \frac{q_{t1}^2}{2} \right)$$

where q_{s1} and q_{t1} denote the quantities demanded in the markets s and t , respectively, and α_s and α_t denote the “importance” of consumption utility in the respective markets. We further assume that $\psi_s, \psi_t > 0$ implying that the marginal utility from consuming an extra unit of this product decreases with consumption in both markets.

We model the game in three stages and we solve for the subgame-perfect equilibrium of the game using backward induction. In Stage 3, given the posted prices and the cross-market discount, the

²¹ We include the subscript 1 for the firm identity to be consistent with the notation in the main paper where we have competitor firms.

consumer decides the optimal quantities to purchase by maximizing her surplus. This gives the following functions for quantities consumed in each market:

$$q_{s1} = \frac{\alpha_t \delta + \alpha_s \alpha_t \psi_t}{\alpha_s \alpha_t \psi_s \psi_t - \delta^2} - \frac{\delta}{\alpha_s \alpha_t \psi_s \psi_t - \delta^2} p_{t1} - \frac{\alpha_t \psi_t}{\alpha_s \alpha_t \psi_s \psi_t - \delta^2} p_{s1}$$

$$q_{t1} = \frac{\alpha_s \delta + \alpha_t \alpha_s \psi_s}{\alpha_s \alpha_t \psi_s \psi_t - \delta^2} - \frac{\delta}{\alpha_s \alpha_t \psi_s \psi_t - \delta^2} p_{s1} - \frac{\alpha_s \psi_s}{\alpha_s \alpha_t \psi_s \psi_t - \delta^2} p_{t1}$$

The model is well defined if $\delta < \sqrt{\alpha_s \alpha_t \psi_s \psi_t}$. Notice that if $\delta = 0$, the demand functions are given by $q_{s1} = 1/\psi_s - 1/(\alpha_s \psi_s) p_{s1}$ and $q_{t1} = 1/\psi_t - 1/(\alpha_t \psi_t) p_{t1}$, and the quantities demanded in the two markets are completely independent. In this case, we note that as the rate of satiation becomes larger, the base demand becomes smaller and the demand becomes less price sensitive in the market in question, all else equal. In Stage 2, the firm foresees the above response by the consumers and sets the optimal posted prices, p_{s1} and p_{t1} , by solving its profit maximization problem. The prices and the corresponding quantities consumed are given in Table TA1.

PRICES	
p_{s1}	$\frac{\alpha_s \alpha_t \psi_s (2\alpha_s \psi_t + \delta)}{4\alpha_s \alpha_t \psi_s \psi_t - \delta^2}$
p_{t1}	$\frac{\alpha_s \alpha_t (2\alpha_t \psi_s + \delta) \psi_t}{4\alpha_s \alpha_t \psi_s \psi_t - \delta^2}$
$p_{s1} - \delta q_{s1}$	$\frac{\alpha_t (2\alpha_s \alpha_t \psi_s \psi_t - \alpha_s \psi_t \delta - \delta^2)}{4\alpha_s \alpha_t \psi_s \psi_t - \delta^2}$
QUANTITIES	
q_{s1}	$\frac{\alpha_t (2\alpha_s \psi_t + \delta)}{4\alpha_s \psi_s \psi_t - \delta^2}$
q_{t1}	$\frac{\alpha_s (2\alpha_t \psi_s + \delta)}{4\alpha_s \alpha_t \psi_s \psi_t - \delta^2}$

Table TA1: Prices and quantities in Stage 2 in the monopoly-monopoly scenario.

To formally evaluate how the size of cross-market discount affects prices, demand and profits, we inspect the partial derivatives with respect to δ .

$$\begin{aligned}\frac{\partial p_{s1}}{\partial \delta} &= \frac{\alpha_s \alpha_t \psi_s (\delta^2 + 4\alpha_s (\delta + \alpha_t) \psi_t)}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2} \\ \frac{\partial p_{t1}}{\partial \delta} &= \frac{\alpha_s \alpha_t \psi_t (\delta^2 + 4\alpha_t (\delta + \alpha_s) \psi_s)}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2} \\ \frac{\partial q_{s1}}{\partial \delta} &= \frac{\alpha_t (\delta^2 + 4\alpha_s (\delta + \alpha_s \psi_s) \psi_t)}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2} \\ \frac{\partial q_{t1}}{\partial \delta} &= \frac{\alpha_s (\delta^2 + 4\alpha_t (\delta + \alpha_s \psi_t) \psi_s)}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2} \\ \frac{\partial \Pi_{st1}}{\partial \delta} &= \frac{\alpha_s \alpha_t (\delta + 2\alpha_t \psi_s) (\delta + 2\alpha_s \psi_t)}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2}\end{aligned}$$

Given $\alpha_s, \alpha_t, \psi_s, \psi_t > 0$, we verify that posted prices and purchased quantities are increasing in δ and, therefore, the total profit of the firm is monotonically increasing in δ . (Notice that the rates of variation of prices, quantities and total profit as functions of δ are decreasing in ψ_s and ψ_t). To find the optimal rate of the cross-market discount, the firm will keep increasing δ while ensuring two constraints --- nonnegative consumer surplus and nonnegative effective price in target market. It turns out that both consumer surplus and effective price in the target market are decreasing in δ (as shown below) and, therefore, the optimal is achieved when one of those constraints is binding.

$$\begin{aligned}\frac{\partial CS}{\partial \delta} &= \frac{\alpha_s \alpha_t (\delta^3 + 12\alpha_s \alpha_t \delta \psi_s \psi_t + 3\delta^2 (\alpha_t \psi_t + \alpha_s \psi_s) + 4\alpha_s \alpha_t \psi_s \psi_t (\alpha_t \psi_s + \alpha_s \psi_t))}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^3} \\ \frac{\partial (p_{t1} - \delta q_{s1})}{\partial \delta} &= -\frac{\alpha_s \alpha_t \psi_t (\delta^2 + 4\alpha_t \delta \psi_s + 4\alpha_s \alpha_t \psi_s \psi_t)}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2}\end{aligned}$$

From the above, notice that the rates at which consumer surplus and effective price in the target market decrease with δ are decreasing in the rates of satiation. Consequently, larger values of ψ_s and ψ_t expand the set of feasible values of δ . The optimal values of δ that complete the characterization

of the equilibrium depend on which constraint is binding. Let δ^{ep} and δ^{cs} be the smallest positive roots of the equations enforcing the conditions that effective price in target market is zero and consumer surplus across the two markets is zero, respectively. To find the regions of the parameter space corresponding to the case in which the nonnegativity of the effective price in the target market is the binding constraint, we impose $\delta^{ep} \leq \delta^{cs}$.

$$\delta^{ep} = \frac{1}{2} \left(-\alpha_s \psi_t + \sqrt{\alpha_s \psi_t (8\alpha_t \psi_s + \alpha_s \psi_t)} \right)$$

$$\delta^{cs} = \frac{1}{2} \left(-\alpha_s \psi_t - \alpha_t \psi_s + (\Delta_1 + 4\Delta_2)^{1/3} + (\Delta_1 - 4\Delta_2)^{1/3} \right)$$

where

$$\Delta_1 = -\alpha_t^3 \psi_s^3 + 5\alpha_s \alpha_t^2 \psi_s^2 \psi_t + 5\alpha_s^2 \alpha_t \psi_s \psi_t^2 - \alpha_s^3 \psi_t^3$$

$$\Delta_2 = \sqrt{-\alpha_s \alpha_t \psi_s \psi_t (\alpha_t^2 \psi_s^2 - \alpha_s^2 \psi_t^2)^2}$$

To shed further light on these conditions, in Figure TA1 we display for several values of ψ_s and ψ_t the region of values of α_s and α_t where the positive effective price in the target market is the binding condition (the shaded regions). We conclude that, in general, when the target market is more important than the source market, the monopolist can fully extract consumer surplus.

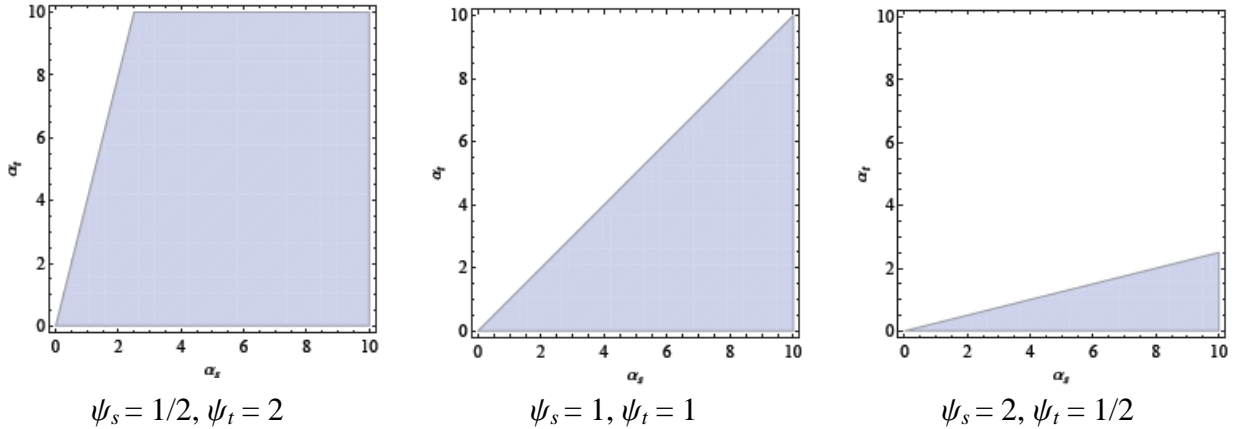


Figure TA1: The shaded regions correspond to the parameter space where the nonnegativity of the effective price in target market is the binding constraint.

