

Do Switching Costs Make Markets Less Competitive?

by

Jean-Pierre Dubé
Günter J. Hitsch
Peter E. Rossi*

Graduate School of Business
University of Chicago

June 2006
Revised, September 2008

Keywords: switching costs, dynamic oligopoly, pricing
JEL Classifications: D1, D4, L1, M3

Abstract

The conventional wisdom in economic theory holds that switching costs make markets less competitive. This paper challenges this claim. We formulate an empirically realistic model of dynamic price competition that allows for differentiated products and imperfect lock-in. We calibrate this model with data from frequently purchased packaged goods markets. These data are ideal in the sense that they have the necessary variation to separately identify switching costs from consumer heterogeneity. Equally important, consumers exhibit inertia in their brand choices, a form of psychological switching cost. This makes our results applicable to the broad range of products that are distinctly identified (i.e. branded) rather than just to those products for which there is a product adoption cost or explicit switching fee. In our simulations, prices are as much as 18 per cent lower with than without switching costs.

*Jean-Pierre Dubé is Sigmund E. Edelstone Professor of Marketing, University of Chicago Graduate School of Business, 5807 S. Woodlawn Avenue, Chicago, IL 60637 (e-mail: jdube@gsb.uchicago.edu). Günter J. Hitsch is Associate Professor of Marketing, University of Chicago Graduate School of Business, 5807 S. Woodlawn Avenue, Chicago, IL 60637 (e-mail: guntj@gsb.uchicago.edu). Peter E. Rossi is Joseph T. and Bernice S. Lewis Professor of Marketing and Statistics, University of Chicago Graduate School of Business, 5807 S. Woodlawn Avenue, Chicago, IL 60637 (e-mail: peter.rossi@chicagogsb.edu). The authors are grateful for comments and suggestions from Alan Sorenson, Matt Gentzkow, Ariel Pakes, Peter Reiss, and seminar participants at the University of British Columbia, the University of Chicago structural IO lunch, Duke University, Erasmus University in Rotterdam, HEC Montreal, New York University, Northwestern University, the University of Minnesota, Universiteit van Tilburg, Yale University, the Canadian Competition Bureau, the Chicago Federal Reserve Bank, the 2006 NBER summer I.O. meetings, the 2006 SED conference, the 2007 Summer IO conference at UBC and Yahoo! Inc.. Support from the Kilts Center for Marketing, Graduate School of Business, University of Chicago is gratefully acknowledged.

Introduction

Surveying the theoretical literature on switching costs, Klemperer (1995, p. 527) and Farrell and Klemperer (2006, p. 23) conclude that there is a “strong presumption” that switching costs make markets less competitive. We propose a model with switching costs that can be calibrated from actual consumer panel data. For levels of switching costs found in our data, we find that equilibrium prices fall in the presence of switching costs. We argue that the conventional wisdom may not be applicable to empirically relevant models even with high switching costs.

We work with a model that captures the main elements of empirical environments in which switching behavior is observed. Typically, we observe switching costs in markets with differentiated products and with sellers that are not subject to some terminal (i.e. finite horizon) trading period. Switching in spite of the presence of switching costs is an empirical regularity in many consumer product markets. Often, switching occurs even though the relative prices of products remain roughly constant. Thus, an empirically viable model must allow for differentiated products (in contrast to Farrell and Shapiro (1988) and Padilla (1995)) and imperfect lock-in (compare with Beggs and Klemperer (1992)) in an infinite horizon setting.

Numerical simulations with a simplified version of our empirical model reveal that prices fall for intermediate levels of the switching cost relative to an environment with zero switching costs. The incentive for a firm to lower its price and “invest” in customer acquisition is found to outweigh the incentive for a firm to raise its price and “harvest” its existing customer base. This seemingly counter-intuitive finding reflects the strategic effects of firms lowering their prices to defend themselves against other firms’

attempts to steal customers. However, for large enough switching costs levels, the strategic effects are dampened and equilibrium prices rise. In general, these preliminary numerical findings suggest that the impact of switching costs on equilibrium prices is an empirical question about the magnitude of switching costs.

