Pricing Decisions Under Demand Uncertainty: A Bayesian Mixture Model Approach

Kirthi Kalyanam

Dept. of Marketing
Leavey School of Business
Santa Clara University
Santa Clara, CA-95053
Tel: 408-554-2705
Fax: 408-554-5056
Internet: kkalyanam@mailer.scu.edu

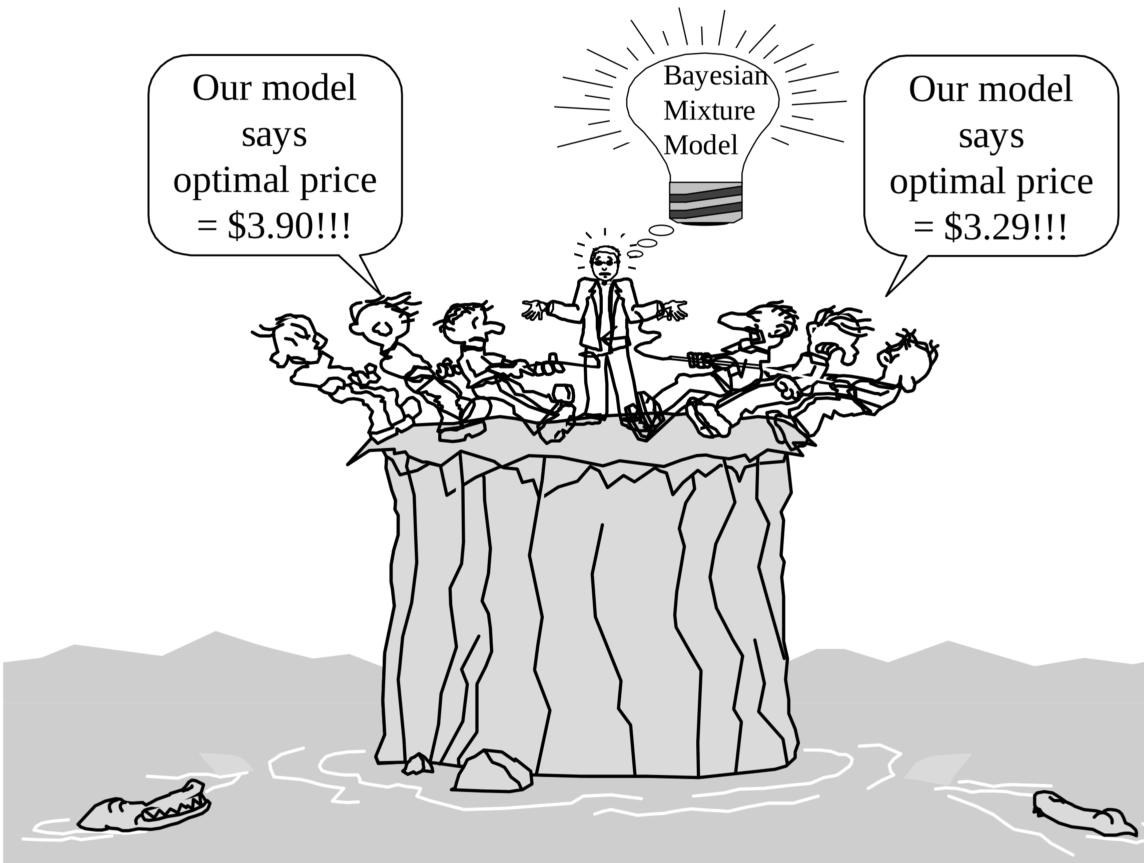
First Draft:	**June 1993**
Revised:	**October 1994, March 1995**
This Version:	**September 1995**

Keywords: Pricing , Bayesian Econometrics, Decision Theory, Regression, Non-nested Models, Demand Structures, Scanner Data.

One Line Abstract: How to make pricing decisions *in spite of* uncertainty about the demand function.

Acknowledgments: Comments by the Editor, the Area Editor and two anonymous reviewers greatly improved previous versions of this paper. This paper is based on my dissertation at Purdue University. I am indebted to Ward Hanson for motivating and guiding this work. The members of my dissertation committee Dan Putler, Jim Berger and Mark Moriarty provided valuable input. I have also benefited from the comments of Arnold Zellner, Luis Pericchi, Greg Allenby, Dipak Jain, Rabikar Chatterjee and Shelby McIntyre. All remaining errors are mine.

Suggested Cover Art



Abstract

In using an estimated demand function to evaluate the impact of pricing decisions a manager has to contend with the fact that *the form of the true demand function is unknown*. Often, many demand functions with different pricing prescriptions can be found to be consistent with the data. A Bayesian mixture model (BMM) is proposed to evaluate the impact of pricing decisions when there is parametric uncertainty about demand. The framework involves: (1) The decision maker proposing a set of demand specifications, priors on the specifications and parameters, and his/her utility structure; (2) Estimating and computing posterior probabilities for the specifications using Bayesian techniques; (3) Diagnostic testing of the specifications; and (4) Combining the estimated specifications and posterior probabilities with the utility structure to arrive at optimal pricing decisions. An empirical implementation using store level scanner data on ground coffee indicates that for the brands examined: (1) There is considerable demand uncertainty; (2) Classical tests are often inconclusive; and (3) The BMM approach allows the decision maker to formulate pricing strategy even when there is uncertainty about competitive interactions, cannibalization, and the profit maximizing price. Results from a Monte Carlo experiment suggest that the BMM approach performs well in terms of recovering potential improvements in profits.

1. Introduction

The advent of optical scanning devices and decreases in the cost of computing power have made it possible to process demand information in an accurate and timely fashion. Databases with sales and marketing mix information can be collected quite easily. These databases enable the estimation of response functions and pricing/promotion decisions in "real" time. For example, *The Wall Street Journal* reports[1] that Dominick's of Chicago recently conducted "real" time pricing experiments about the profitability of everyday low pricing versus high-low pricing formats. Commercial suppliers of marketing research like A.C.Nielsen are embedding estimated response functions in spreadsheets to create promotion planning and pricing tools (SCAN*PRO©, Wittink, et.al.1988) for brand managers. As price and private label competition intensifies[2], pricing and promotion planning tools that are based on estimated demand functions can play an increasing role as decision aids.

This explosion in the estimation and use of demand functions makes it timely and appropriate to re-examine several fundamental issues. In particular, demand functions are latent theoretical constructs whose exact *parametric form is typically unknown.* Estimates of price elasticities, profit maximizing prices, competitive interactions and other policy implications are *conditional* on the parametric form assumed in estimation. In practice, many forms may be found that are not only theoretically plausible but also consistent with the data. The different forms could suggest different profit maximizing prices leaving it unclear as to what is the appropriate pricing action.

Using store level scanner data, I demonstrate that there can be considerable uncertainty about the optimal price due to uncertainty about the parametric form of the demand function. This points to the importance of formalizing the impact of parametric demand uncertainty on pricing decisions. By formalizing the impact of demand uncertainty, and accounting for it, managers can make decisions that are more robust. I also find that existing classical tests may lack the power to resolve parametric demand

---

1 See Gibson, R., "Broad Grocery Price Cuts May Not Pay", *The Wall Street Journal*, Friday, May 7 1993, Page B1.

2 On this point see a recent series of articles in the *The Wall Street Journal* , dated April 21, April 26 and May 13 1993, under Marketing, Marketscan and Advertising.

uncertainty, particularly for non-nested comparisons. The structure of these tests does not permit seamless integration of estimation, specification analysis and optimal pricing into a unified framework. These shortcomings point to the need for other approaches that allow integration into a unified framework.

As an alternative to the existing approaches, I propose a Bayesian mixture model (BMM) approach that draws on Bayesian estimation, inference, and decision theory[3]. The BMM approach consists of input, estimation, diagnostic and optimal pricing modules. In the input module, alternate parametric models of demand are specified along with priors. Utility structures representing the decision maker's attitude towards risk can be explicitly specified. In the estimation module, the inputs are combined with data to compute parameter estimates and posterior probabilities for the models. The diagnostic module involves testing the statistical assumptions underlying the models. In the optimal pricing module the estimates and posterior probabilities are combined with the utility structure to arrive at optimal pricing decisions.