1.2. Analysis of quantity discounts limited to the source market

In this section, we present a detailed analysis of the case in which discounts are redeemable in the same market in which they are accumulated. For simplicity, we use the monopoly-monopoly scenario to study this strategy and compare it against cross-market discounts. For this comparison, it is useful to use the extended model with “satiation parameters” $\psi_s > 0$ and $\psi_t > 0$ as we did in 1 in this Technical Appendix. Then, if δ_s denotes the rate of the “self-market discount”, the consumer utility and expenditure are given by

$$\begin{aligned}\mathcal{U}(q_{s1}, q_{t1}, q_{t2}) &= \alpha_s \left(q_{s1} - \psi_s \frac{q_{s1}^2}{2} \right) + \alpha_t \left(q_{t1} - \psi_t \frac{q_{t1}^2}{2} \right) \\ \mathcal{E}(q_{s1}, q_{t1} | p_{s1}, p_{t1}, \delta_s) &= (p_{s1} - \delta_s q_{s1}) q_{s1} + p_{t1} q_{t1}\end{aligned}$$

Notice that under this scheme, the second market t is rather irrelevant. However, we keep it in the analysis to make proper comparisons of profits with the cross-market discount scheme. The profit of the firm is given by $\Pi_{st,1}(p_{s1}, p_{t1}, \delta_s) = (p_{s1} - \delta_s q_{s1}) q_{s1} + p_{t1} q_{t1}$.

As before, we solve the game in three stages and find the equilibrium by backward induction. In Stage 3, consumers maximize their surplus. In Stage 2, the firm sets its posted prices by maximizing its profit given the discount rate δ_s . The resulting prices, demand and profits in Stage 2 are displayed on the left side of Table TA2. We ensure nonnegative and finite prices, demand, consumer surplus and profits. Note that since purchase decisions can be made separately, the relevant consumer surplus is that associated with market s only. This surplus decreases in δ_s while firm profit increases in δ_s . Therefore, the optimal value of δ_s in Stage 3 will be the value at which the surplus becomes zero, which is given by $\delta_s = \alpha_s \psi_s / 2$. The right side of Table TA2 displays the prices, demand and profits at the global optimum.

Quantity	Conditional on δ_s		Global	
	Market s	Market t	Market s	Market t
Price	$\frac{\alpha_s^2 \psi_s}{-2\delta_s + 2\alpha_s \psi_s}$	$\frac{\alpha_t}{2}$	α_s	$\frac{\alpha_t}{2}$
Demand	$\frac{\alpha_s}{-2\delta_s + 2\alpha_s \psi_s}$	$\frac{1}{2\psi_s}$	$\frac{1}{\psi_s}$	$\frac{1}{2\psi_t}$
Profits	$\frac{\alpha_s^2}{-4\delta_s + 4\alpha_s \psi_s}$	$\frac{\alpha_s}{2\psi_s}$	$\frac{\alpha_s}{2\psi_s}$	$\frac{\alpha_t}{4\psi_t}$

Table TA2: Results for self-market discounts.

As is well documented in nonlinear pricing literature, a self-market discount (basically, a quantity discount) could result in higher prices and higher profits than uniform monopoly prices, as we also find in our model. However, the relevant question here is whether a firm will be better off by offering a cross-market discount or a self market discount. We compare against the corresponding cross-market discount monopoly-monopoly model with parameters ψ_s and ψ_t added as we described in 1 in this Technical Appendix. Recall that according to that appendix, the exact expression of optimal profits using cross-market discounts depends on which constraint is binding. For simplicity we restrict our description to the case in which the effective price in the target market is zero. The resulting prices, demand and profits are displayed in Table TA3.

Quantity	Conditional on δ_s		Global	
	Market s	Market t	Market s	Market t
Price	$\frac{\alpha_s \alpha_t \psi_s \delta + 2\alpha_s \psi_t}{-\delta^2 + 4\alpha_s \alpha_t \psi_s \psi_t}$	$\frac{\alpha_s \alpha_t (\delta + 2\alpha_t \psi_s) \psi_t}{-\delta^2 + 4\alpha_s \alpha_t \psi_s \psi_t}$	$\frac{\alpha_s}{4} + \frac{\sqrt{\alpha_s (8\alpha_t \psi_s + \alpha_s \psi_t)}}{4\sqrt{\psi_t}}$	α_t
Demand	$\frac{\alpha_t (\delta + 2\alpha_s \psi_t)}{-\delta^2 + 4\alpha_s \alpha_t \psi_s \psi_t}$	$\frac{\alpha_s (\delta + 2\alpha_t \psi_s)}{-\delta^2 + 4\alpha_s \alpha_t \psi_s \psi_t}$	$\frac{1}{4\psi_s} + \frac{\sqrt{8\alpha_t \psi_s + \alpha_s \psi_t}}{\sqrt{\alpha_s \psi_t}}$	$\frac{1}{\psi_t}$
Profits	$\frac{\alpha_s \alpha_t^2 \psi_s (\delta + 2\alpha_s \psi_t)^2}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2}$	$\frac{\alpha_s \alpha_t (\delta + 2\alpha_t \psi_s) (\delta^2 + \alpha_s (\delta - 2\alpha_t \psi_s) \psi_t)}{(\delta^2 - 4\alpha_s \alpha_t \psi_s \psi_t)^2}$	$\frac{4\alpha_t \psi_s + \alpha_s \psi_t + \sqrt{\alpha_s \psi_t (8\alpha_t \psi_s + \alpha_s \psi_t)}}{8\psi_s \psi_t}$	0

Table TA3: Results for cross-market discounts.

We find that the parameters ψ_s and ψ_t have two effects. They affect the rate at which firm profit increases with the discounts and the domain in which the model is well defined. Given that profits are increasing in δ_s , these boundaries directly affect the global optimum of the model. We find that while the self-market discount scheme is very sensitive to the parameter ψ_s , cross-market discounts can sustain relatively higher profits. This leads to the interesting insight that consumption is greater in cross-market discounts because they distribute the additional consumption (motivated by the price discount) in two markets and, therefore, they can lead to higher profits for the firm. Figure TA2 displays this pattern by plotting profit for both schemes as a function of discounts where the function is well defined, for several values of ψ_s and ψ_t (for clarity of exposition we have fixed $\alpha_s=\alpha_t=1$).

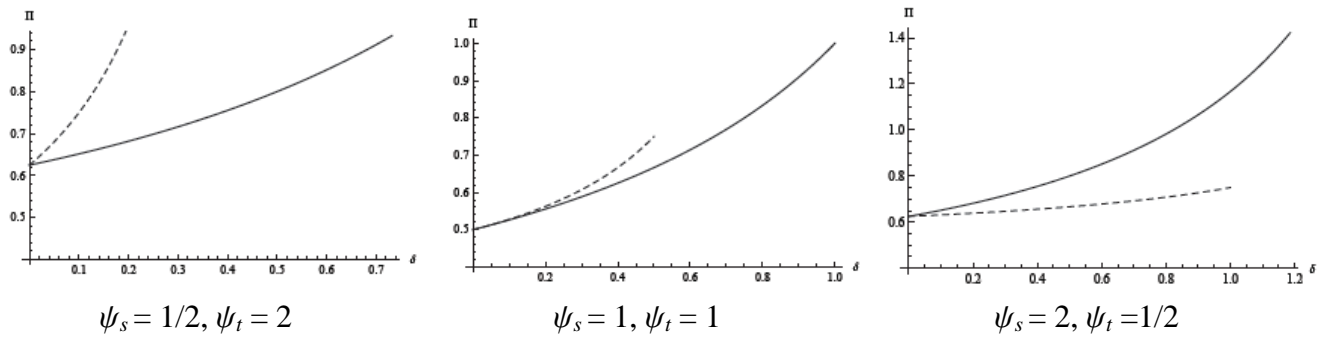


Figure TA2: In the above plots, the solid line shows total profit for the firm from a cross-market discount scheme, and the dashed line shows total profit for the firm from a self-market discount scheme.

We can also verify that a cross-market discount could generate higher profits. To derive conditions under which this happens, we need to compare optimal profits for each strategy. In this case, in addition to the conditions above characterizing the solutions, we note that the condition under which the optimal profit from a cross-market discount strategy ($\Pi_{st,1}^{CMD}$) is greater than the optimal profit from a self-market discount strategy ($\Pi_{st,1}^{SMD}$) is given by

$$\Pi_{st,1}^{CMD} - \Pi_{st,1}^{SMD} = \frac{2\alpha_t\psi_s - 3\alpha_s\psi_t + \sqrt{(\alpha_s\psi_t)^2 + 8\alpha_t\psi_s + \alpha_s\psi_t}}{8\psi_s\psi_t} > 0$$

To gain further insight into this, in Figure TA3 we plot the regions in the ψ_s - ψ_t plane (left panel) and the α_s - α_t plane (right panel) where a cross-market discount is more profitable than a self-market discount. We conclude that a cross-market discount dominates a self-market discount except for the cases in which the source market is significantly more important or saturates at a significantly lower rate than target market. In particular, for equal rates of diminishing returns ($\psi_s = \psi_t = \psi$) and equal market importance ($\alpha_s = \alpha_t = \alpha$), a cross-market discount is always the preferred strategy.