Switching costs are typically not directly observed. Instead, the analyst must infer their magnitude from the observed switching behavior of consumers. Consumers are also typically heterogeneous in their baseline preferences for products. To separate consumer-specific switching costs from brand preferences, panel data with a long time dimension and some source of exogenous switching is required. Panel data on the purchases of branded, frequently purchased products is well-suited to this task. Switching costs enter demand models used for these types of products in the same way as in the applied theory literature on switching costs. Frequent price reductions or sales induce switching in the panels, allowing us to separately identify heterogeneity and switching costs. We estimate the demand model from data on two categories of frequently purchased consumer products (refrigerated orange juice and margarine), and then compute the price equilibrium¹.

We build on a large empirical literature that has documented the existence of *brand loyalty* or simply *state-dependence* in consumer choice (c.f. Erdem (1996), Roy, Chintagunta, and Haldar (1996), Keane (1997), Seetharaman, Ainslie, and Chintagunta (1999), Seetharaman 2004 and Dubé et al. (2006)). However, this literature has not explored the implications of state dependence for equilibrium pricing. Che et al. (2006) investigate how far boundedly rational firms look ahead when making pricing decisions

in markets with state dependence. They do not consider how consumer loyalty affects equilibrium prices.

Our work is also related to two recent working papers, Doganoglu (2004) and Viard (2003), both of which provide theoretical counter-examples in which switching costs can lower equilibrium prices. However, their examples consist of stylized OLG specifications chosen based on analytical tractability, rather than their ability to produce empirical realistic behavior. In contrast, we work with the classic discrete-choice specification and long-lived consumers, the foundation of the majority of demand systems used in current marketing research and applied industrial organization. We show that switching costs can in fact lower equilibrium prices in the context of state-dependent demand for consumer package goods. In contrast, Viard's (2005) empirical results do not contradict the conventional wisdom that switching costs make markets less competitive.

We interpret our estimates of state dependence as arising from a psychological switching cost. Of course, switching costs can come from a variety of sources including product adoption costs, shopping/search costs, and psychological sources. For example, it is true that razor companies create switching costs between brands of razor blades by making handles or razors that only fit their blades. This certainly is one source of switching costs, but not necessarily the largest or most universal. The mere purchase/consumption of a distinctly identified (e.g. branded) product can create a switching cost. Klemperer (p. 518) points this out when he cites “psychological costs of switching, or non-economic “brand-loyalty” as an important example of switching costs (see also Farrell and Klemperer p. 15). These psychological costs are often thought to

¹ In previous work, Dubé et al. (2006), we used the same demand specification to study the category monopolist's problem. Clearly, competitive forces can alter equilibrium prices due to strategic effects and this

come from the well-known phenomenon of “cognitive dissonance” in which consumers change their preferences to “rationalize” previous choices. These psychological sources of switching costs are applicable to a much broader array of products than a switching cost narrowly defined as a monetary fee or learning costs. Both psychological switching costs and a monetary switching fee will give rise to an observationally equivalent form of inertia in consumer purchases.

Our estimated switching costs are on the order of 15 to 19 per cent of the purchase price of the goods. When these switching costs are used in model simulations, equilibrium prices decrease relative to prices without switching costs. This prediction is very robust to variation in the parameter values. In particular, if switching costs are scaled up to four times those inferred from our data, we still find that prices decline. We observe price reductions of up to 18 percent in the presence of switching costs. *Model*

The model consists of single-product firms competing for consumers with switching costs by pricing differentiated products. Each firm sets a pricing policy to maximize the discounted sum of profits over an infinite horizon. The solution concept for this game is Markov perfect equilibrium (MPE). The goal is to study the effects of consumer switching costs on pricing in the context of a model that generalizes much of the empirical research on differentiated products demand estimation. Unlike much of the established theoretical literature, we allow for product differentiation and imperfect lock-in—the possibility that consumers switch away from products they have previously purchased—which are features commonly present in actual markets. In addition, we include a random utility component which allows for the possibility that consumers switch products even when relative prices are not changing.

is the focus here.

We emphasize that this paper focuses on markets where switching costs are actual opportunity costs to consumers, not on markets with customer recognition where firms charge different prices to new and previous customers. In the literature on customer recognition, unlike in the literature on switching costs, it is well-established that equilibrium prices can be lower than if firms are unable to price discriminate between new and old customers (see Fudenberg and Villas-Boas 2006 for a survey).

After developing the model in general, we briefly explore a simple case to illustrate how these features of consumer choice influence pricing in the presence of switching costs. The advantage of this simple model is that it simplifies computation of an equilibrium, which means we can easily explore a number of comparative static exercises.