Formalizing demand uncertainty in this manner has many important payoffs. While the classical approaches emphasize choosing a demand specification, the BMM approach emphasizes constructing an objective function that represents a mixture of the specifications. Hence, pricing decisions can be arrived at *even* when there is no consensus among the different parametric specifications. The pricing decisions will reflect parametric demand uncertainty, and hence, be more robust than those based on a single demand model. Finally the BMM approach provides a unified theoretical framework for estimation, inference and optimal decision making. This is in contrast to previous research that has tended to focus on only one aspect, be it estimation, testing or optimal pricing. The empirical implementation confirms the benefits of using the BMM approach. In a Monte Carlo experiment, on average the BMM approach recovers 75% of the potential improvement in profits whereas pricing decisions based on a single specification average a 67% recovery rate.

[3] While the context and specifics of this application are new, Bayesian approaches have a fairly long history in marketing. A non-exhaustive list of some applications include model selection (Barry & Wildt 1977; Rust & Schmittlein 1985), hypothesis testing (Allenby 1991) and improved estimation of diffusion models (Lenk & Rao 1990; Lilien, Rao and Kalish 1981).

The rest of this paper is organized as follows: Section 2 is a review of related literature. Section 3 outlines the BMM approach. Section 4 describes an empirical implementation on a scanner data base. Sections 5 and 6 contain the estimation and diagnostic testing results. Section 7 explores the pricing implications. Section 8 presents the BMM optimal pricing solutions. Section 9 examines the ability of the BMM approach to recover potential improvements in profits. The paper concludes with a summary and directions for future research.

2. Related Literature

The marketing literature was surveyed with a focus on the twin issues of specification (the form of the demand relationship) and estimation (the parameters of the specification) uncertainty, and their impact on pricing decisions. The survey[4] revealed that past work has tended to utilize linear, double-log, semilog and attraction forms. Only about 5 studies explored multiple specifications and that too in the context of demonstrating statistical tests (e.g. Jain & Vilcassim 1989). None of these studies examined pricing implications of the different specifications.

In spite of the paucity of research on this issue, some formal and informal approaches to deal with parametric demand uncertainty do emerge from the literature. Informal approaches include comparison of goodness of fit measures like the R-squares of the different specifications. A commonly used approach is a face validity check by comparing the signs and magnitudes of the coefficients with *a priori* theoretical expectations. Restrictions based on theory do define the appropriate neighborhood in which to conduct the search. In practice, this neighborhood may be quite broad and need not imply a unique parametric specification. For example, theory tells us that demand should be downward sloping in price. However, many specifications may satisfy this requirement. Under these circumstances, the appropriate specification is an empirical question, and many researchers in marketing (Blattberg & Sen 1975; Jain & Vilcassim

[4] I started with 42 studies that were included in the Tellis (1988 Table 1, p.333) study. These studies covered the period 1961-1985. This list was augmented with an ABI/INFORM search and a review of all copies of JMR, Marketing Science and Management Science over 1986-1991. This yielded 8 more studies, bringing the total to 50 studies. After deleting studies that did not adopt a regression approach, I was left with 44 studies of sales response and/or market share. For a tabulation of these studies see Kalyanam (1993).

1993) and elsewhere (Theil 1971, p.540-43) have adopted the perspective that statistical testing procedures are in order.

Relevant statistical procedures include nested (likelihood ratio) or non-nested classical tests[5]. MacKinnon (1983) surveys this material and Balasubramanian and Jain (1994) demonstrate the application of these tests in a marketing context. However, the classical testing approach suffers from several drawbacks. Numerous tests tailored to specific situations exist, making the implementation cumbersome. Second the different tests may have different power characteristics, and one may not be able to make statements about the power of the battery of tests as a whole. Third, as Allenby (1991 p.379) points out, classical tests lack a natural metric for non-nested testing. Fourth, tests non-nested classical tests have low power (MacKinnon et.al. 1983, p.56; Balasubramanian & Jain 1994) and can often be *inconclusive*.

Finally, there is the issue of choosing a significance level for these tests and correctly accounting for uncertainty. The goal of the analyst is, given the data at hand and the uncertainty about the demand structure, to make the "best" pricing decisions. The decision maker's attitude towards specification and estimation uncertainty needs to be incorporated in a utility structure. This utility structure then needs to be combined with the uncertainty about the demand model to arrive at the appropriate pricing decisions. While the uncertainty due to estimation is embedded in the variance of the parameter estimates, the classical tests do not offer a measure of residual specification uncertainty. Rather, the structure of the classical tests is such that one is compelled to choose a specification at a significance level, while what we are really interested in is carrying forward the uncertainty due to estimation and specification to the utility structure.

The BMM approach proposed in this paper overcomes some of these drawbacks by adopting a Bayesian decision theoretic approach. A unified theoretical framework is proposed for estimation, inference, and decision making. Utility functions representing the decision maker's risk attitude can also be incorporated, thereby allowing for decision makers with varying risk preferences. The impact of parametric demand uncertainty on

---

[5] Examples of non-nested classical tests include the Cox Procedure, (1962); the P test, Davidson and MacKinnon (1981).

pricing decisions can be formally examined. Residual specification uncertainty can be quantified by computing posterior probabilities. Optimal pricing decisions can be arrived at *even* when demand uncertainty is unresolved. Rather than choose a specification at a significance level, all existing specifications can be carried forward to the decision stage with the posterior probabilities acting as weights. Such weighting of specifications has been extensively used in the combination of forecasts literature (Moriarty 1990) to improve forecasting accuracy. In contrast the objective of the BMM approach is to improve the effectiveness of pricing. Also the statistical foundations of the posterior probabilities are quite different from those for computing forecasting weights.

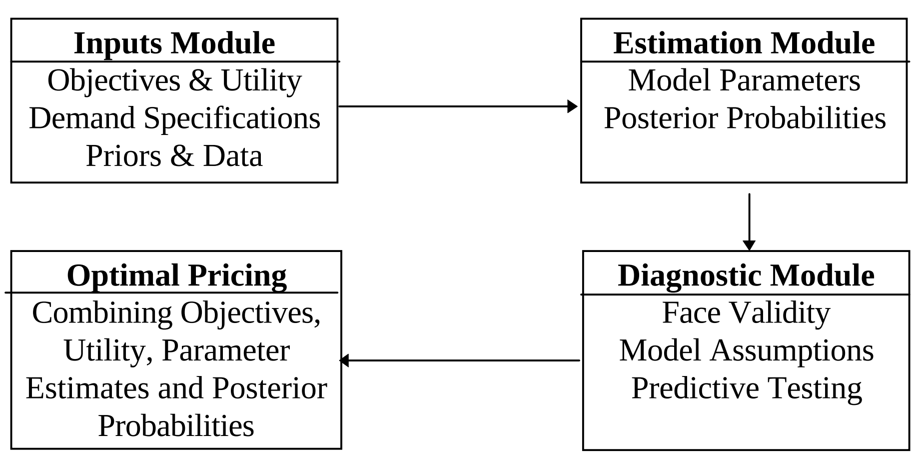
3. The Bayesian Mixture Model Approach

The BMM approach draws on Bayesian estimation, inference and decision theory. Figure 1 portrays the various modules. This section describes the modules.

3.1 Inputs Module: The inputs module consists of formalizing the decision maker's objectives and utility structure, identifying a set of demand specifications and forming priors for the specifications and the model parameters.

The first step is to formalize the decision maker's objectives and attitude towards risk into a utility function. Formally, let U(Ω,P) be the decision maker's utility (payoff) function, where Ω is the decision maker's objective and P is the action. The form of U(Ω,P) reflects the decision maker's attitude towards risk. In practice, one may have to elicit and estimate U(Ω,P) using algebraic procedures (Hauser and Urban 1979) or measurement error approaches (Eliashberg and Hauser 1985).

Figure 1: Overview of the BMM Approach



Next, the decision maker proposes parametric specifications that are viewed as reasonable representations of the demand relationship. This is an important step, since all decisions are conditional on the demand specifications proposed. It may be useful to compile a list of commonly used functional forms[6] and selecting from them keeping in mind what these functions imply regarding the sales response to price changes. Note that this *a priori* subjective specification of a set of demand relations would be necessary even under non-Bayesian approaches. Formally, we wish to model the relationship between K independent variables (x's) and a dependent variable y. We are interested in different specifications that may involve non-linear transformations on the y's and the x's as follows:

$$f_j(y_n) = \sum_{k=1}^{K} \theta_{kj} g_{kj}(x_{nk}) + \varepsilon_{nj}, \qquad (1)$$

where $f_j()$ is the transformation on the dependent variable in specification j, y_n is the nth observation on the dependent variable, $g_{kj}()$ is the transformation on the kth independent variable in the jth specification, θ_{kj} is the coefficient of the kth independent variable in the jth specification, x_{nk} is the nth observation on the kth independent variable, and ε_{nj} is a random disturbance for observation n in the jth specification.

