Print Date: 4/20/2010

Making Better Pricing Decisions with Implied Priors

Alan L. Montgomery and Marcel Goic

Working Paper

April 2010

Alan L. Montgomery (e-mail: alan.montgomery@cmu.edu) is an Associate Professor at the Tepper School of Business, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, and Marcel Goic (email: mgoic@cmu.edu) is a doctoral candidate from the same institution and Instructor at Department of Industrial Engineering, University of Chile. The corresponding author is Alan L. Montgomery

Making Better Pricing Decisions with Implied Priors

Abstract:

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 this model is used for inference to make decisions about the optimal price. Often the optimal pricing solutions from the estimated demand model are non-sensical. Hence, the manager imposes a set of constraints on the feasible price space to find 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. This information should more appropriately be incorporated into the prior distribution of the parameter estimates. We show how constraints and statements about optimal prices imply informative prior distributions that can be used in a Bayesian approach. These prior distributions in turn influence the parameter estimates and subsequently the posterior distribution of optimal prices. In effect the Bayesian approach constrains the parameter space instead of the pricing decision space. The Bayesian method improves the quality of the pricing decisions made by managers and offers a consistent and scientific approach for incorporating managerial expertise into the pricing problem.

Keywords: Optimal Pricing, Constrained Optimization, Bayesian Methodology, Informative 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):

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

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:

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

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:

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

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:

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

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

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:

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

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

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

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.

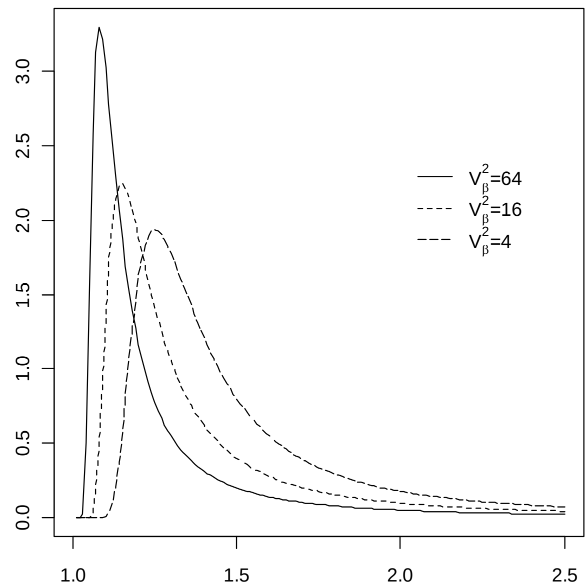

Figure 1. Distribution of optimal prices for several values of the variance of the price sensitivity parameter.

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 prior[1] 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

[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.

information in the previous example that $\overline{\beta} = -3$ 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:

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

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:

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

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 elasticity[2].

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 obtain[3]. Table 1 summarizes posterior means, standard errors, and 95% credible intervals for prices and β parameter for all scenarios as well as the percentage

[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).

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, seemly non-informative priors on the price elasticity combined with optimal price constraints are in truth quite informative.

	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).