Demand

Demand derives from a population of consumers who make discrete choices from J product alternatives and an outside option (i.e. no-purchase). For simplicity, we drop the consumer-specific index. In each period t , a consumer is loyal to one product, $s_t \in \{1, \dots, J\}$. If the consumer is currently loyal to product j , $s_t = j$, and purchases product $k \neq j$, then her loyalty state changes, $s_{t+1} = k$. If the consumer chooses product j or the outside option, then $s_{t+1} = s_t$, i.e. the consumer's brand loyalty remains

unchanged. Conditional on price p_{jt} and her current loyalty state s_t , the consumer's utility index from the choice of a product j at time t is

$$u_{jt} = \delta_j + \alpha p_{jt} + \gamma I\{s_t \neq j\} + \varepsilon_{jt}. \quad (1)$$

The demand model in (1) has been used extensively in the empirical literature on consumer package goods (c.f. Erdem (1996), Keane (1997), and Shum (2004)). We assume that the random utility component ε_{jt} is i.i.d. Type I Extreme Value distributed. If the consumer is loyal to j but buys product $k \neq j$, she foregoes the utility component γ . Thus, she implicitly incurs a switching cost. Note that the consumer behavior associated with switching cost/product loyalty, γ , is different from the consumer behavior associated with the brand intercept, δ_j . An increase in the brand intercept will always increase the probability of purchase of the j th brand, while the switching cost parameter will only increase the purchase probability if the consumer is currently loyal to j .

Let $U(j, s_t, p_t)$ denote the deterministic component of the utility index, such that $u_{jt} = U(j, s_t, p_t) + \varepsilon_{jt}$. The utility from the outside alternative is $U_{0t} = U(0, s_t, p_t) = \varepsilon_{0t}$.

The consumer's choice probability has the logit form

$$P_j(s_t, p_t) = \frac{\exp(U(j, s_t, p_t))}{\sum_{k=0}^J \exp(U(k, s_t, p_t))}. \quad (2)$$

Demand parameters in (2), $\theta = (\delta_1, \dots, \delta_J, \alpha, \gamma)$, are consumer-specific with N "types" in the market. The behavior of each consumer type n is fully summarized by the taste vector, θ^n . The probability of buying product j by a consumer of type n in state

s_t , for example, is denoted by $P_j(s_t, p_t; n)$. We assume that for each consumer type, there is a continuum of consumers in the market with mass μ_n .

At any point in time, the market is summarized by the distribution of consumers over types and loyalty states. Let $x_{jt}^n \in [0, 1]$ be the fraction of consumers of type n who are loyal to product j . The vector $x_t^n = (x_{1t}^n, \dots, x_{Jt}^n)$ summarizes the distribution over loyalty states for all consumers of type n , and $x_t = (x_t^1, \dots, x_t^N)$ summarizes the state of the whole market. X denotes the state space.

Aggregate demand is obtained by summing consumer level demand over consumer types and loyalty states. Aggregate demand for product j is given by

$$D_j(x_t, p_t) = \sum_{n=1}^N \mu_n \left(\sum_{k=1}^J x_{kt}^n P_j(k, p_t; n) \right). \quad (3)$$

Evolution of the State

The distribution of consumer loyalty states, $x_t = (x_t^1, \dots, x_t^N)$, summarizes all current-period payoff-relevant information for the firm, and describes the state of the market. The transition of the aggregate state can be derived from the transition probabilities of the individual states. Conditional on a price vector p_t , we can define a Markov transition matrix $Q(p_t; n)$ with elements

$$Q_{jk}(p_t; n) = \Pr\{s_{t+1} = j \mid k, p_t; n\},$$

where

$$\Pr\{s_{t+1} = j \mid s_t, p_t, n\} = \begin{cases} P_j(s_t, p_t; n) + P_0(s_t, p_t; n) & \text{if } j = s_t, \\ P_j(s_t, p_t; n) & \text{if } j \neq s_t. \end{cases} \quad (4)$$

$Q_{jk}(p_t; n)$ denotes the probability that a consumer of type n who is currently loyal to k will become loyal to product j . The whole state vector for type n then evolves according to the Markov chain

$$x_{t+1}^n = Q(p_t; n)x_t^n.$$

Consumers can change loyalty states but not types such that the overall market state vector x_t also evolves according to a Markov Chain with a block diagonal transition matrix. The evolution of the state vector is deterministic, and we denote the transition function by f , $x_{t+1} = f(x_t, p_t)$.