Equation (1) is very flexible in that it permits any transformation of the dependent or the independent variables. Within a specification, one can choose a different transformation for the dependent variable and each of the independent variables. This flexibility (being able to specify a different transformation for each independent variable) is important for marketing research since there are binary variables like features, quarterly dummies, and coupon drops that are not amenable to certain transformations like the log or the inverse. To make the notation more concise, the coefficients and the data in specification j are henceforth referred to as:

$\Theta_j = [\theta_{1j},\ldots, \theta_{Kj}]^t$ is a Kx1 vector of the model parameters and ^{t} is the transpose operator,

Y_j = Nx1 vector of the dependent variables for the jth specification, with $f_j(y_n)$ as elements,

X_j = NxK matrix of independent variables for the jth specification, with $g_{kj}(x_{nk})$ as elements.

[6] Discussions of power, reciprocal, logarithmic, square root transformations and of non-linear specifications that are intrinsically linear can be found in Draper and Smith (1981), Chapter 5.

A prior probability, PR_j for each of the j specifications and a prior density $\Pi(\Theta_j, \sigma_j^2)$ is required. Raiffa and Schlaiffer (1961, p. 343) and Zellner (1971, Ch. 3) discuss various forms of non-diffuse priors in multiple regression contexts. However, the cognitive burden involved in specifying informative priors suggests that the ability to specify diffuse priors would have some practical value. Specifying diffuse priors for the parameters can be complicated due to the fact that the specifications can be non-nested and involve transformations on the dependent and independent variables. For example, one of the specifications can be linear, and the other can be semi-log involving a logarithmic transformation on the dependent variable. The diffuse priors specified need to be *invariant* across the specifications. This problem can be resolved by assigning diffuse priors in accordance with Jeffreys (1939, p.179) multiparameter rule. Priors assigned using this rule are invariant with respect to any differentiable one-to-one transformation and are given by:

$$\Pi(\Theta_j, \sigma_j^2) \propto \left| I_{\Theta_j, \sigma_j^2} \right|^{\frac{1}{2}}, \qquad (2)$$

where I_{Θ_j, σ_j^2} is the Fisher information. An alternative justification for the use of Jeffreys' multi-parameter rule from the perspective of the Box and Cox (1964) analysis of transformations is presented in section A of the Appendix.

3.2 Estimation Module: In the standard Bayesian manner the posterior density of the parameters is given by

$$\Pi(\Theta_j, \sigma_j^2 | Y_j, X_j) \propto L(Y_j | \Theta_j, \sigma_j^2, X_j)\Pi(\Theta_j, \sigma_j^2), \qquad (3)$$

where $L(Y_j | \Theta_j, \sigma_j^2, X_j)$ is the likelihood. Posterior probabilities for the specifications can be computed by integrating the right hand side of (3). Let

$$M(Y_j, X_j) = \iint_{\Theta_j, \sigma_j^2} L(Y_j | \Theta_j, \sigma_j^2, X_j)\Pi(\Theta_j, \sigma_j^2)\, d\Theta_j d\sigma_j^2. \qquad (4)$$

The scalar $M(Y_j,X_j)$ represents the evidence in support of specification j provided by the data. PT_j, the posterior probability of specification j, is obtained by combining $M(Y_j,X_j)$ with the prior probabilities PR_j in Bayes formula, as follows:

$$PT_j = \frac{PR_j \; M\left(Y_j, X_j\right)}{\sum_{j=1}^{J} PR_j \; M\left(Y_j, X_j\right)}. \qquad (5)$$

3.3 Diagnostic Module: The parameter estimates are compared with *a priori* theoretical expectations in order to establish their face validity. Specifications with questionable estimates can be eliminated at this stage. Further, the validity of the statistical assumptions are tested via analysis of the residuals. This includes examining the sensitivity of the model parameters and posterior probabilities to outlying observations, and predictive testing in holdout samples. If more than one specification survives the diagnostics, they are carried forward to the optimal pricing stage.

3.4 Optimal Pricing Module: The posterior probabilities for the specifications and the posterior densities for the parameters can now be combined with the utility structure to make optimal decisions in accordance with the *conditional Bayes principle.*[7] The optimal Bayes pricing action maximizes the Bayesian expected utility.[8] Formally, if $\Pi(\Omega_j)$ is the probability distribution of the payoff under specification j at the time of the decision, and P is the pricing action, then the optimal pricing action conditional on specification j maximizes $\int U\left(\Omega_j, P\right)\Pi\left(\Omega_j\right) d\Omega_j$. The unconditional (on specification) optimal pricing action is

$$\max_{P} \sum_{j=1}^{J} PT_j \left[\int U\left(\Omega_j, P\right)\Pi\left(\Omega_j\right) d\Omega_j \right]. \qquad (6)$$

4. Empirical Implementation

The goals of the implementation are: (1) Demonstrate an empirical application of the BMM framework; (2) Highlight the ramifications of demand uncertainty on pricing decisions; (3) Establish the inadequacies of the existing classical approaches; and (4) Compare and contrast the BMM results with the other approaches.

4.1 Data: The scanner data base on ground coffee sales was collected by SAMI from the Kansas city market for sixty five weeks (February 15 1979 to May 15 1980). The database

[7] See Berger (1985) p.16-19 for a discussion of various Bayesian and classical decision principles.
[8] provided the maximum exists.

provides store level data on the retail prices of the UPC's in the coffee category in each week and the corresponding number of units sold at each price level. The database also includes descriptions of advertisements that have appeared in a newspaper, flyer or bag stuffer. The top 6 brand size combinations[9] accounting for 87.1% of the category sales volume are retained for analysis. Table 1 contains summary statistics for the brands.

Table 1: Description of the Ground Coffee Brands

Brand	Market Share (Units)	Average Price ($/lb.)	Feature (% Weeks)
B1	20.97	3.12	20.00
B2	5.98	3.09	23.07
B3	37.88	3.13	16.92
B4	8.99	3.10	13.85
B5	20.06	3.12	18.46
B6	6.10	3.07	15.38

4.2 Objectives and Utility: The empirical implementation will be conducted from the perspective of a manufacturer who is interested in determining the profit maximizing wholesale prices for the ground coffee brand-sizes[10] subject to the following assumptions: (1) the retailers price the category by adding a markup to wholesale prices[11]; (2) all competitors have access to the same demand information and react to price changes. Modifying Equation (6) to reflect the manufacturers' pricing problem yields:

$$\max_{w_i} \sum_{s=1}^{S}\sum_{j=1}^{J} PT_{ijs}\left[\int U\left(\Omega_{ijs}, w_i\right)\Pi\left(\Omega_{ijs}\right)d\Omega_{ijs}\right], \qquad (7)$$

where i indexes brandsize, s indexes stores and w denotes wholesale price. There are several aspects of (7) that require elaboration. First, the summation over stores reflects the pricing problem faced by manufacturer i, that is to set a wholesale price w_i such that the expected total profit obtained by selling through the multiple stores in the data set is

[9] The six brand sizes were constructed following Guadagni & Little (1983). Since grinds of an UPC are priced and promoted together, treating each grind as a separate brand would result in perfect collinearity of own price and promotion and prices and promotion of other grinds of the same brand. Hence demand specifications cannot be estimated at the grind level and different grinds of a given brand-size are aggregated together.

[10] I would like to thank Richard Staelin for directing my attention to this formulation.

[11] For some empirical justification of this assumption see Kotler (1994, pg. 499-500), Hanson (1992), Shaw (1994), Merrin (1991) and the cites therein. These cites indicate that markup pricing is frequently used by multi-product businesses like members of the channel of distribution.

maximized. Second, equation (7) blends the information about the demand specification across the various stores in the market while allowing the parameters and posterior probabilities for each specification to be idiosyncratic to each store. The choice of a functional form for U( ) completes the specification of (7). To model the decision maker's risk preference, I choose the constant relative risk aversion (CRRA) utility function $U(\Omega_{ijs},w_i)= (\Omega_{ijs})^{\kappa}/\kappa$, $\kappa \le 1$. This function is also referred to as the power utility function (Ingersoll 1987, p.39). Setting $\kappa = 1$ yields the linear, risk neutral case, whereas $\kappa<1$ yields the risk averse case. The CRRA utility function is analytically tractable[12] with any probability distribution that has well defined moments.