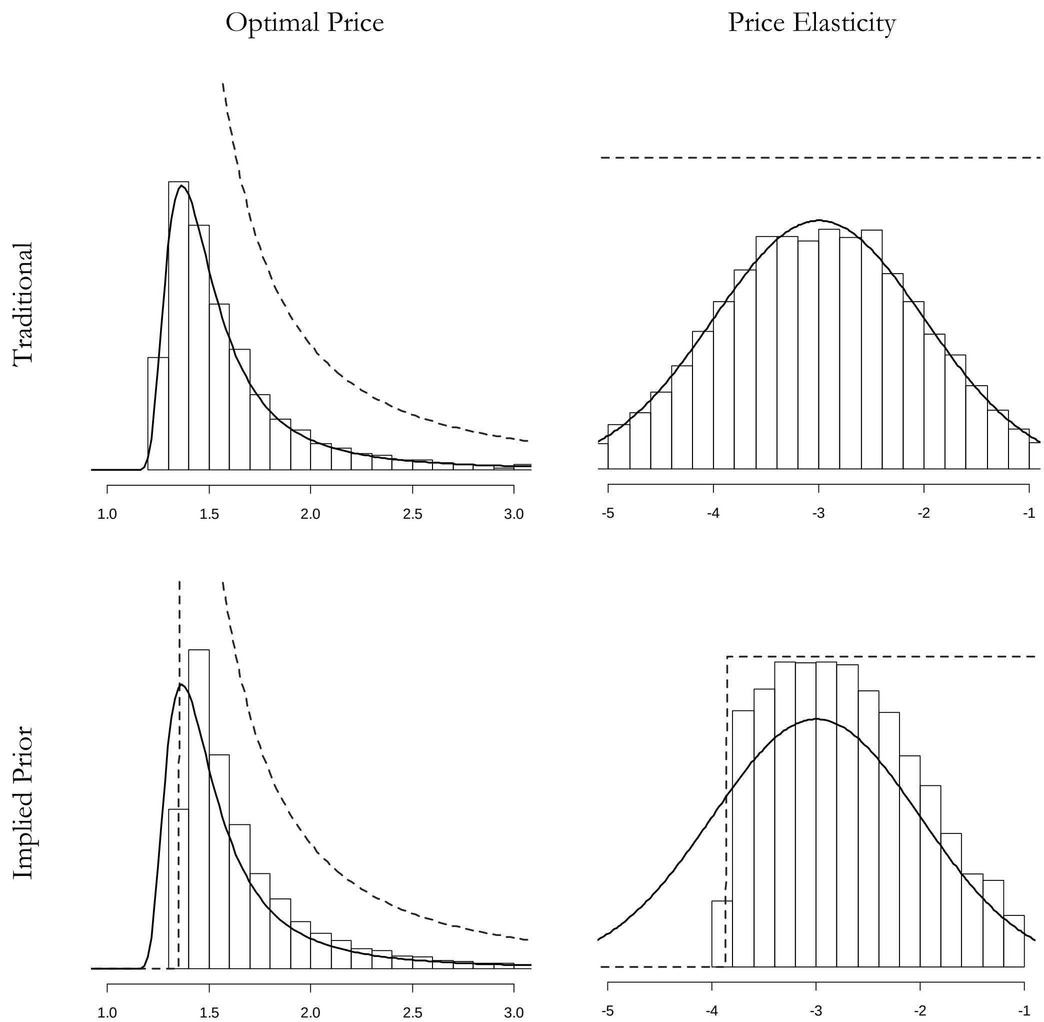

Figure 2. The histograms represent the posterior distribution, the solid lines represent the likelihood, and the dashed lines represent the prior distributions.

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 Essegaier 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: $q \leq q_u$. 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 prices[4].

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

[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 we the set of constraints we want to impose.

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 products[5]. 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 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,

---

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

Mulhern and Leone, 1991; and Kalyanam, 1996), we model the demand of each product i in each week t by a semilog model[6] ($i \in \{1,\ldots,N\}$ and $t \in \{1, ..., T\}$):

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

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:

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

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 ith element is p_i.

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

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

$$\boldsymbol{\theta}_i \mid \sigma_i^2 \sim \mathrm{N}\left(\mathbf{0},\ \sigma_i^2 V_i\right) \cdot \mathrm{I}\left[\left\{p_{0i}^* \le p_i^* \le p_{1i}^*\right\}\right] \cdot \mathrm{I}\left[\max\left(\mathbf{p}_A^*\right) \le \min\left(\mathbf{p}_B^*\right), \max\left(\mathbf{p}_B^*\right) \le \min\left(\mathbf{p}_C^*\right)\right] \tag{12}$$

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

---

[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.

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: $p_{0i}^* \leq p_i^* \leq p_{1i}^*$. 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:

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

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.