Firms

We consider a market with J competing single-product firms. Time is discrete, $t = 0, 1, \dots$. Conditional on all product prices and the state of the market, x_t , firm j 's current-period profit function is: $\pi_j(x_t, p_t) = D_j(x_t, p_t) \cdot (p_j - c_j)$. c_j is the marginal cost of production, which does not vary over time. Firms compete in prices and choose Markovian strategies,² $\sigma_j: X \rightarrow \mathbb{R}$, that depend on the current payoff-relevant information, summarized by x . Firms discount the future using the common factor β , $0 \leq \beta < 1$. For a given profile of strategies, $\sigma = (\sigma_1, \dots, \sigma_J)$, the PDV of profits,

² This assumption rules out behavior that conditions current prices also on the history of past play, and thus collusive strategies in particular.

$\sum_{t=0}^{\infty} \beta^t \pi_j(x_t, \sigma(x_t))$, is well-defined. Conditional on a profile of competitor's strategies, σ_{-j} , firm j chooses a pricing strategy that maximizes its expected value. Associated with a solution of this problem is firm j 's value function, which satisfies the Bellman equation

$$V_j(x) = \max_{p_j \geq 0} \left\{ \pi_j(x, p) + \beta V_j(f(x, p)) \right\} \quad \forall x \in X. \quad (5)$$

In this equation, the price vector consists of firm j 's price and the prices prescribed by the competitor's strategies, $p = (\sigma_1(x), \dots, \sigma_{j-1}(x), p_j, \sigma_{j+1}(x), \dots, \sigma_j(x))$. Therefore, the Bellman equation (5) depends on the pricing strategies chosen by the competitors.

We use Markov perfect equilibrium as our solution concept. A Markov perfect equilibrium in pure strategies is defined by a pricing strategy for each firm, σ_j^* , and an associated value function, V_j , such that

$$V_j(x) = \max_{p_j} \left\{ \omega_j(x, p_j, \sigma_{-j}^*(x)) + \beta V_j(f(x, p_j, \sigma_{-j}^*(x))) \right\}$$

for all states, x , and firms. That is, in each subgame starting at x , the firm's strategy is a best response to the strategies chosen by its competitors. For a simple version of the model which we will explore in the next section, we can prove the existence of a pure strategy price equilibrium. However, we cannot prove that a pure strategy equilibrium exists in general.³ We establish the existence of a pure strategy equilibrium computationally on a case-by-case approach. In Appendix E (on-line), we describe the numerical algorithm used to compute the price equilibrium.

A Simplified Model

Let us briefly consider a simplified variant of the model discussed above for the purpose of building an intuition as to why switching costs can lead to lower equilibrium prices. In section 3, we will return to the model with many consumer types and base our pricing computations on empirical estimates of this model.

We assume there is exactly one consumer in the market, who chooses among the J products and an outside option in each period as in (1) above. With only one consumer, equilibrium computations are simplified greatly, facilitating the comparative statics necessary to develop an intuition as to the role of switching costs. Here we are following a number of recent papers that use computational methods to establish properties of various theoretical models (c.f. Doraszelski and Satterthwaite (2005) and Besanko, Doraszelski, Kryukov and Satterthwaite (2007)).

The loyalty variable of this consumer, $s_t \in X = \{1, \dots, J\}$, summarizes all current-period payoff-relevant information, and describes the state of the market. Conditional on all product prices and the state of the market, firm j receives the expected current-period profit $\pi_j(s_t, p_t) = P_j(s_t, p_t) \cdot (p_{jt} - c_j)$. While we can show the existence of a pure-strategy equilibrium (see appendix A.1), we cannot characterize the equilibrium policies analytically. Instead, we solve the game numerically for different parameter values.

³ Even in static games of price competition, restrictions on the distribution of consumer tastes need to be imposed to establish the existence of a pure strategy equilibrium (Caplin and Nalebuff 1991). In general, the “non-parametric” distribution of tastes that our model allows for does not obey these restrictions.

Motivated by this paper, Cabral (2007) derives theoretical results which support our computations.

Equilibrium Price Computations with the Simplified Model

We now explore the predictions of the simplified pricing model. To keep the exposition as simple as possible, we focus on symmetric games with two firms. Each firm has the same utility intercept and marginal production cost. In a symmetric equilibrium, $\sigma_1^*(1) = \sigma_2^*(2)$ and $\sigma_1^*(2) = \sigma_2^*(1)$. We therefore only need to know firm 1's pricing policy to characterize the market equilibrium.