4.3 Demand Specifications: Since the pricing problem being addressed is that of determining the wholesale price w_i, the specifications to be estimated should relate the demand faced by manufacturer i to the wholesale price w_i. The wholesale demand specifications are derived in a two step fashion. First, the retail demand for brand i in each store is specified. Second, the wholesale demand is derived from retail demand by recognizing that the retail price is a function of the wholesale price. The function that maps wholesale prices into retail prices reflects the retailers' price setting behavior. With respect to the retail demand functions, the literature survey suggested that the linear, semi-log and the double-log are three popular specifications. However, since there is ample evidence that the linear specification is a poor choice for a decision model (Little, 1970), only the semi-log and the double-log are retained for analysis as follows:

$$\ln(sales)_{its} = \alpha_{1is} + \sum_{b=1}^{6}\left(\beta_{1ibs}P_{bst} + \gamma_{1ibs}F_{bts}\right) + \sum_{q=2}^{4}\delta_{1iqs}Q_{qts} + \varepsilon_{1its}\,, \qquad (8)$$

$$\ln(sales)_{its} = \alpha_{2is} + \sum_{b=1}^{6}\left(\beta_{2ibs}\ln(P_{bts}) + \gamma_{2ibs}F_{bts}\right) + \sum_{q=2}^{4}\delta_{2iqs}Q_{qts} + \varepsilon_{2its}\,. \qquad (9)$$

In (8) and (9) i and b are brand indices, t is a time index, s is the store index, α_{1is} is the base level of sales for brand i in specification 1, β_{1ibs} is the coefficient of the price of brand b (b=1,...,6) in the sales equation of brand i in specification 1, P_{bts} is the retail price of brand b in time t, γ_{1ibs} is the coefficient of feature activity of brand b (b=1,...,6) in the sales of

[12] Hansen and Singleton (1983, p.250), note that this property of the power utility function has been exploited in the theoretical finance literature to arrive at closed form solutions for log normal returns.

brand i in specification 1, F_{bts} is a binary variable indicating whether brand b was featured in a retail flyer in time t, Q_{qts} are seasonal dummies, and ε are normally distributed error terms.

The choice of the independent variables requires some comments and justification. To capture the effects of own prices and feature ads and the effects of competitive prices and feature ads a complete set of price and feature effects were included in each specification. The dummy variables Q_{qt} were included to capture the seasonality that typically characterizes coffee sales (Sun 1985). The four seasonal quarters are defined as March-May, June-August, September-November and December-February. Three quarterly dummies are used with the intercept being the March-May quarter. In addition to these predictor variables the sales of the brands could have varied due to other marketing activities like displays and coupon drops, information about which is not available in the data. The impact of such omitted variables and the generalizability of the results are investigated in the diagnostic module. Finally, since the retailer's price setting rule is a markup over the wholesale price, the retail price $P_{bst} = m_{bst} w_{bt}$, where m_{bst} is the retail markup on brand b.[13] Substituting for P_{bst} in (8) and (9) yields the wholesale demand functions.

4.4 Priors: Diffuse priors are utilized for the specifications and the parameters. For the parameters diffuse priors are assigned on the basis of Jeffreys' multiparameter rule. For the specifications in (8) and (9), under the assumption of normal errors, application of Jeffreys' multi-parameter rule yields $\Pi\left(\Theta_{ijs}, \sigma^2_{ijs}\right) = \frac{1}{\sigma^{K+1}_{ijs}}$, where K is the number of independent variables. The use of diffuse priors will permit an objective comparison of the results of the Bayesian analysis with classical procedures.

5. Estimation Results

The posterior density of the parameters in (8) and (9) will be proportional to Equation(11) (ignoring the integral) in part B of the appendix. The marginal density of

[13] The July issue of Progressive Grocer used to be a "Supermarket Manual" issue, with data on category gross margins. This data indicates that the gross margin in supermarkets for ground coffee was about 10% in the early 1980's, yielding m_{bst}=1.1

Θ_{ijs} is multinormal with mean vector μ_{ijs} and variance-covariance matrix Σ_{ijs}. An interesting consequence of the diffuse priors is that the Bayesian parameter estimates will coincide with OLS estimates. Some key aspects of the estimation results for one of the stores[14] are presented in Table 2.

Table 2: Estimation Results

	R Squares		Posterior Probabilities		Price Elasticity		Equilibrium Wholesale Price	
Brand	Slog	Dlog	Slog	Dlog	Slog	Dlog	Slog	Dlog
B1	0.86	0.85	0.96	0.04	-4.61	-4.00	3.02	3.28
B2	0.78	0.77	0.84	0.17	-10.62	-9.68	2.67	2.69
B3	0.75	0.75	0.37	0.63	-3.24	-2.74	3.29	3.90
B4	0.73	0.73	0.51	0.49	-5.50	-4.33	2.92	3.18
B5	0.79	0.78	0.67	0.33	-5.82	-5.17	2.89	2.99
B6	0.70	0.69	0.77	0.23	-6.95	-5.20	2.83	3.15

Table 2 indicates that the R Squares across both specifications are quite comparable for all the brands. The posterior probabilities are interesting in that they lack a systematic pattern of strong updating towards any one specification. That is, none of the priors are updated to zero or one. For B4, there seems to be no updating at all. Overall across all brands, considerable residual uncertainty exists. The empirical findings of Bolton (1989) suggest that the finding of considerable residual uncertainty is not unique to the database used in this paper[15]. These findings suggest that there is not one specification that uniquely characterizes demand.

Another interesting feature of the estimation is that neither the semilog nor the double log specification dominates across all the brands. For example, for B1 and B2, the semilog seems slightly preferred whereas for B4, it seems to be an even split. Very slightly preferred are the semilog for B5 and B6, and the doublelog for B3. If one specification seemed to be more preferred consistently over all the brands, a case could be made that the choice of that specification is a safe bet. The posterior probabilities dispute

[14] The results from this store are qualitatively representative of the other stores.

[15] Bolton (1989), in testing a linear versus a double log specification for 44 sales equations, found that the likelihood test was inconclusive for 40 equations.

this notion and suggest such a "global bet" on a single specification across stores could lead to erroneous decisions for some of the brands. This finding points to the need for the testing of demand specifications on a brand by brand basis at each store. The posterior probabilities do provide a measure of specification uncertainty, which we can carry forward to the decision stage.

6. Diagnostic Testing

Before one proceeds to the pricing implications or the optimal pricing module, a closer look at some of the assumptions used in the model building process is necessary. Overall the signs and the magnitudes of the parameter estimates were consistent with theoretical expectations and what has been reported in the literature. The coherence of the statistical assumptions is explored by analysis of the residuals, predictive performance in holdout samples and tests of functional form. It is possible that only one of the specifications survives diagnostic testing. If this is indeed the case, it would simplify the optimal pricing module.

6.1 Residual Analysis: The statistical assumptions underlying the model are those of iid normal residuals with constant variance. The appropriateness of these assumptions was explored by graphical analysis of the residuals and formal tests. Increasing or decreasing bandwidth in a plot of the residuals against the dependent variable can be indicative of heteroskedasticity. Such patterns were not discerned for any of the specifications. Normal probability plots[16] did not indicate any gross departures from normality. Durbin Watson tests were conducted to ascertain if autocorrelation exists in the data. The Durbin-Watson statistic is reported in Table 3. The appropriate upper and lower critical values are 2.075 and 0.919 respectively. For brands 1, 2 and 6 the appropriate conclusion is that 1st order autocorrelation is not different from zero. For brands 3, 4, and 5 the D-W statistic falls slightly below the upper critical value but well over the lower threshold of

[16] Plot of the actual residual versus it's expected value under normality. If the residuals are approximately normal, the points should fall reasonably close to a straight line.