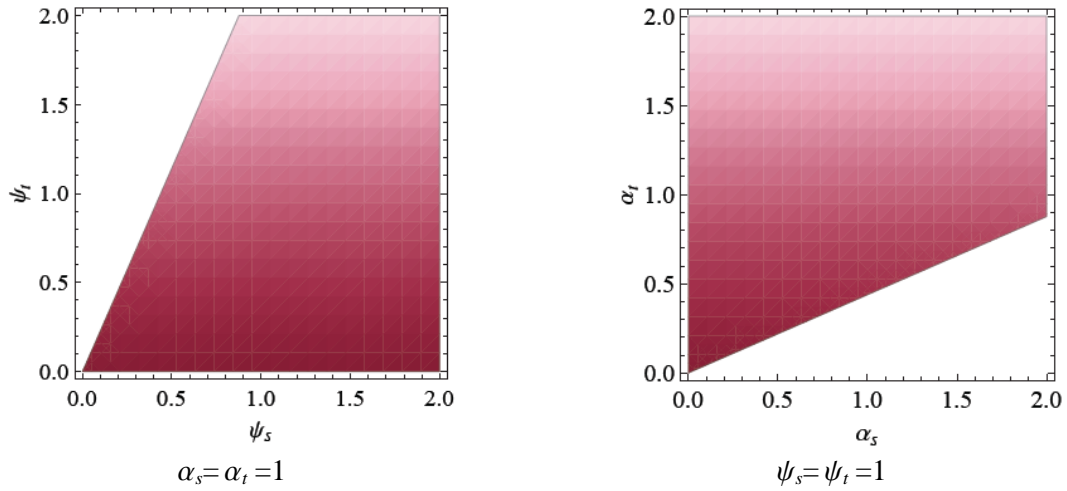


Figure TA3: In the shaded regions, a cross-market discount is preferred over a self-market discount

1.3. Analysis with nonzero marginal cost

In this section, we provide details for the analysis of the monopoly-monopoly scenario with nonzero marginal costs. Let c_s and c_t be the marginal costs for the source and the market target, respectively.

By adding such costs, there is no change in consumer utility maximization problem. The firm's total profit is given by $\Pi_{st,1}(p_{s1}, p_{t1}, \delta) = (p_{s1} - c_s) q_{s1} + (p_{t1} - c_t - \delta q_{s1}) q_{t1}$.

We again solve the model in the same three stages as the basic model in Section **Error! Reference source not found.** in Chapter 3. The resulting prices, demand and profits as functions of δ after Stage 2 are displayed in Table TA4. We verify that the main characteristics of this model are the same as of the basic model --- the optimal posted prices in both markets increase in δ , the optimal effective price in the target market decreases in δ , the quantities demanded in both markets increase in δ , and the total profit of the firm from the two markets increases in δ .

Quantity	Market s	Market t
Price	$c_s + \frac{\alpha_s(-c_t\delta + (-2c_s + 2\alpha_s + \delta))}{4\alpha_s\alpha_t - \delta^2}$	$c_t + \frac{\alpha_t(-c_s\delta + \alpha_t(-2c_t + 2\alpha_t + \delta))}{4\alpha_s\alpha_t - \delta^2}$
Demand	$\frac{2c_s\alpha_t + c_t\delta - \alpha_t(2\alpha_s + \delta)}{-4\alpha_s\alpha_t + \delta^2}$	$\frac{2c_t\alpha_s + c_s\delta - \alpha_s(2\alpha_t + \delta)}{-4\alpha_s\alpha_t + \delta^2}$
Profits	$\frac{\alpha_s(2c_s\alpha_t + c_t\delta - \alpha_t(\delta + 2\alpha_s))^2}{(\delta^2 - 4\alpha_s\alpha_t\psi_s\psi_t)^2}$	$\frac{(2\alpha_s(c_t - \alpha_t)\alpha_t + (-c_s + \alpha_s)\alpha_t\delta + (-c_t + \alpha_t)\delta^2)(2c_t\alpha_s + c_s\delta - \alpha_s(2\alpha_t + \delta))}{(-4\alpha_s\alpha_t + \delta^2)^2}$

Table TA4: Stage 2 results for the monopoly-monopoly scenario with positive marginal costs.

The validity of the model requires positive and finite prices, demand, consumer surplus and profits. Interestingly, given that a cross-market discount couples the demand in the two markets, it could generate positive demand in a market that would not exist independently. By inspection, finite quantities are guaranteed if $\delta_s < \sqrt{\alpha_s\alpha_t}$. To check positiveness, given the model is well defined at $\delta = 0$, we need to only monitor those quantities that decrease as δ increases.

Given that the profit of the focal firm is increasing in δ , and the effective price in target market and the consumer surplus are decreasing in δ , by figuring out which one of the latter two is smaller and limiting it at zero, we can directly identify the optimal discount in Stage 1.

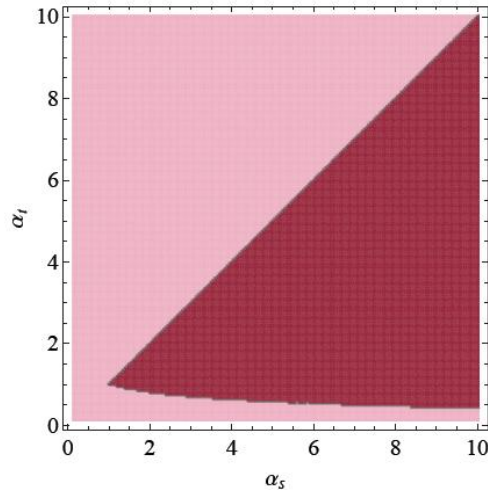


Figure TA4: In the dark-shaded region, the price in the target market is below the marginal cost in this market.

The main new insight that we obtain from this extension is that the firm can find it advantageous to set a cross-market discount large enough that the effective price in the target market is below marginal cost (but still nonnegative). The dark-shaded region in Figure TA4 shows the region in the $\alpha_s - \alpha_t$ plane (with $c_s = c_t = 1$) in which the optimal cross-market discount implies an effective price below marginal cost in the target market. Broadly speaking, when the target market is less important than the source market in the consumption utility function, the firm can have the incentive to take a loss in the target market to make overall higher total profit by selling more in the source market. On the other hand, if the target market is more important than the source market, it is not profitable to sacrifice this source of profit and the firm chooses a less aggressive discount policy.

2. Analysis of the marginal effects of parameters in the monopoly-duopoly scenario

To analyze how each of the parameters of the model affects the cross-market discount decisions, we computed the optimal rate δ for different points of the parameter space. To conduct a comprehensive analysis we need to define a fine grid for each of the parameters of the demand. The

parameter θ_t is naturally bounded between 0 and 1, but the parameters α_s and α_t could take any positive value. Therefore, we reparametrized them as $\alpha_s = a_s / (1 - a_s)$ and $\alpha_t = a_t / (1 - a_t)$, where $a_s, a_t \in [0, 1)$. Therefore, by varying each of θ_t, α_s and α_t on a grid of values between 0 and 1, we can conduct a complete numerical analysis. Figures TA5, TA6 and TA7 display the optimal rate of the cross-market discount as a function of θ_t, α_s and α_t , respectively, for various representative values of the other parameters. We conclude that when competition is present in the target market, the optimal rate of the cross-market discount decreases in the intensity of competition, and increases in the importance of the source and target markets.