We first consider the case of homogenous products (i.e. with $\varepsilon_{ji} = 0$). In this case, we can establish theoretically (see appendix A.2) that switching costs allow firms to raise prices above the baseline Bertrand outcome, where $p = c$. In particular, there is an equilibrium where the firm that possesses the loyal customer increases its price above cost by the value $v = (1 - \beta)\gamma$. v is the flow value of the switching cost. If the firm charges an even higher price, the competitors could poach the customer by subsidizing the switching cost, incurring a loss in the current period, and recouping this loss by pricing above cost in the future. In summary, if products are not differentiated, then we find that switching costs make markets less competitive, as predicted by much of the previous literature.

We now turn to the case of differentiated products and switching in equilibrium. In the case of product differentiation, the customer sometimes switches, and therefore we

characterize the equilibrium outcome by the average transaction price paid, conditional on a purchase:

$$\hat{p}^a = \frac{P_1(1, \sigma^*(1)) \cdot \sigma_1^*(1) + P_2(1, \sigma^*(1)) \cdot \sigma_2^*(1)}{P_1(1, \sigma^*(1)) + P_2(1, \sigma^*(1))}.$$

That is, \hat{p}^a is the expected price paid in state $s_t = 1$, which—due to symmetry—is the same as the expected price paid in state $s_t = 2$. $P_1(\cdot, \cdot)$, $P_2(\cdot, \cdot)$ are the probabilities of choice of each product conditional on price and loyalty state.

Figure 1 shows the relationship between the level of switching costs and the average transaction price for the case of $\delta_j = 1$, $c_j = 0.5$, $\alpha = 1$. We find that prices initially fall and then rise for larger switching cost levels. Indeed, only for switching cost levels larger than 4 does the average transaction price exceed the transaction price without switching costs. Although not reported, the main result that prices are decreasing-then-increasing in the level γ was found to be robust to the exclusion of the outside good and to the degree of switching (the scale of the intercept and price coefficients in (1)). These results show that the conjectured effect of switching costs on prices—switching costs make markets less competitive—need not be true in a model that is a simplified variant of a widely used class of empirical demand models.

To understand our results, recall that under competition with switching costs, firms face two incentives that work in opposite directions (c.f. Klemperer 1987). First, firms can “harvest” a loyal customer by charging higher prices. Second, firms can “invest” in future loyalty by lowering current prices. Our results imply that either force can dominate in equilibrium. Imperfect lock-in ($\gamma < \infty$) and the random component in consumer tastes, features common to many empirical models of differentiated products

demand, stimulate the incentive to cut prices to attract competitors' loyal customers. Anticipating this incentive, the competitor lowers its price to prevent the customer from switching. In some instances, this downward-pressure on prices overshadows the upward-pressure from harvesting. In contrast, Beggs and Klemperer (1992) only consider the case of perfectly locked-in consumers ($\gamma = \infty$). In their specification, the incentive to harvest will always outweigh the incentive to invest.

To illustrate the role of the investment motive, we examine the extreme case where only the “harvesting” incentive is present. To exclude the investment motive, we consider competitors who do not anticipate the future benefits from lowering current prices, and hence set prices in a myopic ($\beta = 0$) fashion. Figure 1 shows the average transaction price paid under this scenario, and allows us to compare the pricing outcomes with fully rational, forward-looking firms and myopic decision makers. After eliminating the investment motive, prices always rise in the degree of switching costs—switching costs make markets less competitive. In general, the average transaction price under competition with forward-looking firms is always lower than the average price under myopic competition.

One of the difficulties in interpreting the comparative static result displayed in Figure 1 is that an increase in γ not only increases the switching cost, it also leads to an increase in the total market size (i.e., a decrease in the outside good share). Since our interest lies in the role of switching behavior, we construct an adjusted comparative static exercise in which γ does not influence the outside share and leaves the total market size constant. Table 1 indicates that we still observe the main result of decreasing-then-increasing prices. In addition, Table 1 reports the probability that the consumer will

remain loyal. It is remarkable that even when the consumer will remain loyal with probability .981, equilibrium prices are still below those that would occur without switching costs. This suggests that results in the literature that prices rise are closely linked with assumptions of perfect lock-in.