0.919 thereby rendering the test inconclusive. For these brands it seems that some extremely mild 1st order autocorrelation pervades both specifications.[17]

Table 3: Diagnostic Testing Results

	Durbin Watson Statistic		**Holdout MSE**		**P Test (α=0.01)**		
Brand	Slog	Dlog	Slog	Dlog	Slog	Dlog	Conclusion
B1	2.23	2.19	0.212	0.233	0.10	0.00	Slog
B2	2.33	2.39	0.905	0.988	0.35	0.06	Both
B3	1.89	1.88	0.195	0.188	0.24	0.55	Both
B4	2.02	2.05	0.382	0.384	0.64	0.60	Both
B5	1.82	1.84	0.225	0.227	0.70	0.25	Both
B6	2.53	2.54	0.986	1.03	0.32	0.09	Both

Another aspect of concern is the presence of outlying observations. Omitted variables may beresponsible for these outliers. The model may be more complete for the observations that remain after the outliers are deleted. Outliers were identified by various techniques,[18] the observations deleted and the specifications re-estimated. As expected the R-Squares improved for all the brand sizes, but the posterior probabilities, and the optimal prices were fairly insensitive to these outliers. In summary, the residual analysis does not reveal any serious violations of the statistical assumptions, and any one specification does not emerge as a clear winner based on these results.

6.2 Holdout Prediction: Since only 65 observations are available, it was infeasible to set aside a large part of the data set for predictive purposes. As an alternative I created 65 holdout observations by deleting one observation at a time, estimating the model on the

[17] A plot of category volume over time indicates that there may be some very mild trend in the data. When a trend variable is included in the model, it is not statistically significant, but it does result in a DW statistic that leads to the conclusion of no 1st order autocorrelation.

[18] In one approach, outliers were identified by graphing the standardized residuals. Interestingly the same observations were outliers for both specifications. In a second approach I deleted the 7 largest outliers (10% of the observations) and refit the models. In this case the outliers for the two specifications could be different sets of observations. In either case I was unable to eliminate a specification based on the refit results.

remaining 64 observations, and predicting the holdout observation. Table 3 contains the predictive mean square error obtained from this process. The MSE indicate that the specifications compare very well with each other, and neither specification shows overwhelming predictive superiority. These results are similar to the posterior probabilities tending not to update overwhelmingly towards any one specification. Comparing the MSE from Table 3 to the posterior probabilities in Table 2 reveals that the even though the Bayesian posterior probabilities were computed on the entire sample, they seem to be able to pick up the specification that predicts better in the holdout sample. For example, for brand 1 the semilog has lower MSE (0.212) versus the doublelog (0.233). The corresponding posterior probabilities are 0.96 and 0.04 respectively. This is reassuring, since it indicates that the posterior probabilities may be more than merely 'goodness of fit' criteria.

6.3 Tests of Functional Form: The semi-log and the double-log specifications are not nested within each other, but have the same dependent variable. The P test (Davidson and MacKinnon 1981) can be used to discriminate between them. The test involves a two way implementation. That is, to test the semilog against the double-log, we regress the residuals from the semilog against the residuals obtained from the doublelog. If the regression coefficent of the residuals is statistically significant, the null model is rejected. Then we reverse the test with the double-log as the null and the semilog as the alternate. What is reported in Table 3 is the P value of the regression coefficient with the column heading as the null model. For example, the second entry in the third column of the table is 0.00, and is the P value with the double-log as the null model and the semilog as the alternate. Since the P value is less than 0.01 we can reject the double-log model against the semilog alternative. At α=0.01 the test is inconclusive for five out of six brands. Increasing α to 0.05 does not improve the situation. However, the residual uncertainty at 5 percent may be too large to be ignored.[19]

[19] More importantly it is not clear that, in a "posterior sense", the uncertainty about the decision is 5 percent. This is because there may not exist a coherent many-to-one mapping that maps the p values from each pairwise test into a p-value for the battery of tests. Even if such a mapping exists, it may not correctly represent uncertainty in a posterior sense. See Berger and Selke (1987), for some theoretical evidence about the irreconcilability of p-values and posterior probabilities in testing point null hypothesis.

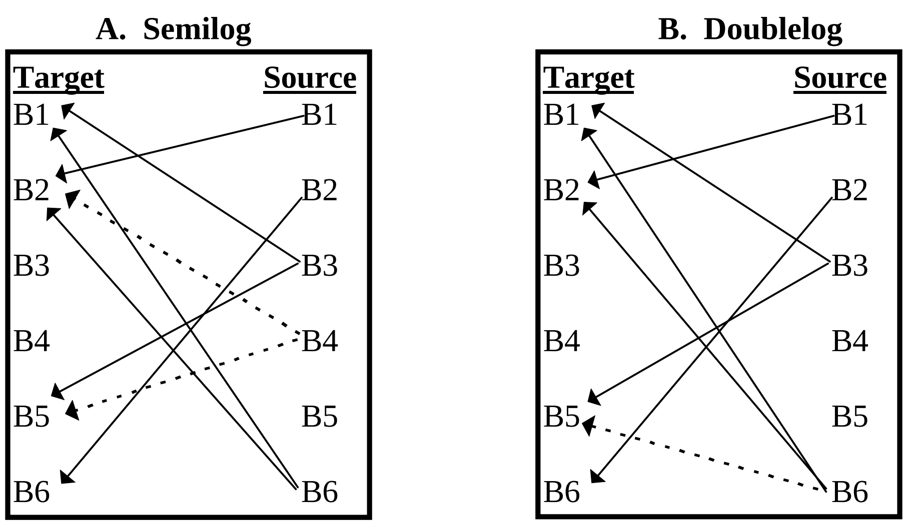
7. Pricing Implications

The previous sections have demonstrated that both specifications perform well in estimation and holdout samples and survive rigorous diagnostic testing. Before we proceed further it is natural to ask: What is the impact of specification uncertainty on pricing? If the different specifications make policy recommendations for pricing that are fairly similar, then, from a practical perspective, resolving specification uncertainty is a moot point. In what follows, I focus on what the different specifications indicate with respect price elasticities, cannibalization, the structure of price competition, and the profit maximizing price. Table 2 contains the price elasticity estimates for the various brands for each demand specification. The double-log has a constant elasticity equal to the price coefficient. For the semi-log, the price elasticity is the product of the price coefficient and the wholesale price. Table 2 reports the price elasticity averaged over all 65 observations. The table confirms that there is considerable inter-specification variation in price elasticities. For example, for brand 1, the estimated elasticity for the semilog model is -4.61, while for the doublelog it is -4.00, a considerable difference. This inter-specification variability in price elasticity is not unique to brand 1, and seems to pervade across all the brands examined. A pattern that does emerge across the brands is that the semilog estimates are more negative relative to the double log.

Table 2 also contains the equilibrium wholesale prices given that manufacturers sell across multiple stores. Parts C, D and E of the appendix contain the computational details. Again, a key feature of the table is the inter-specification variability regarding the equilibrium price. For example for B3, the slog model suggests \$3.29 per lb. whereas the double-log suggests \$3.90. While both models suggest price increases, the differences between the models are significant. So the question: *Which price should we implement*? remains unresolved. A similar dilemma arises when it comes to choosing the optimal price for the other brands.

Another issue of importance is the structure of price competition among the brands. I explore this issue by examining the statistical significance and the magnitude of the cross price coefficients. Figure 2 is a graphical representation of inter-brand competition in one store. Solid arrows are cross effects that are common across specifications. For example, the solid arrow from B1 to B2 in panel A indicates that the price coefficient of B1 was statistically significant in the sales equation of B2 at α=0.1. Dashed arrows are cross effects that are unique to the specification. For example the dashed arrow from B4 to B2 in panel A that this competitive interaction is unique to the semi-log model. Other unique interactions include the pair B4 and B5 in panel A and B6 and B5 in panel B. Taken together, the semi-log model attributes a higher level of competitive "clout" to B4 than does the double log. The semi-log also portrays B2 as being much more vulnerable. The double log on the other hand attributes more clout to B6 than does the semi-log. To summarize there is considerable variation in the inter-brand competitive structure across the specifications, and it seems that the question of "*who are my competitors?*" is not clearly resolved. In conclusion, with respect to cannibalization, competition and profit maximization the specifications present a mixed bag of results, sometimes complementing and at other times contradicting each other. This suggests that procedures that formally model the impact of demand uncertainty on pricing decisions could add value.

Figure 2: Inter-Brand Competition



8. Optimal Pricing Decisions

The time frame for the pricing is set to be Week 66, which corresponds to the first quarter, and accordingly, the seasonal dummies are set to zero. The BMM pricing solutions are computed for a risk neutral (κ=1) and a risk averse decision maker (κ=-1) by maximizing equation (7) for all brands (computational details in parts C, D and E of the appendix). The pricing solutions in Table 4 provides a number of insights about the functioning of the BMM approach.