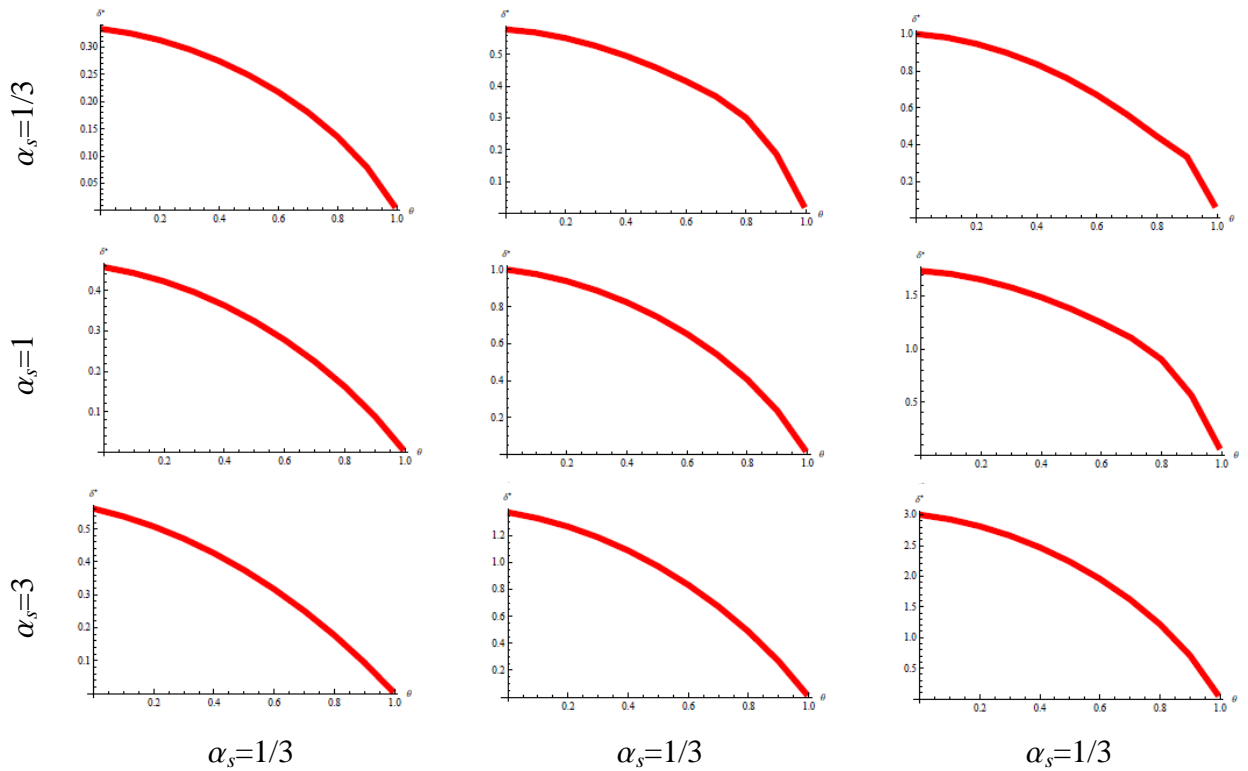


Figure TA5: Optimal rate of the cross-market discount as a function of θ_t

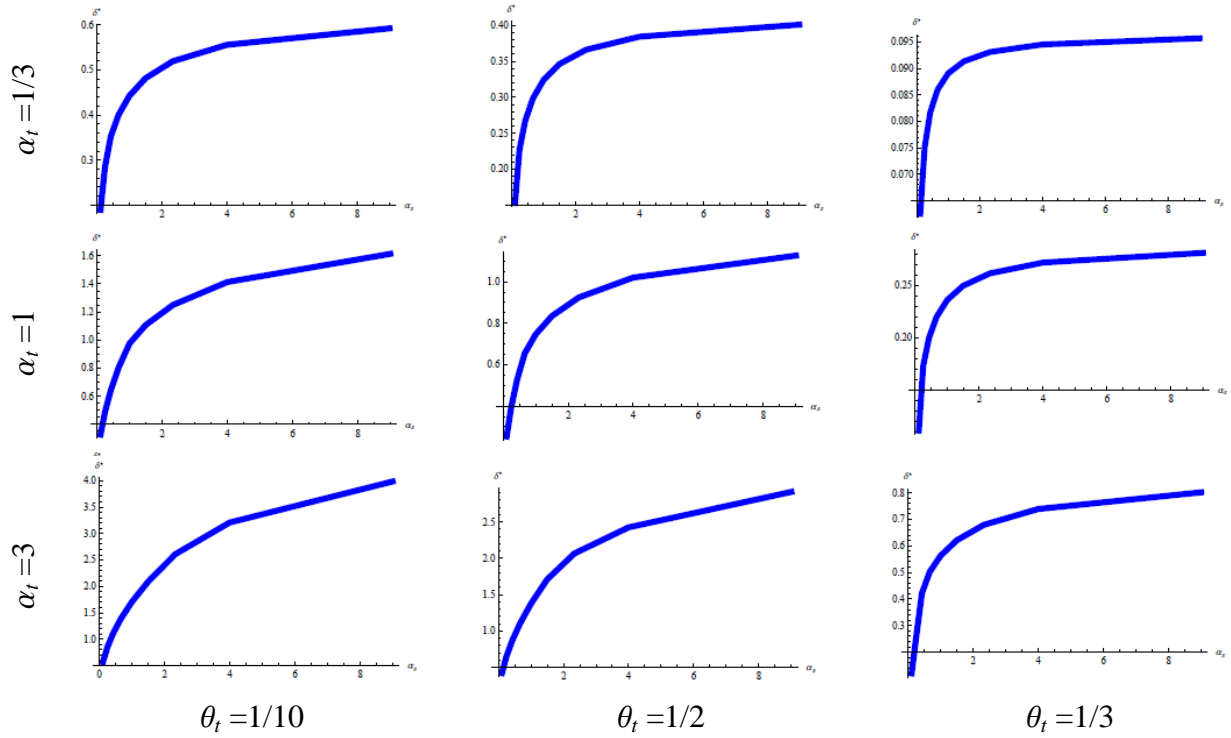


Figure TA6: Optimal rate of the cross-market discount as a function of θ_t

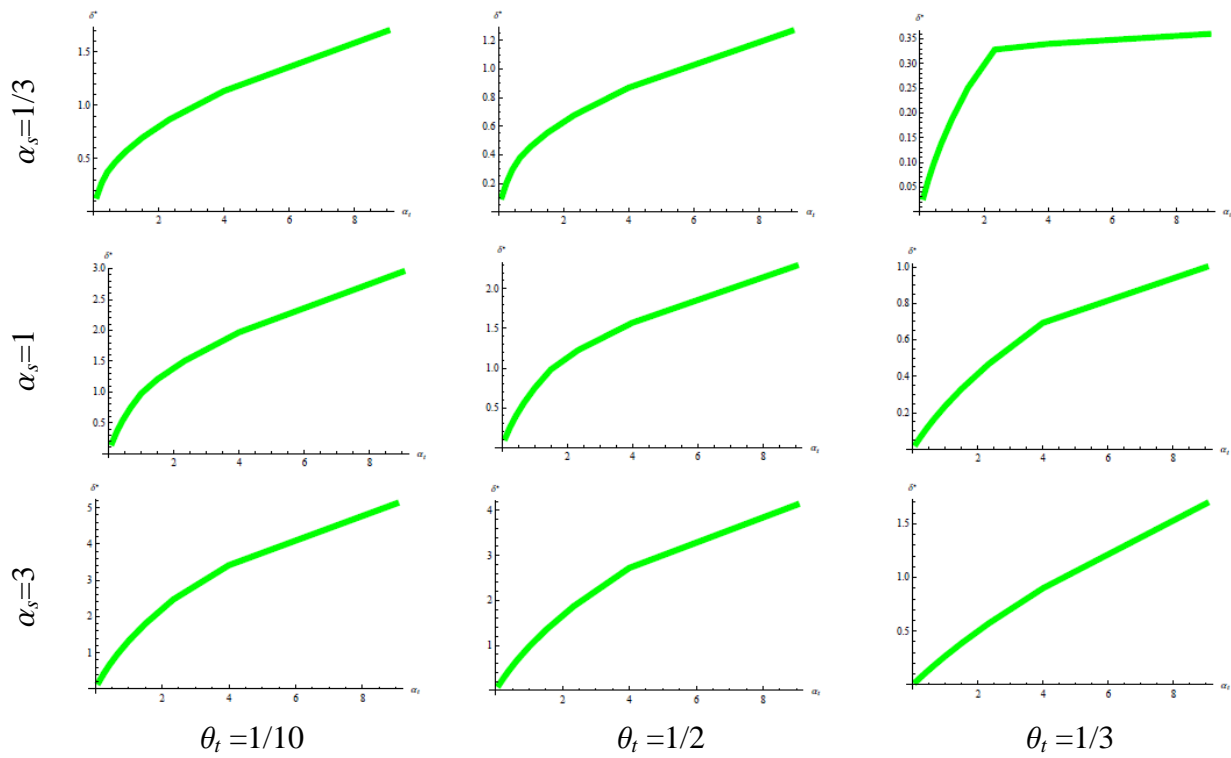


Figure TA7: Optimal rate of the cross-market discount as a function of θ_t

3. Analysis of the duopoly-duopoly scenario

3.1. Without competing cross-market discount program

In this section, we present a detailed analysis of the case in which the focal firm faces competition in both markets. We assume that Firm 1 is the focal firm which operates in both markets, Firm 2 competes in market t , and a new firm, Firm 3, competes in market s . We keep all notation used in previous scenarios and introduce a new parameter θ_s that corresponds to the degree of competition in market s . Also, p_{s3} and q_{s3} denote the price and the quantity demanded from Firm 3 in the source market. Then, in the duopoly-duopoly scenario the consumer utility and expenditure are given by

$$\mathcal{U}(q_{s1}, q_{t1}) = \alpha_s \left(q_{s1} - \frac{q_{s1}^2}{2} + q_{s3} - \frac{q_{s3}^2}{2} - \theta_s q_{s1} q_{s3} \right) + \alpha_t \left(q_{t1} - \frac{q_{t1}^2}{2} + q_{t2} - \frac{q_{t2}^2}{2} - \theta_t q_{t1} q_{t2} \right)$$

$$\mathcal{E}(q_{s1}, q_{s3}, q_{t1}, q_{t2} | p_{s1}, p_{s3}, p_{t1}, p_{t2}, \delta) = p_{s1} q_{s1} + p_{s3} q_{s3} + (p_{t1} - \delta q_{s1}) q_{t1} + p_{t2} q_{t2}$$

Quantities demanded as a function of posted prices and the cross-market discount are obtained by solving the consumer surplus maximization problem and given by

$$q_{s1} = \frac{\alpha_t (p_{s1} - p_{s3} \theta_s + \alpha_s \theta_s - \alpha_s - \delta) + \alpha_t \theta_t^2 (-p_{s1} + p_{s3} \theta_s - \alpha_s \theta_s + \alpha_s) + p_{t1} \delta + \delta \theta_t (\alpha_t - p_{t2})}{\delta^2 - \alpha_s \alpha_t (\theta_s^2 - 1) (\theta_t^2 - 1)}$$

$$q_{s3} = \frac{\alpha_s (p_{s1} \alpha_t \theta_s (1 - \theta_t^2) + \delta \theta_s (p_{t1} - p_{t2} \theta_t + \alpha_t (\theta_t - 1)) + \alpha_s \alpha_t (\theta_s - 1) (\theta_t^2 - 1) - \delta^2) + p_{s3} (\alpha_s \alpha_t (\theta_t^2 - 1) + \delta^2)}{\alpha_s (\alpha_s \alpha_t (\theta_s^2 - 1) (\theta_t^2 - 1) - \delta^2)}$$

$$q_{t1} = \frac{\delta (p_{s1} - p_{s3} \theta_s) + p_{t1} (\alpha_s - \alpha_s \theta_s^2) + \alpha_s (\theta_s^2 - 1) \theta_t (p_{t2} - \alpha_t) + \alpha_t (\theta_s - 1) (\alpha_t \theta_s + \alpha_t + \delta)}{\delta^2 - \alpha_s \alpha_t (\theta_s^2 - 1) (\theta_t^2 - 1)}$$

$$q_{t2} = \frac{\alpha_t \theta_t (\delta (p_{s1} - p_{s3} \theta_s) + p_{t1} (\alpha_s - \alpha_s \theta_s^2) + \alpha_s (\theta_s - 1) (\alpha_t \theta_s + \alpha_t + \delta)) + (p_{t2} - \alpha_t) (\alpha_s \alpha_t (\theta_s^2 - 1) - \delta^2)}{\alpha_t (\alpha_s \alpha_t (\theta_s^2 - 1) (\theta_t^2 - 1) - \delta^2)}$$