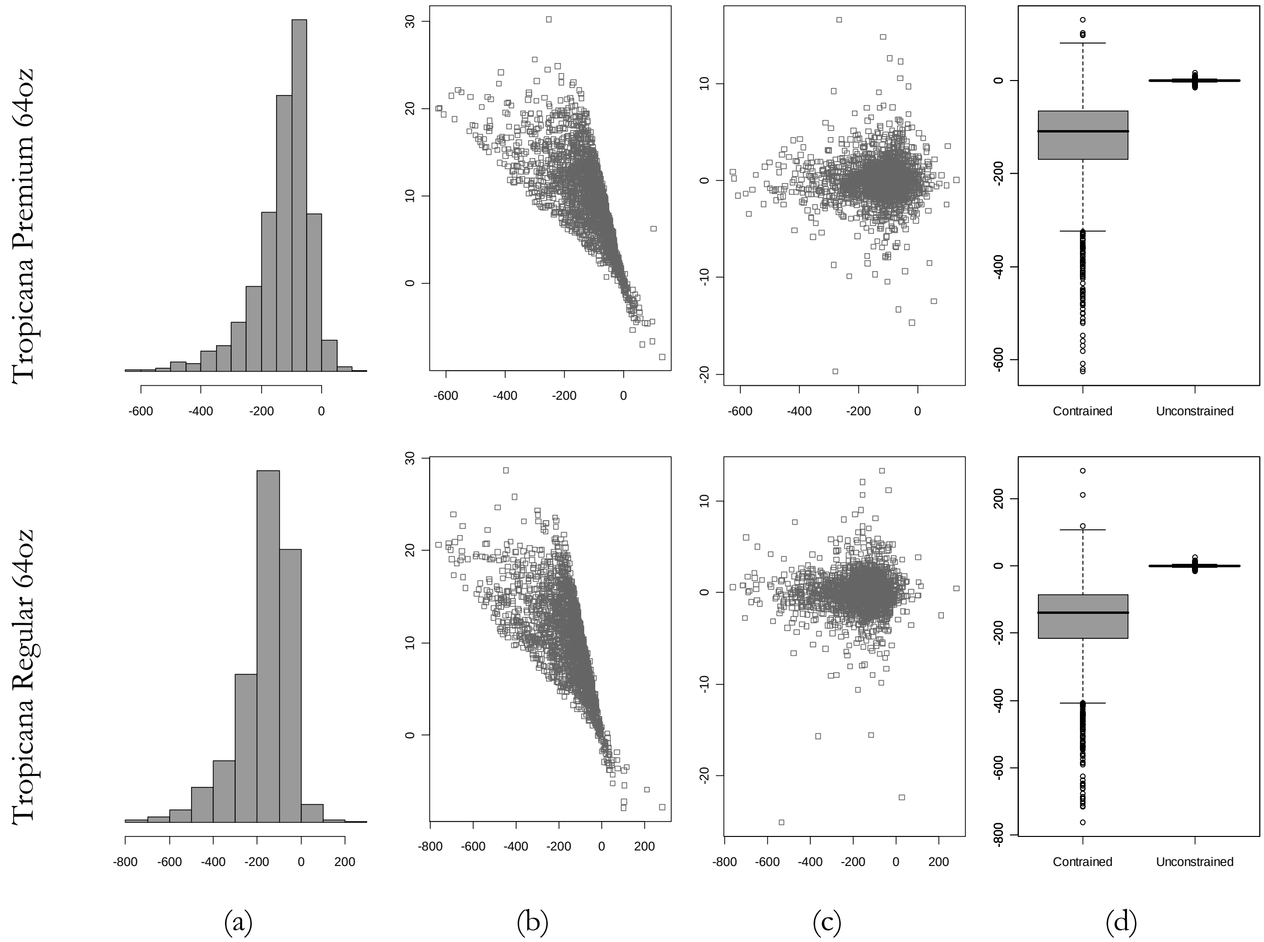

Figure 3. Prior distribution when considering price constraints (a) marginal distribution of own price coefficients, (b) joint distribution between own price coefficient and intercept, (c) joint distribution between own price coefficient and other brand price coefficient and (d) boxplots of own price 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 estimate[7] 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 Figure 4 we display boxplots of the parameters for a sub-set of the products. Here we observe that

[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.

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.