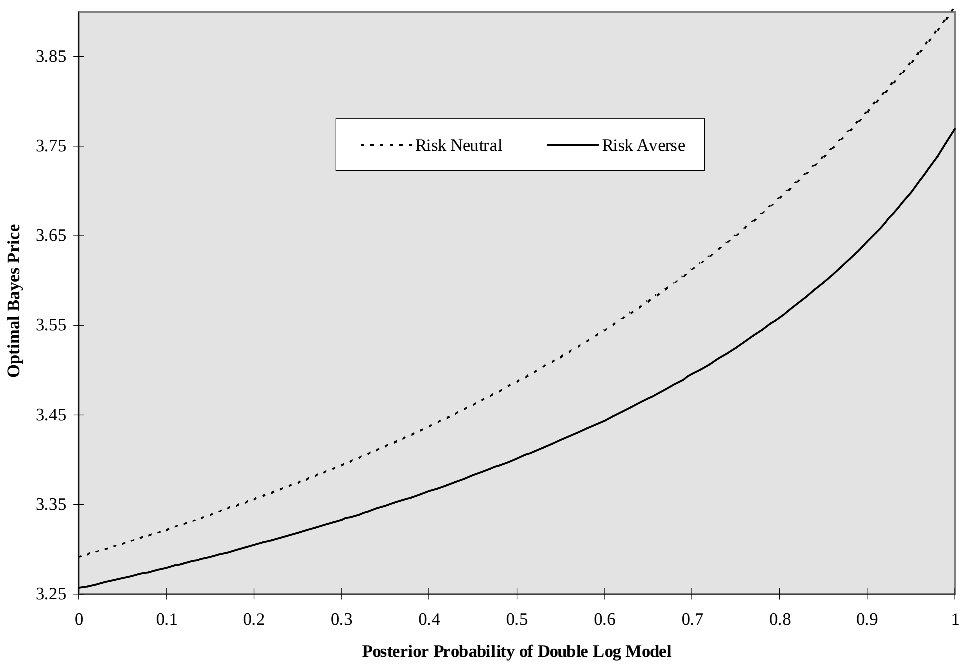
First and most importantly, irrespective of whether the posterior probabilities are conclusive or not, the BMM recommends a price that can be implemented. For example, for B3, the BMM pricing solution for the risk neutral decision maker is $3.57 which is a specific pricing recommendation even though there is no clear recommendation about which specification is appropriate. Another aspect of the solution is that the posterior probabilities do act as weights, tempering the influence a specification has on optimal pricing.

Table 4: BMM Pricing Solutions

Brand	Risk Neutral	Risk Averse
B1	3.02	2.95
B2	2.67	2.65
B3	3.57	3.46
B4	3.01	2.99
B5	2.92	2.91
B6	2.87	2.83

Another insight from the table is the divergence in the pricing actions of the risk neutral and the risk averse decision maker. Varying risk preferences do seem to have an impact on the pricing decision. While Table 4 presents the BMM pricing solutions, it does not provide insight into the divergences in the optimal pricing behavior of the risk averse and the risk neutral clients. In order to examine this closely the posterior probabilities were varied over a grid of points (ranging from 0 to 1) for κ=1 and κ=-1. In Figure 3 the optimal Bayes price for B3 is plotted against the grid of posterior probabilities of the double-log model. Figure 3 suggests that for this brand the risk neutral decision maker always charges a higher price than the risk averse decision maker. This result seems to be due to the mean-variance nature of the solution for expected utility. As we examine prices away from the mean price, the variance of the payoff increases due to increasing prediction error and makes a price increase less attractive from the perspective of the risk averse decision maker.

Figure 3: BMM Pricing Solutions for B3



9. A Monte Carlo Experiment

This section explores the *actual optimality* of the prices determined using the BMM approach. While the holdout predictions did provide some evidence that the posterior probabilities seemed to be able to detect the specification with better predictive performance in holdout samples, it does not establish to what extent the BMM approach recovers possible improvements in profits. This issue of recovery is explored in a Monte Carlo demand experiment. In order to ensure some degree of realism, the experiment was based on the parameters of the semilog and the doublelog models estimated using the six brands in the ground coffee data base. The procedure involved the following steps:

1. A specification (semilog or doublelog) and a brand were chosen and the posterior joint density of the specification parameters $\Pi(\Theta,\sigma^2)$ estimated. The posterior joint density is multivariate normal-inverse gamma.
2. Make a random draw from the inverse gamma density $\Pi(\sigma^2)$.

3. Make a random draw[20] from $\Pi(\Theta|\sigma^2)$ conditional on the value of σ^2 drawn in step 2.
4. Draw iid normal variates ε from $N(0,\sigma^2)$, where σ^2 is from the random draw in step 2. Compute $Y=X\Theta+\varepsilon$. Θ was drawn in step 3, and X is the matrix of the independent variables observed in the data and transformed for the specification.
5. Compute the current profit level based on current prices (Π_{CP}) and the true optimal profits (Π_{MAX}) for the vector Θ.
6. Implement the BMM approach (the specifications estimated are double log and semilog) on the data set created in step 4. Since the true underlying specification is known the actual profits that result from implementing the BMM (Π_{BMM}), the double log (Π_{DL}) and the semilog (Π_{SL}) prices can be computed.
7. Compute a recovery fraction of the procedure for this run. For example the recovery fraction of the BMM approach is defined as R_{BMM} and is given by $\frac{\Pi_{BMM}-\Pi_{CP}}{\Pi_{MAX}-\Pi_{CP}}$. In other words R_{BMM} measures the fraction of the total possible improvement recovered by the BMM procedure, and is bounded from above by 1.

Steps 1 to 7 were repeated for all specifications for all brands. In the actual implementation of step 3, thirty draws were made from $\Pi(\sigma^2)$. For each value of σ^2, thirty draws were made from $\Pi(\Theta|\sigma^2)$. The brand, specification, precision draw, parameter draw combination (6x2x30x30) resulted in 10,800 runs[21]. Figure 4 graphs the mean recovery rates.

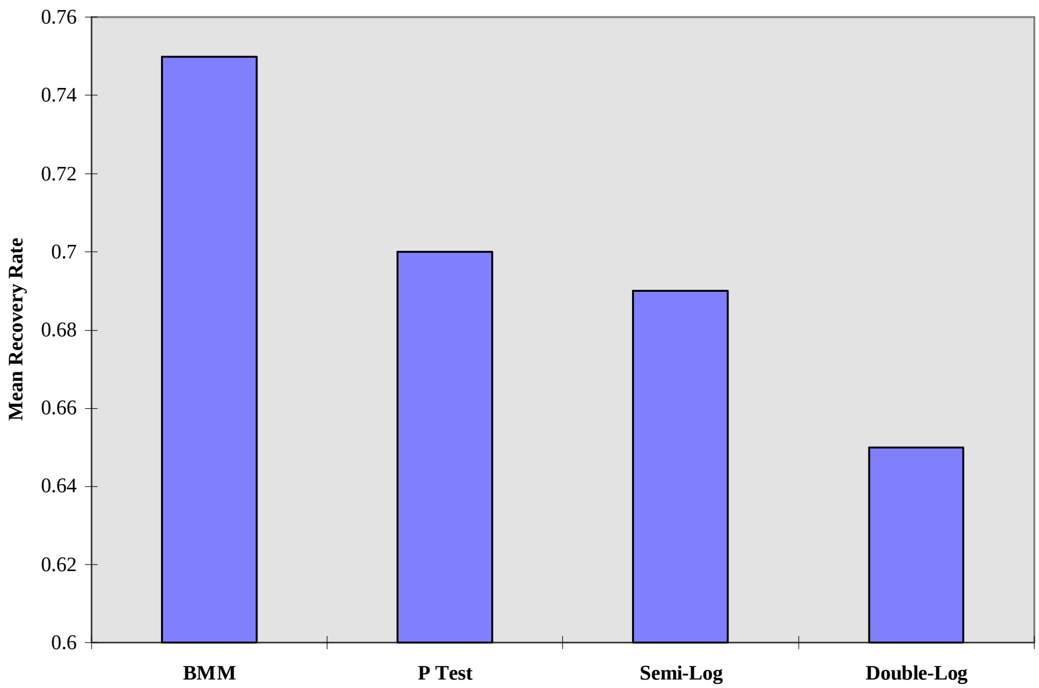
As the figure indicates, not only does the BMM approach perform better than the other approaches compared, it performs well in an absolute sense. The mean recovery rate for the BMM procedure is 0.75. That is on average, the BMM procedure recovered 75% of the potential profit improvement. Compared to this, the mean recovery rate is 0.69 for the semilog model, 0.65 for the double log model, and 0.70 for the P test. These

[20] This procedure (like Gibbs sampling) circumvents the need to make random draws from the multivariate normal-inverse gamma density. Draws are made from the tractable conditional and marginal densities. For an intuitive discussion of why such procedures will converge to the true density, see Casella and George (1992).