Foreseeing consumer responses, all firms simultaneously determine their posted prices to maximize their respective profits which are given by

$$\Pi_{s,1}(p_{s1}, p_{t1}, \delta) = p_{s1} q_{s1} + (p_{t1} - \delta q_{s1}) q_{t1}, \quad \Pi_{t2}(p_{t2}) = p_{t2} q_{t2}, \quad \Pi_{s3}(p_{s3}) = p_{s3} q_{s3}$$

The model is well defined for $\delta < \sqrt{\alpha_s \alpha_t (1 - \theta_s^2)(1 - \theta_t^2)}$, but the nature of the resulting equilibrium now depends on four parameters in a nonlinear fashion. For the special case in which $\alpha_s = \alpha_t = 1$ and $\theta_s = \theta_t = \theta$, Table TA5 displays the posted prices, quantities demanded and profits for the focal and competitor firms²².

Quantity	Focal Firm	Competitor
Price	$\frac{(\theta^2 - 1)(\delta^2(\theta - 2) + (\theta - 1)^2(1 + \theta)(2 + \theta))}{2(2 - \delta)(\delta - 1)(1 + \delta) - 3(\delta^2 + \delta - 3)\theta^2 + (\delta - 6)\theta^4 + \theta^6}$	$\frac{(\theta^2 - 1)(\delta - 1 + \theta^2)(\delta - 2 + \theta + \theta^2)}{2(2 - \delta)(\delta - 1)(1 + \delta) + 3(\delta^2 - \delta + 3)\theta^2 + (\delta - 6)\theta^4 - \theta^6}$
Demand	$\frac{\delta^2(\theta - 2) + (\theta - 1)^2(1 + \theta)(2 + \theta)}{2(2 - \delta)(\delta - 1)(1 + \delta) + 3(\delta^2 + \delta - 3)\theta^2(6 - \delta)\theta^4 - \theta^6}$	$\frac{(\delta^2 - 1 + \theta^2)(\delta - 2 + \theta + \theta^2)}{2(2 - \delta)(\delta - 1)(1 + \delta) + 3(\delta^2 + \delta - 3)\theta^2(6 - \delta)\theta^4 - \theta^6}$
Profits	$\frac{(2 - \delta - 2\theta^2)(\delta^2(\theta - 2) + (\theta - 1)^2(1 + \theta)(2 + \theta))^2}{(2(2 - \delta)(1 - \delta)(1 + \delta) + 3(3 - \delta - \delta^2)\theta^2 + (\delta - 6)\theta^4 + \theta^6)^2}$	$\frac{(\delta + 1 - \theta^2)(\delta - 1 + \theta^2)(\delta^2 - 1 + \theta^2)(\delta - 2 + \theta + \theta^2)^2}{(2(2 - \delta)(1 - \delta)(1 + \delta) + 3(3 - \delta - \delta^2)\theta^2 + (\delta - 6)\theta^4 + \theta^6)^2}$

Table TA5: Results for the duopoly-duopoly scenario.

Because of the more intense competition, posted prices and quantities demanded for the focal firm are not monotonically increasing in the rate of the cross-market discount. For moderate values of δ , prices and quantities demanded increase. However, when facing deeper discounts, both competitors react stronger, hurting the profitability of the strategy. As a consequence we observe that prices, quantities and therefore profits exhibit an inverted-U shape as a function of the cross-market discount.

We then proceed to numerically compute the optimal discount. As in the monopoly-duopoly scenario, the optimal cross-market discount is increasing in the importance of both markets (α_s and α_t), but decreasing in the degree of competition in the target market (θ_t). Increasing the degree of

²² For this special case, the equilibrium in the two markets is identical. This is a direct consequence of the simplifying assumption that markets have the same importance and intensity of competition.

competition in the source market (θ_s) leads to a similar decreasing trend. Figure TA8 displays contour plots summarizing these relationships.

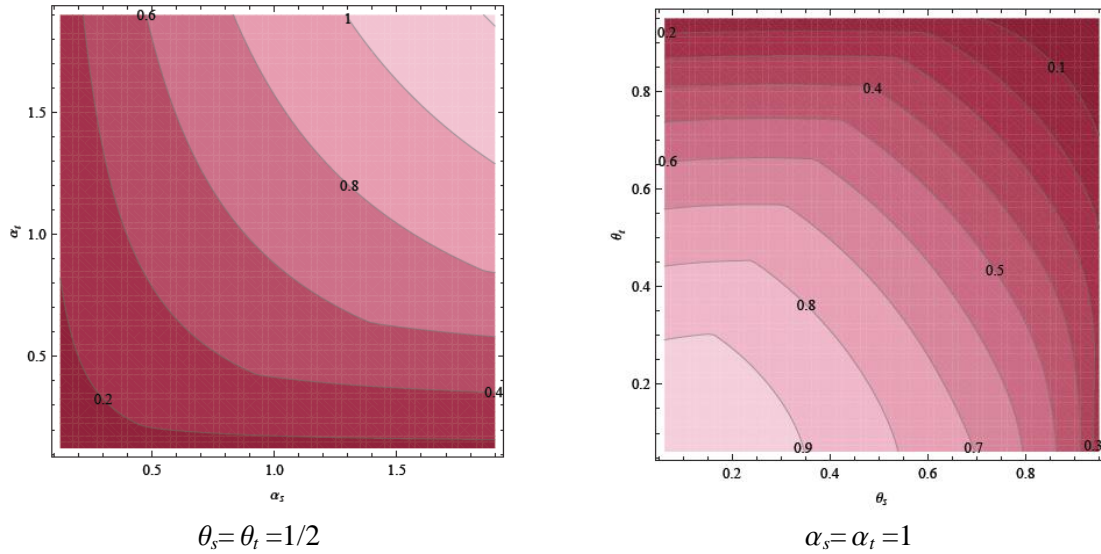


Figure TA8: Contour plots for optimal discount as a function of market importance

Extending previous results, we find that competition in any market prevents firms from fully extracting consumer surplus. A situation in which consumers have no surplus only occurs when competition parameters in both markets are close to zero, resembling the monopoly-monopoly scenario. Also, we verify that competitor firms in both markets make positive profits in the interior of the parameter space. The profit of Firm 2, which competes in the target market, approaches zero when its market has no attractiveness to the consumer ($\alpha_t \rightarrow 0$), or asymptotically when the importance of the source market dominates the consumer's decision ($\alpha_s \rightarrow \infty$). Similarly, the profit of Firm 3, which competes in the source market, asymptotically goes to zero when the importance of the target market dominates consumer's decision ($\alpha_t \rightarrow \infty$). Regarding the degree of competition, profits approach zero when the corresponding parameter approaches its maximum (θ_s or $\theta_t \rightarrow \infty$).

Figure TA9 displays total change in profits for all firms with respect to the case in which no cross-market discount is used, as a function of all four parameters of the model. Not surprisingly, the focal firm is always better off by offering a cross-market discount. At the same time, competitors in both markets are worse off when facing such a strategy.

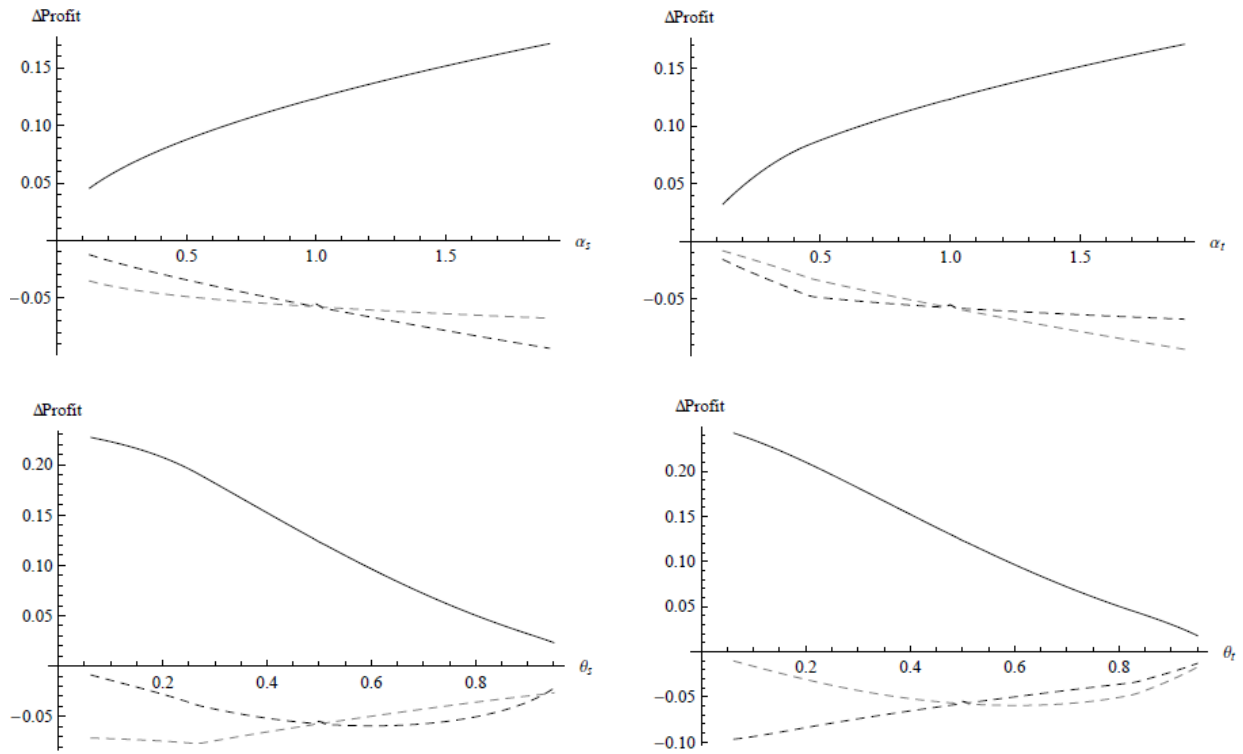


Figure TA9: Variation of profits with respect to the case in which no cross-market discount is used. In each panel, full black lines represent profit of firm 1 (focal firm), black-dashed lines represent profit of firm 3 (competitor in source market) and gray-dashed lines represent profit of firm 2 (competitor in target market).