In Table 2, we report the impact of switching costs on firm profits. We consider profits from both the oligopolistic equilibrium prices as well as profits that would occur if prices were fixed at the levels obtained under zero switching costs ($\gamma = 0$). These profit values are found in the rightmost two columns of the table. Interestingly, in this latter scenario, the firm with the loyal customer is strictly better-off when the switching cost increases. However, when the firms re-optimize their prices, both firms can be strictly worse-off if the switching cost leads to lower equilibrium prices. That is, the *strategic effect* of price competition on profits outweighs the *direct effect* of switching costs on profits, holding prices constant. Thus, the investment motive under competition outweighs the harvesting incentive. This outcome is an instance of a “Bertrand supertrap,” as analyzed in Cabral and Villas-Boas (2005) for finite-horizon games.

In Appendices C and D (on-line), we discuss two additional variants of the simple model to check the robustness of our results to other model features typically considered in the theoretical literature. We consider separately the impact of forward-looking consumer behavior and overlapping generations of consumers on equilibrium prices. The main conclusion, that switching costs do not necessarily lead to higher prices, is robust to these different model formulations as well as a wide range of parameter values.

In summary, we have shown that, contrary to the conventional wisdom, switching costs can toughen price competition. We have also shown that this result arises from the

dynamics associated with the investment motive. In particular, when there is some random switching in equilibrium, firms may compete both to attract and retain the customer. When the strategic effects are large enough, they can cause the investment motive to outweigh the harvesting motive, leading to lower equilibrium prices. In the next section, we investigate whether this result will still hold when we consider a richer model that generalizes much of the empirical research on demand estimation.

Empirical Model and Estimation

We now explore the impact of consumer switching costs on our full model with many “types” of consumers. Recall that for the simple version of the model with a single consumer, we observed that equilibrium prices can be *lower* with switching costs than without for a wide range of parameter values. Therefore, the impact of switching costs on prices is an empirical matter regarding the magnitude of switching costs consistent with actual consumer behavior. Hence, we calibrate our analysis of the full model using actual empirical estimates of the joint-distribution of preferences and switching costs, θ , for a population of heterogeneous consumers. In the sections below, we describe the data and the procedure used to estimate the demand parameters.

Econometric Model

For the full model, the probability that consumer h chooses alternative j given loyalty to product k is given by

$$P(j | s = k; \theta^h) = \frac{\exp(\delta_j^h + \alpha^h p_j + \gamma^h I\{s = j\})}{1 + \sum_{k=1}^J \exp(\delta_k^h + \alpha^h p_k + \gamma^h I\{s = k\})}. \quad (6)$$

One might argue that the switching cost, γ^b , could be modeled using a much more complex function of a consumer's purchase history. However, as discussed previously, there is a well-established precedent for using this specification in the empirical literature devoted to package goods demand. Furthermore, this specification is identical to the one routinely used in the applied theory literature on switching costs. Other specifications of switching costs or state dependence involve more general functions of past choices. While these are easily estimable (c.f. Dube et al (2008)), computing equilibrium prices may be difficult due to an increase in the dimension of the state space required.

To accommodate differences across consumers in (6), we use a potentially large number of consumer types and a continuum of consumers of each type. A literal interpretation of this assumption is that the distribution of demand parameters is discrete but with a very large number of mass points. In the consumer heterogeneity literature (c.f. Allenby et al 1999), continuous models of heterogeneity have gained favor over models with a small number of mass points. The distinction between continuous models of heterogeneity and discrete models with a very large number of mass points is largely semantic. In fact, some non-parametric methods rely on discrete approximations. Our approach will be to specify a very flexible, but continuous model of heterogeneity and then exploit recent developments in Bayesian inference and computation to use draws from the posterior of this model as “representative” of the large number of consumer types. Each consumer in our data will be viewed as “representative” of a type. We will use MCMC methods to construct a Bayes estimate of each consumer's coefficient vector.

It is well known (c.f. Heckman 1981 and Keane 1997) that state dependence and heterogeneity can be confounded. In the current context, a mis-specified tightly

parametric model of heterogeneity could lead to spurious findings of switching costs. The state-of-the-art in this literature (cf. Keane 1997) is to use normal models of heterogeneity. There is good reason to believe that there may be substantial departures from normality for the distribution of choice model parameters across consumers. For example, there may be sub-populations of consumers with different preferences for different brands. This might lead to multimodality in the distribution of the intercepts.