[21] Longer run sequences were conducted with this dataset and with data from the other 3 stores. The substantive implications are similar.

results offer some insight into the actual optimality of the BMM prices and the ability to recover potential profit improvements.

Figure 4: Recovery Rates in the Monte Carlo Experiment



10. Conclusions

Scanner data based models are being increasingly used for managerial decision making. It is important to realize that such decisions are conditional on the parametric specification. This paper proposes a Bayesian mixture model framework to formalize the impact of parametric uncertainty on pricing decisions. Key features of the framework are:

1. It presents a *unified theoretical approach* to estimation and pricing decisions,
2. The analysis can be conducted using subjective or diffuse priors,
3. Choice of a significance level is not required,
4. Utility structures reflecting the decision maker's risk attitude can be incorporated. The power utility function resulted in closed form solutions for expected utility,
5. Pricing decisions can be arrived at even when parametric demand uncertainty is unresolved.

The empirical implementation led to a number of insights some of which are listed below:

1. The posterior probabilities reveal that there can be considerable residual demand uncertainty in marketing data. This can translate into uncertainty about price elasticities, the structure of price competition and the optimal price.
2. Face validity measures, residual analysis or classical tests may lack the power to resolve this uncertainty. The "either or" structure of these tests emphasizes choosing one specification over the other, thereby being ill suited to situations characterized by residual demand uncertainty.
3. The results of a battery of these tests cannot be readily translated into p-values. Moreover the p-values may not correspond to posterior probabilities.
4. More than one specification may be necessary to characterize demand. Consequently, specification testing needs to be conducted for each brand at the store level. The posterior probabilities were more conclusive than the classical tests. Thus, the Bayesian approach seems to be better suited and more powerful for non-nested testing.
5. Using the BMM approach one can arrive at optimal pricing decisions even when demand uncertainty is unresolved. The posterior probabilities act as weights in the BMM pricing solutions. Specifications influence the optimal pricing in accordance to their weights.
6. There can be considerable divergence in the pricing actions of the risk averse and the risk neutral decision maker.
7. The results of the Monte Carlo experiment provide some reassurance about the ability of the BMM approach to recover potential profit improvements over current pricing.

The goal of the BMM framework is to assist managers in making robust pricing decisions by elaborating specifications via mixtures. In this paper the demand specifications utilized were intrinsically linear with conjugate prior densities. Recent developments in Bayesian computing (e.g. Gibbs sampling) make it possible to extend the BMM approach to include more complex models. Finally, though the focus in this paper has been on pricing, the BMM approach is potentially applicable to other elements of the marketing mix.

APPENDIX

A. The Box-Cox Analysis of Transformations: Equation (1) is a fairly general specification that can accommodate various transformations. Alternately, one can adopt the analysis of transformations presented by Box and Cox (1964). The Box-Cox model is given

$$y_n^{\lambda_{oj}} = \theta_{0j} + \sum_{i=1}^{I} \theta_{0j} x_{in}^{\lambda_{ij}} + \varepsilon_{in} \tag{10}$$

Although (10) has it's limitations in that it cannot accommodate certain specifications like the translog, it does encompass a fairly large class of transformations. Thus there will be considerable overlap between (10) and (1). In addition to presenting the power transformations on y, Box and Cox (1964) developed both likelihood and Bayesian estimation procedures. In their Bayesian approach they use an "arbitrary" diffuse prior by using an approximate consistency argument. This choice has been criticized on the grounds of outcome dependency (Nelder 1964; Fraser 1968; Lindley 1972, p.48; and Pericchi 1981). Also, Bayesian and Likelihood approaches could lead to contradictory inferences, leaving open the question as to which is the right inference to follow (Pericchi 1981). The alternate choice of prior, assigned in accordance with Jeffreys' multiparameter rule is not outcome dependent. There is no assumption of *a priori* dependence between the parameters. The posterior density can be derived by letting the prior information in the multivariate normal inverse-gamma posterior density approach the non-informative situation (Pericchi 1981). Further, likelihood and Bayesian inference will coincide.

B. Posterior Evidence for a Specification: Given the diffuse prior based on Jeffreys' multiparameter rule and a normal likelihood for the data, the evidence in favor of specification j is:

$$M(Y_{ijs}, X_{ijs}) = \iint \frac{1}{(2\pi)^{N/2}} \frac{1}{\sigma^{N+K+1}} \exp\left\{ -\frac{\left(Y_{ijs} - X_{ijs}\Theta_{ijs}\right)^t \left(Y_{ijs} - X_{ijs}\Theta_{ijs}\right)}{2\sigma_j^2} \right\} \left|J_{ijs}\right| d\Theta_{ijs} d\sigma_{ijs} \tag{11}$$

where $\left|J_{ijs}\right|$ is the Jacobian of the transformation on the dependent variable for brand i in the jth specification in store s. On solving the integral we get

$$M\left(Y_{ijs}, X_{ijs}\right) = \left\{2\pi^{\frac{K}{2}}\right\} \left\{ \frac{\Gamma\left(\frac{N}{2}\right)}{4\pi^{\frac{N}{2}}} \right\} \left\{ \frac{\left|J_{ijs}\right|}{\left|X_{ijs}^t X_{ijs}\right|^{-\frac{1}{2}} \left(S_{ijs}\right)^{\frac{N}{2}}} \right\}, \tag{12}$$

where $S_{ijs} = \left(Y_{ijs} - X_{ijs}\tilde{\Theta}_{ijs}\right)^t \left(Y_{ijs} - X_{ijs}\tilde{\Theta}_{ijs}\right)$, and $\tilde{\Theta}_{ijs} = \left(X_{ijs}^t X_{ijs}\right)^{-1} X_{ijs}^t Y_{ijs}$. Note that the term in the denominator of (12) involves the cross product matrix of the data. The magnitude of this term will reduce with data reducing transformations (for example the logarithmic transformation in the double log model) with a consequent favorable impact on the posterior probabilities. Posterior probabilities that are invariant to such transformations can be computed by appropriately loading the prior probabilities. In particular specifying $PR_{ijs} = \left| X_{ijs}^{t} X_{ijs} \right|^{\frac{1}{2}}$ provides the necessary invariance. The posterior probabilities for the specifications in the empirical implementation are computed using this approach.

C. Expected Utility for Log-Normal Returns: For the vector of independent variables $\bar{X}_{ijs}$, the predictive density of $\ln\left(\bar{y}_{ijs}\right)$ is normal with mean $\mu\left(\bar{X}_{ijs}\right) = \bar{X}_{ijs}\tilde{\Theta}_{ijs}$ and variance given by $V(\bar{X}_{ijs}) = \bar{X}_{ijs}^{t}\Sigma_{ijs}\bar{X}_{ijs}$. The density of $\bar{y}_{ijs}$ is lognormal (Aitchison & Brown 1957, p.7). Further from theorem 2.1 of Aitchison and Brown, it follows that the density of profits $\Pi(\Omega_{ijs})$ is lognormal. Further for $U\left(\Omega_{ijs}, w_i\right) = \frac{\Omega_{ijs}^{\kappa}}{\kappa}$, the density of utility is equal to the moment generating function of the log-normal distribution around the origin. This integral evaluates to

$$\exp\left[\kappa\mu\left(\bar{X}_{ijs}\right) + \frac{1}{2}\kappa^{2}V\left(\bar{X}_{ijs}\right)\right]\frac{(w_i - c_i)^{\kappa}}{\kappa}. \qquad (13)$$

(13) is substituted in lieu of the integral in (7). (13) is quasi concave and hence (7) is quasi concave since it is the sum of convex combinations of quasi concave functions. This guarantees that a unique maximum for (7) exists. However the solution for the maximum does not exist in closed form and I solve (7) numerically.