Not obvious are the results derived from observing how the differential profit gains are affected by variations in the parameters of the model. When looking at importance parameters, we see that the focal firm benefits from an increase in the importance parameter in *either* market, and this also leads to a greater reduction in profits of the competitors in *both* markets. Variations derived from different degrees of competition also demonstrate the competitive advantage of the cross-market strategy, but the effect of competition is not symmetric between markets as we show with the impact of

parameter θ_s . The largest loss that a cross-market discount could cause to the competitor in the source market is achieved for medium degrees of competition. When θ_s approaches 0, there is almost no interaction between the firms and what the focal firm does barely affects its competitors. When θ_s approaches 1, the competition is so strong and prices are so low to start with that neither firm can make any profits. Interestingly, the degree of competition in the source market also affects profits of the competitor in the target market. The smaller θ_s is, the larger is the reduction in profit for Firm 2, because the focal firm can leverage its cross-market discount strategy without worrying about Firm 3's competitive reactions. The impact of the parameter θ_t is analogous.

3.2. With competing cross-market discount program

In this section, we present a detailed analyzed of the case in which the firms compete in both source s and target t markets. We use the indices 1 and 2 to denote the two firms. Both firms have the possibility to implement cross-market discounts and, therefore, the game is symmetric. We keep all notation used in previous scenarios except for the need introduce an index on the cross-market discount offered by each firm, which we denote by δ_1 and δ_2 . Then, in the symmetric duopoly-duopoly scenario the consumer utility and expenditure are given by

$$\mathcal{U}(q_{s1}, q_{t1}) = \alpha_s \left(q_{s1} - \frac{q_{s1}^2}{2} + q_{s2} - \frac{q_{s2}^2}{2} - \theta_s q_{s1} q_{s2} \right) + \alpha_t \left(q_{t1} - \frac{q_{t1}^2}{2} + q_{t2} - \frac{q_{t2}^2}{2} - \theta_t q_{t1} q_{t2} \right)$$

$$\mathcal{E}(q_{s1}, q_{s2}, q_{t1}, q_{t2} | p_{s1}, p_{s2}, p_{t1}, p_{t2}, \delta) = p_{s1} q_{s1} + p_{s2} q_{s2} + (p_{t1} - \delta_1 q_{s1}) q_{t1} + (p_{t2} - \delta_2 q_{s2}) q_{t2}$$

Quantities demanded as a function of posted prices and cross-market discounts are obtained by solving the consumer surplus maximization problem. Demand functions for the special case in which $\alpha_s = \alpha_t = 1$ and $\theta_s = \theta_t = \theta$ are given by

$$\begin{aligned}
q_{s1} &= \frac{(p_{s1}-1+(p_{r1}-1)\delta_1)(\delta_2^2-1)+(p_{s2}-1-\delta_1+p_{r2}\delta_1+(p_{r2}-1+(p_{s2}-1)\delta_1)\delta_2)\theta+(p_{s1}-1+\delta_2-p_{r1}\delta_2)\theta^2-(p_{s2}-1)\theta^3}{(\delta_1^2-1)(\delta_2^2-1)-2(1+\delta_1\delta_2)\theta^2+\theta^4} \\
q_{s2} &= \frac{(\delta_1^2-1)(p_{s2}-1+(p_{r2}-1)\delta_2)+(p_{s1}-1-\delta_1+p_{r1}\delta_1+(p_{s2}-1+(p_{r2}-1)\delta_1)\delta_2)\theta+(p_{s2}-1+\delta_1-p_{r2}\delta_1)\theta^2-(p_{s1}-1)\theta^3}{(\delta_1^2-1)(\delta_2^2-1)-2(1+\delta_1\delta_2)\theta^2+\theta^4} \\
q_{r1} &= \frac{(p_{r1}-1+(p_{s1}-1)\delta_1)(\delta_2^2-1)+(p_{r2}-1-\delta_1+p_{s2}\delta_1+(p_{s2}-1+(p_{r2}-1)\delta_1)\delta_2)\theta+(p_{r1}-1+\delta_2-p_{s1}\delta_2)\theta^2-(p_{r2}-1)\theta^3}{(\delta_1^2-1)(\delta_2^2-1)-2(1+\delta_1\delta_2)\theta^2+\theta^4} \\
q_{r2} &= \frac{(\delta_1^2-1)(p_{r2}-1+(p_{s2}-1)\delta_2)+(p_{r1}-1-\delta_1+p_{s1}\delta_1+(p_{s1}-1+(p_{r1}-1)\delta_1)\delta_2)\theta+(p_{s2}-1+\delta_1-p_{s2}\delta_1)\theta^2-(p_{r1}-1)\theta^3}{(\delta_1^2-1)(\delta_2^2-1)-2(1+\delta_1\delta_2)\theta^2+\theta^4}
\end{aligned}$$

Foreseeing the above consumer responses, all firms simultaneously determine their posted prices to maximize their respective profits which are given by

$$\Pi_{st,1}(p_{s1}, p_{r1}, \delta_1) = p_{s1}q_{s1} + (p_{r1} - \delta_1 q_{s1})q_{r1}, \quad \Pi_{st,2}(p_{s2}, p_{r2}, \delta_2) = p_{s2}q_{s2} + (p_{r2} - \delta_2 q_{s2})q_{r2}$$

After solving for the optimal posted prices of both firms in both markets, we can derive optimal profits as a function of cross-market discounts. The equilibrium levels of discounts are found by intersecting reaction functions derived from taking first order conditions for each firm:

$$\frac{\partial \Pi_{st,1}(\delta_1, \delta_2)}{\partial \delta_1} = 0 \quad \text{and} \quad \frac{\partial \Pi_{st,1}(\delta_1, \delta_2)}{\partial \delta_2} = 0$$

As in other scenarios, we need to verify nonnegative consumer surplus, nonnegative profits for each firm and positive effective price that the consumer is paying in the target market (this is the price after the cross-market discount is applied). The problem is analytically intractable and therefore we proceed numerically to compute its solutions. For simplicity, we compute the equilibria for the case $\alpha_s = \alpha_r = \alpha$ and $\theta_s = \theta_r = \theta$ for several values of θ . We are interested in analyzing how the situation of the firms and the consumers changes in comparison to the case in which no cross-market discounts are implemented. In the left panel of Figure TA10, we plot the profit of the firms and the consumer surplus w.r.t. θ with and without cross-market discounts for the case of $\alpha=1$. The right panel displays the equilibrium level of the discount δ , the difference in profit for the retailers and the difference in the consumer surplus.

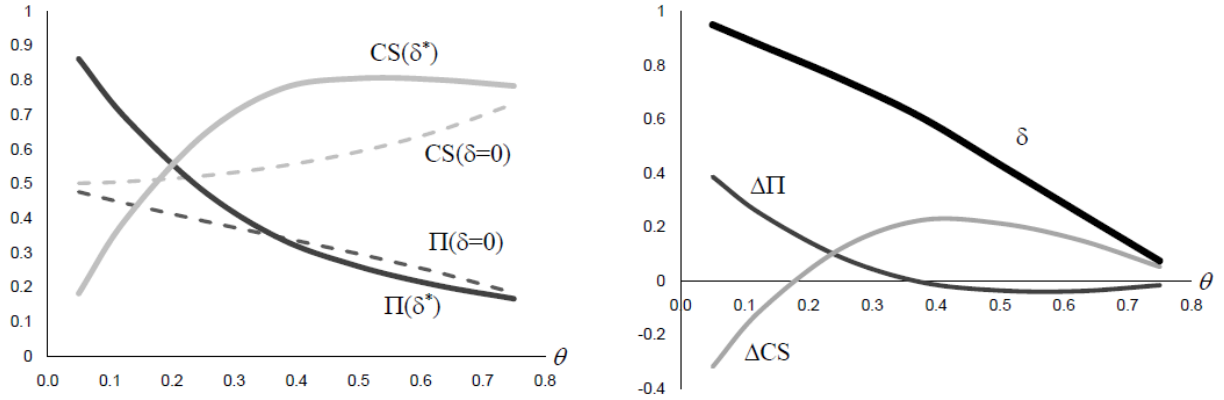


Figure TA10: Change in profits and surplus w.r.t. θ compared to the case in which no cross-market discount is used.

We observe that the value of the cross-market discount rate (δ) decreases with the competitive intensity (θ). As is also expected, when the intensity of competition increases, consumer surplus increases while the firms' profits decrease, but at different rates depending on whether cross-market discounts are used or not.