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

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

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

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

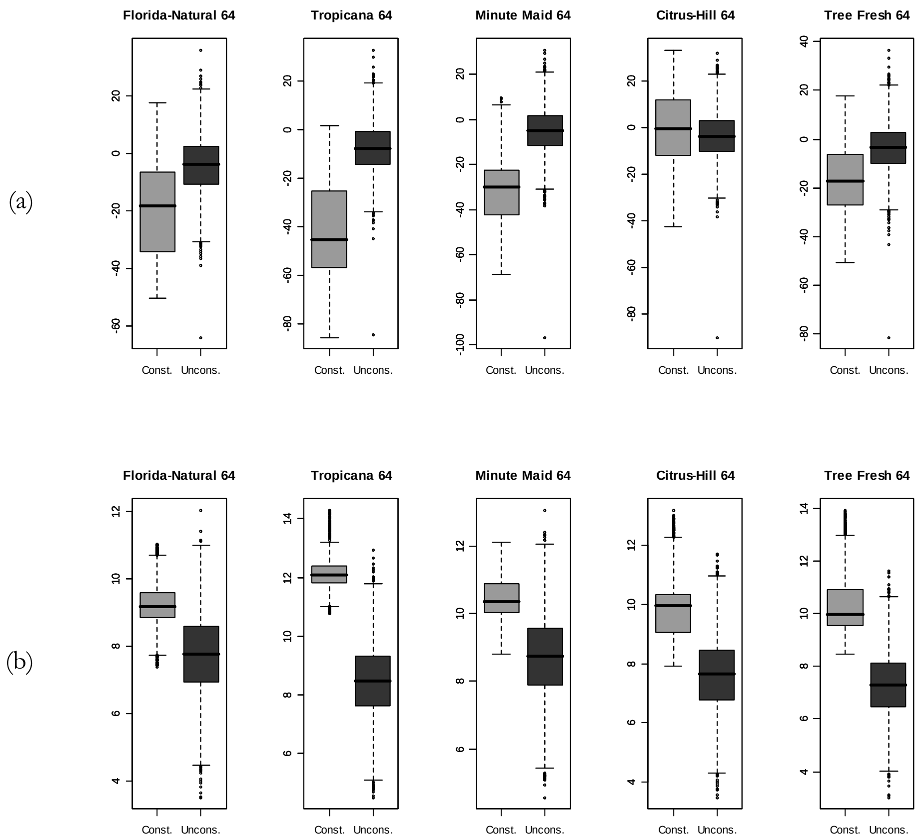

Figure 4. Boxplots of parameter estimates with and without considering price constraints on the definition of prior specification (a) Own price coefficient. (b) Intercept

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 Mean (Std.Err.)	TRADITIONAL 95% CI	IMPLIED PRIOR Mean (Std.Err.)	IMPLIED PRIOR 95% CI	BOUNDS (Lower, Upper)	% CHANGE
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.

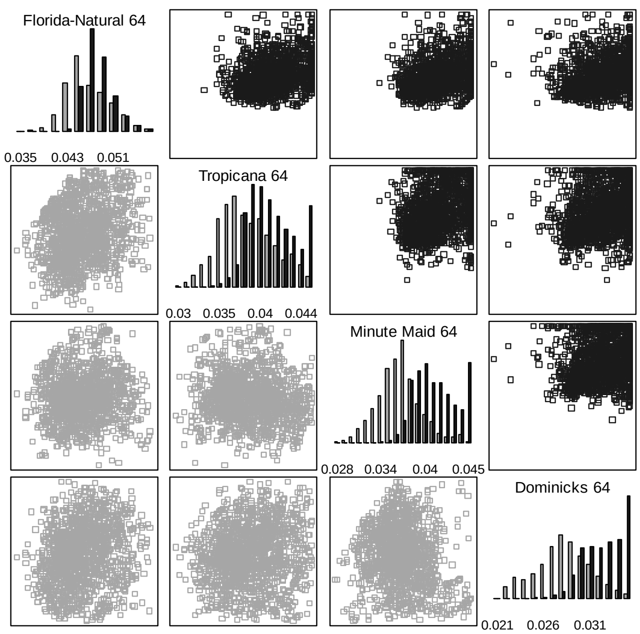

Figure 5. Comparison of optimal prices

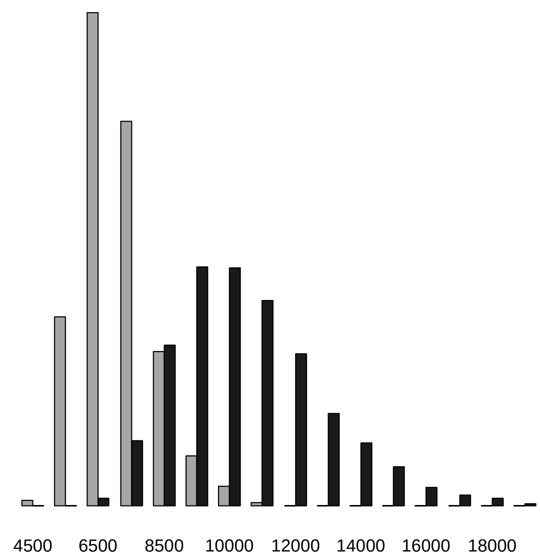

Figure 6. Comparison of optimal revenues.

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.

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.

Essegaier, 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 Uncertainity: 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.