D. Manufacturers' Variable Costs (c_i): The spot prices of coffee beans in the New York exchange for May 1981 (as reported in *Coffee: 1984 Commodity Year Book*) was \$2.18/lb. The correlation between the spot price and the wholesale price of roasted ground coffee (obtained from *U.S. Bureau of Labor Statistics*) was 0.73, indicating the absence of any lagged structure. Progressive Grocer (1984) indicates that the packaging cost of a 1 lb. can was about \$0.25.

E. Solution Procedure to Obtain Equilibrium Prices: Since the objective function for each manufacturer (Eq. 7) is quasi concave, a non cooperative equilibrium exists (Friedman 1977). The solution was obtained using the following algorithm: [1] Solve for the vector of optimal wholesale prices of all brands (maximizing Eq. 7) given an initial set of wholesale prices. Denote the solution vector in this iteration by WP*(1); [2] Update the optimal prices of each manufacturer conditional on the prices of the other 5 brands contained in WP*(1). Denote the updated wholesale prices so obtained by WP*(2); [3] In the nth step the wholesale prices are updated conditioning on WP*(n-1), [4] Iterate till updating ceases. W*, the vector of prices in the final iteration is the non cooperative equilibrium since none of the manufacturers unilaterally alter their response to their competitors' best response. This numerical algorithm is an implementation of the sequential fictitious play solution procedure that have been proposed in the game theory literature. In the implementation the process was iterated upto a convergence tolerance level of $\varepsilon \leq 10^{-10}$. For a large grid of starting values the process converged to the same set of equilibrium prices.

References

Aitchison, J. and J.A.C. Brown (1957), "*The Lognormal Distribution*", Cambridge: University Press.

Allenby, G.M. (1991), "Hypothesis Testing with Scanner Data: The Advantage of Bayesian Methods," *Journal of Marketing Research*, 27, (November), 379-89.

Balasubramanian S.K. and D.C.Jain (1994), "Simple Approaches to Evaluating Competing Non-Nested Models in Marketing", *International Journal of Research in Marketing*, 11, 53-72.

Barry, C.B. and A.R. Wildt (1977), "Statistical Model Comparison in Marketing Research", *Management Science*, 24, (December), 387-92.

Berger, J.O. and T. Selke (1987), "Testing a Point Null Hypothesis: The Irreconcilability of P Values and Evidence", *Journal of the American Statistical Association*, 82, 397, (March), 112-22.

Berger, J.O. (1985), "*Statistical Decision Theory and Bayesian Analysis*", 2nd Edition, New York: Springer Verlag.

Blattberg, R.C. and S.K. Sen (1975), "A Bayesian Technique to Discriminate Between Stochastic Models of Brand Choice", *Management Science*, 21, 6, (February), 682-96.

Bolton, R.N. (1989), "The Relationship Between Market Characteristics and Promotional Price Elasticities", *Marketing Science*, 8, 2,153-169.

Box, G.E.P. and D.R. Cox (1964), "An Analysis of Transformations", (with Discussion) *Journal of the Royal Statistical Society*, Series B, 26, 211-52.

Casella, George and Edward I. George (1992), "Explaining the Gibbs Sampler", *The American Statistician*, 46, (August), 3, 167-73.

Cox, D.R. (1962), "Further Results on Tests of Separate Families of Hypothesis", *Journal of the Royal Statistical Society*, Series B, 24, 406-24.

Davidson, R., and J.G. MacKinnon (1981), "Several Tests for Model Specification in the Presence of Alternate Hypothesis", *Econometrica*, 49, 781-93.

Draper, N.H. and H. Smith (1981), "*Applied Regression Analysis*", Second Edition, New York: John Wiley.

Eliashberg, J. and J.R. Hauser (1985), "A Measurement Error Approach for Modeling Consumer Risk Preference", *Management Science*, 31, 1, (January), 1-25.

Fraser, D.A.S. (1968), "*The Structure of Inference*", New York: John Wiley.

Friedman, J.W. (1977), "*Oligopoly and the Theory of Games*", Amsterdam: New Holland.

Guadagni, P.M. and J.D.C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data", *Marketing Science*, 2, 3, (Summer), 203-38.

Hansen, P.L. and K.J. Singleton (1983), "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns", *Journal of Political Economy*, 91, 2, 249-65.

Hanson, W.A. (1992), "The Dynamics of Cost Plus Pricing", *Managerial and Decision Economics*, 13, 149-161.

Hauser, J.R. and G.L. Urban (1979), "Assessment of Attribute Importances and Consumer Utility Functions: von Neumann-Morgenstern Theory Applied to Consumer Behavior", *Journal of Consumer Research*, 5, (March), 251-62.

Ingersoll, Jr. J.E (1987), "*Theory of Financial Decision Making*", New Jersey: Rowman and Littlefield.

Jain, D.C. and N.Vilcassim (1989), "Testing Functional Forms of Market Share Models Using the Box-Cox Transformation and the Lagrange Multiplier Approach", *International Journal of Research in Marketing*, 6, (March), 95-107.

Jeffreys, H. (1939), "*Theory of Probability*", Third Edition, Oxford: Clarendon Press.

Kalyanam, K. (1993) "Optimal Pricing Decisions Under Demand Uncertainty: A Bayesian Mixture Model Approach", Unpublished P.h.D Dissertation, Purdue University.

Kotler, P. (1994), "*Marketing Management: Analysis, Planning, Implementation, and Control*", Eighth Edition, New Jersey: Prentice Hall.

Lenk, P.J. and A.G. Rao (1990), 'New Models from Old: Forecasting Product Adoption by Hierarchical Bayes Procedures", *Marketing Science,* 9, 1, 42-53.

Lilien, G.L., A.G. Rao and S. Kalish (1981), "Bayesian Estimation and Control of Detailing Effort in a Repeat Purchase Diffusion Environment", *Management Science*, 27, 5, (May), 493-506.

Lindley, D.V. (1976), "Bayesian Statistics, A Review", SIAM, Philadelphia, Regional Conference Series in Applied Mathematics, 2. Philadelphia: Society for Industrial and Applied Mathematics.

Little J.D.C. (1970), "Models and Managers, the Concept of a Decision Calculus", *Management Science*, 16, 8, B466-85.

MacKinnon. J.G. (1983), "Model Specification Tests Against Non-Nested Alternatives", *Econometric Reviews*, 2, 85-110.

Merrin, E. (1991), "Computer Superstores: Mother Lode or Flash in the Pan?", *Channel Marker Publications*, Palo Alto, CA, 41.

Moriarty, M.M. (1990), "Boundary Value Models for the Combination of Forecasts", *Journal of Marketing Research*, 27, (November), 402-17.

Nelder, J.A. (1964), Contribution to discussion of the paper by G.E.P.Box and D.R. Cox, *Journal of the Royal Statistical Society*, Series B, 26, 244-52.

Pericchi, L.R. (1981), "A Bayesian Approach to Transformations to Normality", *Biometrika*, 68, 1, 35-43.

Raiffa, H. and R. Schlaiffer (1961), "*Applied Statistical Decision Theory*", Boston: Division of Research, Graduate School of Business Administration, Harvard University.

Rust, R.T. and D.C. Schmittlein (1985), "A Bayesian Cross-Validated Likelihood Method for Comparing Alternative Specifications of Quantitative Models", *Marketing Science*, 4, 1, (Winter), 20-40.

Shaw, R.T. (1995), "*Marketing: Core Concepts*," Cincinnati: Southwestern College Publishing.

Sun, T. (1985), "The Demand for Coffee and Tea", Staff Report, Economic Research Service, U.S. Department of Agriculture.

Tellis, G. (1988), "The Price Elasticity of Selective Demand: Meta-Analysis of Econometric Models of Sales", *Journal of Marketing Research*, 25, (November), 331-41.

Theil, H. (1971), "*Principles of Econometrics,*" New York: John Wiley.

Wittink, D.R., et.al (1988), "SCAN*PRO©: The Estimation, Validation and Use of Promotional Effects Based on Scanner Data", Working Paper, Johnson Graduate School of Management, Cornell University, Ithaca, New York.

Zellner, A. (1971), "*An Introduction to Bayesian Inference in Econometrics*", New York: John Wiley.