For small values of θ , demand functions are almost independent and the model resembles the monopoly-monopoly scenario in which the retailers are better off when offering cross-market discounts because they can extract a larger surplus from the consumers. The introduction of cross-market discounts increases the rate at which firms' profits decrease with the competition parameter. As a consequence, in the other extreme, for relatively large values of θ , the introduction of cross-market discount leads to more intense competition, decreasing each firm's ability to extract surplus from the consumer, thus making the firms worse off with respect the case in which no cross-market discounts are allowed. A very interesting result is obtained when the intensity of competition is

moderate. Here, all players, i.e., the consumer and both the firms, are better off with the introduction of cross-market discounts.

To confirm the generality of these results, we reparametrized $\alpha=a/(1-a)$, where $a \in [0,1)$. Therefore, by varying each of θ and α on a grid of values between 0 and 1, we can conduct a complete numerical analysis. In all of these cases, we found exactly the same patterns as above.

4. Analysis of cross-market discounts based on expenditure in the source market

In this section, we discuss the case in which discounts are based on the expenditure in the source market and not only on its quantity demanded. Our analysis here shows that, compared to this extension, the basic model is significantly more parsimonious and tractable, yet still captures most of the relevant insights of the discount scheme. In this extension, we base our analysis on the monopoly-duopoly scenario. The consumption utility remains unchanged, but the expenditure function changes. Let δ_r be the rate of cross-market discount. Then, the new expenditure function is given by $\mathcal{E}(q_{s1}, q_{t1}, q_{t2} | p_{s1}, p_{t1}, p_{t2}, \delta_r) = p_{s1}q_{s1} + (p_{t1} - \delta_r p_{s1}q_{s1})q_{t1} + p_{t2}q_{t2}$. In Stage 1, the demand functions are derived from consumer surplus maximization and are given by

$$q_{s1} = \frac{\alpha_s \alpha_t (-1 + \theta_t^2) + p_{s1} (\alpha_t (-1 + \delta_r - \theta_t) (-1 + \theta_t) + \delta_r (p_{t1} - p_{t2} \theta_t))}{p_{s1}^2 \delta_r^2 + \alpha_s \alpha_t (-1 + \theta_t^2)}$$

$$q_{t1} = \frac{p_{t1} \alpha_s + p_{s1}^2 \delta_r - \alpha_s (\alpha_t + p_{s1} \delta_r) + \alpha_s (-p_{t2} + \alpha_t) \theta_t}{p_{s1}^2 \delta_r^2 + \alpha_s \alpha_t (-1 + \theta_t^2)}$$

$$q_{t2} = \frac{(p_{t2} - \alpha_t) (\alpha_s \alpha_t - p_{s1}^2 \delta_r^2) + \alpha_t (-p_{t1} \alpha_s + \alpha_s \alpha_t + p_{s1} (-p_{s1} + \alpha_s) \delta_r) \theta_t}{\alpha_t (p_{s1}^2 \delta_r^2 + \alpha_s \alpha_t (-1 + \theta_t^2))}$$

In Stage 2, the analysis to find equilibrium prices is algebraically intractable and, therefore, we need to rely on numerical methods. We base our analysis on results computed on a grid of values of α_s, α_t

$\in [0,2]$ and $\theta_t \in [0,1)$. To illustrate the nature of the resulting equilibrium, in Figure TA11 we display equilibrium prices, quantities demanded and profits as a function of the discount rate δ_r for the cases $\alpha_s=\alpha_t=1$ and $\theta_t=1/2$.

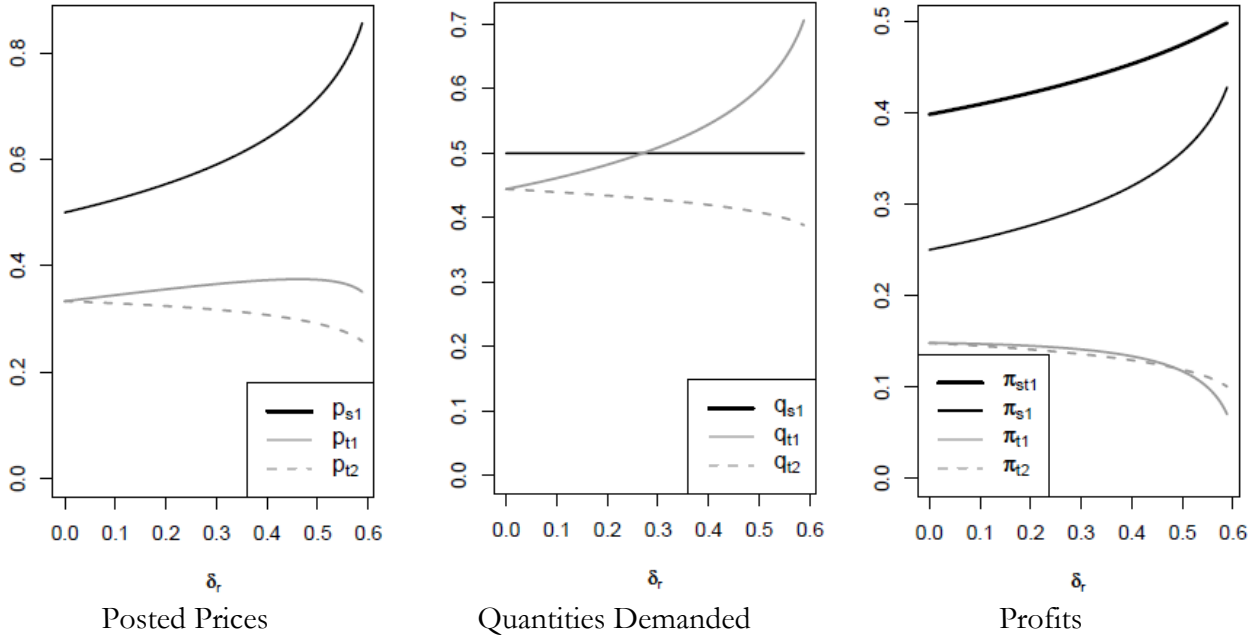


Figure TA11: Prices, quantities and profits as functions of the discount rate δ_r .

First, we note that the optimal posted price in the source market increases in δ_r . The direction for the posted price in the target market depends on the competitive structure: p_{t1} decreases in δ_r if θ_t is small, but in the presence of more intense competition it exhibits an inverted-U shape. Quantities sold in the target market are increasing in δ_r , but constant in the source market. Profits follow exactly the same patterns as in the quantity-based discount scenario. Total profit for firm 1 is increasing if θ_t is small, but as θ_t increases, it exhibits an inverted-U shape if the target market is important enough with respect to the source market. Similarly, as in the quantity-based discount case, consumer surplus is decreasing in δ_r if θ_t is small, but increases if θ_t is large.

Inspecting the optimal rate of discount we verify that, as before, in the monopoly-duopoly scenario the competitor makes positive profits and the consumer has positive surplus. Finally, we analyze how the optimal value of the cross-market discount rate (δ_r) varies with each of the model parameters. Numerical results show that it increases with the importance of target market in consumption utility (α_t), but decreases with the importance of source-market in consumption utility (α_s) and with the competitive intensity in the target market (θ_t)²³. Although the directional effect of α_s on the *rate* of the discount is different from the one we observe in the basic model, its effect on the *monetary size* of the discount and effective price paid in the target market remains unchanged.

Therefore, we confirm that the key insights we derived from the basic model generalize to the case in which the discount is a function of the expenditure in the source market. This provides support to the idea that the properties that we described before using a quantity-based discount model are not an artifact of the specific implementation of the cross-market discount policy.

An interesting feature of this extension is that a discount from the monopoly to the duopoly market is not symmetric with respect to a discount from the duopoly to the monopoly market. The expenditure function when cross discounting from the duopoly to the monopoly market is given by

$\mathcal{E}(q_{s1}, q_{t1}, q_{t2} | p_{s1}, p_{t1}, p_{t2}, \delta_r) = (p_{t1} - \delta_r p_{t1} q_{t1}) q_{s1} + p_{t1} q_{t1} + p_{t2} q_{t2}$. The resulting demand functions are given by

²³ For these numerical solutions, we confirm that we indeed have maximums and all second-order conditions are satisfied.

$$q_{s1} = \frac{p_{r1}\delta_r(p_{r1} + \alpha_t(-1 + \theta_t) - p_{r2}\theta_t) + \alpha_s\alpha_t(-1 + \theta_t^2) + p_{s1}(\alpha_t - \alpha_t\theta_t^2)}{p_{r1}^2\delta_r^2 + \alpha_s\alpha_t(-1 + \theta_t^2)}$$

$$q_{r1} = \frac{p_{r1}(\alpha_s + p_{s1}\delta_r - \alpha_s\delta_r) + \alpha_s(\alpha_t(-1 + \theta_t) - p_{r2}\theta_t)}{p_{r1}^2\delta_r^2 + \alpha_s\alpha_t(-1 + \theta_t^2)}$$

$$q_{r2} = \frac{(p_{r2} - \alpha_t)(\alpha_s\alpha_t - p_{s1}^2\delta_r^2) + \alpha_t(\alpha_s\alpha_t + p_{r1}(\alpha_s(-1 + \delta_r) - p_{s1}\delta_r))\theta_t}{\alpha_t(p_{r1}^2\delta_r^2 + \alpha_s\alpha_t(-1 + \theta_t^2))}$$

Using these demand systems, we can explore which direction of the discount is more profitable for the focal firm by computing optimal profits under both strategies: cross discounting from monopoly to duopoly (M2D) and from duopoly to monopoly (D2M). Figure TA12 displays optimal profits for marginal variations of each parameter of the model around $(\alpha_s, \alpha_t, \theta_t) = (1, 1, 1/2)$.