Our approach is to use a mixture of normals as the distribution of heterogeneity in a hierarchical Bayesian model. As with sufficient components in the mixture, we will be able to accommodate deviations from normality such as multi-modality, skewness, and fat tails. Let θ^h be the vector of choice model parameters for consumer h . The mixture of normals model specifies the distribution of θ^h across consumers as follows:

$$\begin{aligned}\theta^h &\sim N(\mu_{ind}, \Sigma_{ind}) \\ ind &\sim \text{multinomial}(\pi)\end{aligned}$$

π is a vector giving the mixture probabilities for each the K components. We complete the model specification with priors over the mixture probabilities and the mean and covariance matrices:

$$\begin{aligned}\pi &\sim \text{Dirichlet}(\alpha) \\ \mu_k | \Sigma_k &\sim N(\bar{\mu}, \Sigma_k \times a_{\mu}^{-1}) \\ \Sigma_k &\sim IW(\nu, V) \\ \{\mu_k, \Sigma_k\} &\text{ independent}\end{aligned}$$

We implement posterior inference for the mixture of normals model of heterogeneity and the multinomial logit base model along the lines of Rossi et al. (2005). We use a hybrid Metropolis method that uses customized Metropolis candidate densities

for each consumer. Conditional on the draws of θ^b , we use an unconstrained Gibbs sampler. Since our goal is to estimate the distribution of model parameters over consumers, we do not have to impose constraints on this Gibbs sampler to ensure identification. The density of model parameters is identified even if there is label switching.⁴ Moreover, it has been noted (Frühwirth-Schnatter 2001) that the unconstrained Gibbs sampler has superior mixing properties relative to Gibbs Samplers that are constrained in hopes of achieving identification of each component parameters.

Our MCMC algorithm will provide draws of the mixture probabilities as well as the normal component parameters. Thus, each MCMC draw of the mixture parameters provides a draw of the entire multivariate density of consumer parameters. We can average these densities to provide a Bayes estimate of the consumer parameter density. We can also construct Bayesian credibility regions for any given density ordinate to gauge the level of uncertainty in the estimation of the consumer distribution.

We fit models with successively larger numbers of components and assess the adequacy of the number components by examining the fitted density as well as the Bayes factor associated with each number of components. What is important to note is that our improved MCMC algorithm is capable of fitting models with a large number of components at relatively low computational cost.

Description of the Data

⁴ In mixture models, there is a generic identification problem which has been dubbed “label switching.” That is, the likelihood is unchanged if the labels for components are interchanged. This is only a problem if inference is desired for the mixture component parameters. In our application, we are interested in estimating

Switching costs are rarely directly observed (some components may be known, but the “hassle” costs of switching are not). For this reason, we must turn to data on the purchase histories of customers to infer switching costs from the observed patterns of switching between brands in the face of price variation. Consumer panel data on the purchases of packaged goods are ideal for estimating switching costs. The panel length is long relative to the average inter-purchase times and there is extensive price variation that causes frequent brand-switching and, hence, generates variation in consumers’ loyalty states.

For our empirical analysis, we estimate the logit demand model described above using household panel data containing all purchase behavior for the refrigerated orange juice and the 16 oz tub margarine categories. The panel data were collected by AC Nielsen for 2,100 households in a large Midwestern city between 1993 and 1995. In each category, we focus only on those households that purchase a brand at least twice during our sample period. Hence we use 355 households to estimate orange juice, and 429 households to estimate margarine demand. Table 3 lists the products considered in each category as well as the purchase incidence, product shares and average retail and wholesale prices.

Over 85 per cent of the trips to the store recorded in our panel data do not involve purchases in the product category. However, it is unlikely that each observed trip to the supermarket might potentially result in the purchase of either a pack of refrigerated juice and/or a tub of margarine. For a more realistic analysis, we define the outside good in each category as follows. In the refrigerated orange juice category, we define the outside

individual household parameters and the distribution of parameters across households. These quantities are identified even in the presence of label switching.

good as any fresh or canned juice product purchase other than the brands of orange juice considered. In the tub margarine category, we define the outside good as any spreadable product i.e. Jams, Jellies, margarine, butter, peanut butter etc).⁵ In Table 3, we see that under these definitions of the outside good, we see a no-purchase share of roughly 24%, in refrigerated juice, and 46% in tub margarine.