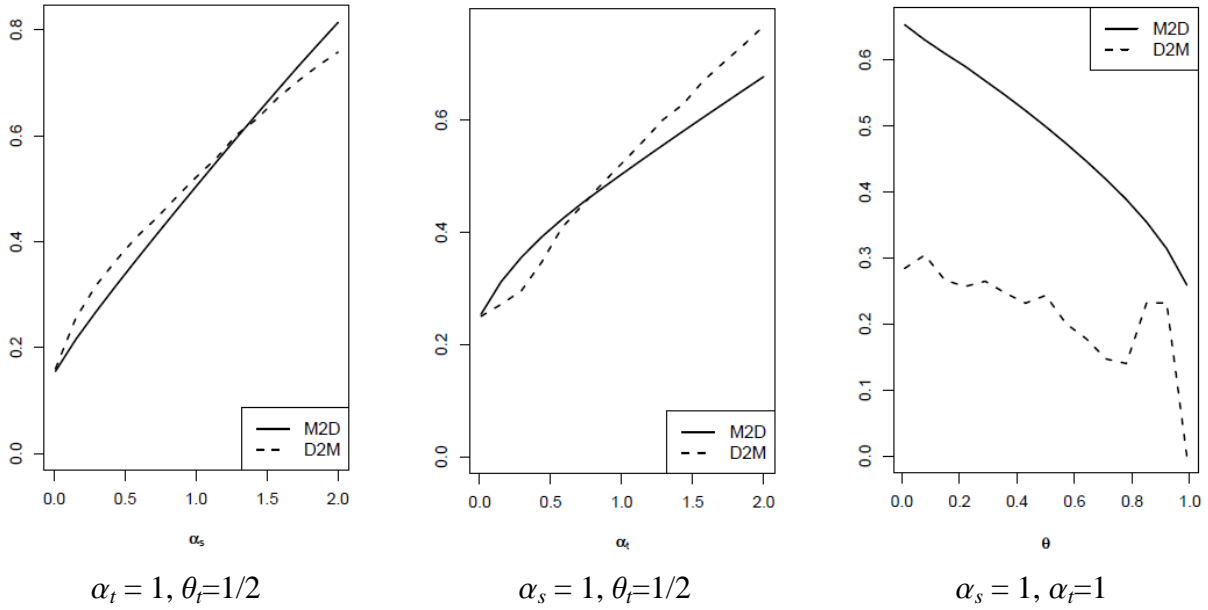


Figure TA12: Variation in profits for discounts from monopoly to duopoly (M2D) and from duopoly to monopoly (D2M) as functions of model parameters.

First, we verify that if one of the market has no importance for the consumer ($\alpha_s = 0$ or $\alpha_t = 0$) the cross-market discount strategy has no value and there is no difference in using either direction. Additionally, we note that profits for both strategies are increasing in market importance in any market. In other words, the focal firm could take advantage of a more valuable market regardless of

whether it is cross discounting from the monopoly to the duopoly market, or from the duopoly to the monopoly market. Upon comparing profitability, we conclude that a cross-market discount should be offered from the more valuable to the less valuable market. For example, the leftmost panel of Figure TA12 displays profits as a function of importance of monopolistic market s (α_s), keeping constant the importance of duopoly market t ($\alpha_t=1$) and competition intensity ($\theta_t =1/2$). We observe that when α_s is small relative to α_t , profits of cross discounting from s to t are higher than discounting from t to s . This relation reverses when α_s gets larger relative to α_t . The center panel displays the marginal variations with respect to α_t and confirms this pattern. Finally, the rightmost panel illustrates how profits of the cross-market discount strategy in each direction vary with the degree of competition (θ_t) in the duopoly market t . We conclude that when importance of both markets is comparable, the focal firm should prefer to offer discounts that are redeemable in the market that faces competition.

Technical Appendix 2:

Inferring Competitor Pricing with Incomplete Information

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 (1.6) to hold and also $q_{imt} = \tau_{imt}$, then

$$\left. \begin{aligned} q_{imt} &= \frac{1}{2} \left(\tau_{imt} + \frac{p_{imt}^c - p_{imt}}{\gamma + \eta} \right) \\ q_{imt} &= \tau_{imt} \end{aligned} \right\} \Rightarrow \gamma + \eta = \frac{1}{\tau_{imt}} (p_{imt}^c - p_{imt})$$

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

$$\left. \begin{aligned} q_{imt}^c &= \frac{1}{2} \left(\tau_{imt} + \frac{p_{imt} - p_{imt}^c}{\gamma + \eta} \right) \\ q_{imt}^c &= \tau_{imt} \end{aligned} \right\} \Rightarrow \gamma + \eta = \frac{1}{\tau_{imt}} (p_{imt} - p_{imt}^c)$$

2. Sampler

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

1. Start with initial values of θ_{im}

2. For each customer

a. Propose a new value θ_{im} according to a random walk process. And accept it with

$$\text{probability} \min \left\{ 1, \frac{p(q_{im} | \theta_{im}^{\text{new}}) p(\theta_{im}^{\text{new}})}{p(q_{im} | \theta_{im}) p(\theta_{im})} \right\}$$

b. draw a value of σ_i conditional on θ_i : $\sigma_{1i}^2 \sim \frac{\nu_1 \cdot ssq_1 + s_i}{\chi_{\nu_1+T}^2}$ and $\sigma_{2i}^2 \sim \frac{\nu_2 \cdot ssq_2 + s_i}{\chi_{\nu_2+T}^2}$

3. Update upper level regression parameters Λ and V_θ following conjugate multivariate regression model. The posterior of the conjugate multivariate linear regression

$$V_\theta | \theta_m, Z \sim IW(\nu + T, V + \tilde{S}) \quad (1.19)$$

$$\Lambda | \theta_m, Z, V_\theta \sim N(\tilde{\lambda}_m, V_\theta \otimes (Z'Z + A)^{-1}) \quad (1.20)$$

Where

$$\tilde{\lambda}_m = \text{vec}(\tilde{\Lambda}), \quad \tilde{\Lambda} = (Z'Z + A)^{-1} (Z'Z\hat{\Lambda} + A\bar{\Lambda}) \quad (1.21)$$

$$\tilde{S} = (\theta_m - Z\tilde{\Lambda})'(\theta_m - Z\tilde{\Lambda}) + (\tilde{\Lambda} - \bar{\Lambda})'A(\tilde{\Lambda} - \bar{\Lambda}) \quad (1.22)$$

4. Repeat as necessary.

In our empirical application we use the following hyper priors parameters: $\nu = \nu_1 = \nu_2 = k + 3$,

$ssq_1 = ssq_2 = 100$, $V = \nu \cdot I_k$, $A = 0.001 \cdot I_{n_z}$ and $\bar{\Lambda} = \mathbf{0}$, the $n_z \times k$ zero matrix.

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.

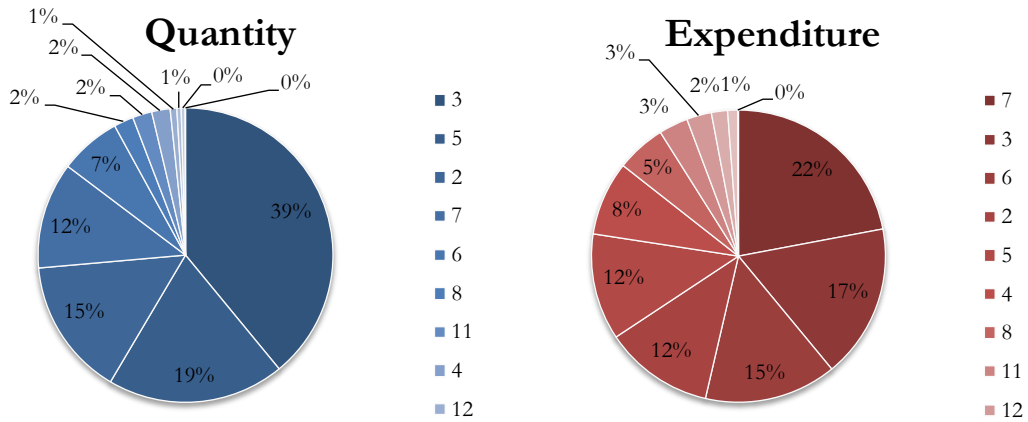


Figure TA2.1: Distribution of transaction per product category

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.

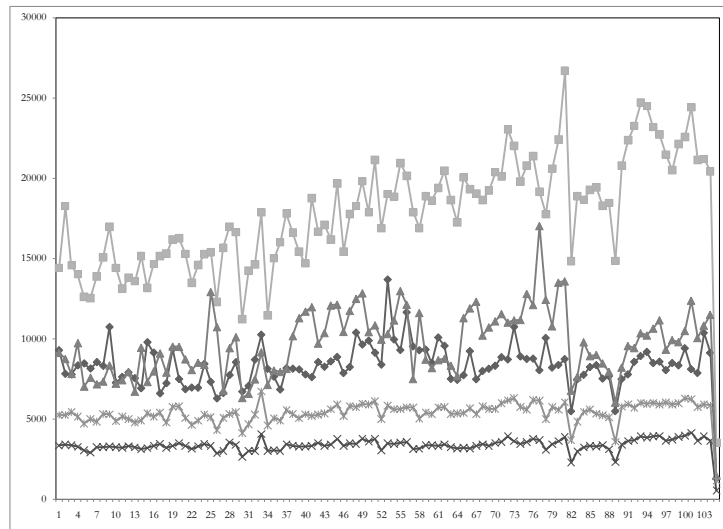


Figure TA2.2: Aggregated time series for the selected product categories.