Demand Estimates

We now report the empirical estimates of demand from the orange juice and margarine data. In Table 5, we report the log-marginal density for several alternative model specifications and for each category. The posterior probability of a model specification is monotone in the log-marginal density, so that by choosing the model with the largest log marginal density we are picking the model with the highest posterior probability. It should be noted that the log-marginal density includes an automatic penalty for adding additional parameters (c.f. Rossi et al. 2005). By comparing models with and without switching costs and with varying degrees of heterogeneity, we can assess the importance of incorporating switching costs and non-normality. We assess the non-normality of the distribution of heterogeneity by comparing the log-marginal density for mixture models with varying numbers of components.

The results in Table 5 indicate several important features of the model. First, heterogeneity clearly leads to a substantial improvement in fit in both categories. Adding a switching cost term to the model also leads to an improvement in fit, albeit smaller.

⁵ In an earlier version of the paper, we defined the outside good as all store trips during which no product from the category of interest was purchased. This broad definition of the outside good led to estimates that were driven more by purchase incidence than brand choice. For instance, households who adopt a pattern of

These results confirm the well-established belief that consumer demand for frequently-purchased CPG products exhibits a switching cost that generates choice inertia even after controlling for persistent differences in consumer tastes.⁶

An interesting finding is the extent to which flexibility in the heterogeneity distribution may be required to “fit” the data. In both categories, we observe considerable improvement in fit by adding more components to the mixture. The improved fit from including five components in the mixture confirms the non-normality of the distribution of tastes in this category.

We now examine the model estimates to assess the non-normality of the fitted distributions of taste parameters. Ultimately, our goal is to estimate the distribution of tastes across households, not to attach any meaning or substantive significance to the parameters of the mixture components. Rather than report parameter estimates for the moments of each of the normal components, we instead plot the fitted marginal densities for several taste coefficients.

In Figures 2 and 3, we plot several fitted densities from the 1, 2 and 5 component mixture models for the margarine data. We also report the 95% posterior credibility region for the 5-component mixture model. This region provides point-wise evidence for the non-normality of the population marginal density for a given coefficient. Figure 2 provides compelling evidence of the need for a flexible model capable of addressing non-normality. In the upper panel, the Shedd’s brand intercepts from the 5-component model exhibit bimodality that cannot be captured by the 1 or 2 component models. The

purchasing a product on a regular cycle will be perceived as relatively price insensitive as the changes in price of the category relative to the outside good will have little influence on purchase incidence for these households.

⁶ Although not reported, our findings are robust to the inclusion of promotional variables such as weekly product features and in-aisle displays.

bimodality implies that there are households who differ markedly in their quality perceptions for margarines. In general, the results suggest that one would recover a very misleading description of the data-generating process if the usual symmetric normal (1-component) prior were used to fit these data.

In Figure 3, the price coefficient (upper panel) for the 5-component model also leads to a highly asymmetric and bi-modal density. In contrast, a symmetric 1-component model has both a mode and tails lying outside the credibility region for the 5-component model. For the switching cost estimates (lower panel), the 5 component model has a slightly lower mean and a thicker left tail than the 1-component model.

In Figure 4, we report fitted densities from the orange juice category. Just as in the margarine category, the densities of the brand intercepts for orange juice exhibit striking forms of non-normality.

Figures 5 and 6 display the fitted densities of the switching cost premium in dollar terms for each category. The inclusion of the outside option in the model enables us to assign money-metric values to our model parameters simply by re-scaling them by the price parameter (i.e. the marginal utility of income). For the switching cost parameter reported in the figures, this ratio represents the dollar cost foregone when a consumer switches to another brand than the one purchased previously. In the graphs, the point-estimate of switching costs from the homogeneous logit specification is denoted by a vertical red line.

Figures 5 and 6 display an entire distribution of switching costs across the population of households. Some of the values on which this distribution puts substantial mass are rather large values, others are small. To provide some sense of the magnitudes

of these values, we compute the ratio of the dollar switching cost to the average price of the products. The ratio of the mean dollar switching cost to average price is 0.19 for margarine and 0.15 for orange juice. It should be emphasized that the entire distribution of switching costs will be used in computation of equilibrium prices. The distribution of dollar switching costs puts mass on some very large values. For example, the ratio of the 95th percentile of dollar switching costs to average prices is 0.85 for margarine and 0.48 for orange juice. In the computations in Section 4 below, we will use this distribution of switching costs as the center point. We will also explore magnifying this distribution by scaling it by a factor of 4.

Pricing Implications of the Demand Estimates

In this section, we use the estimated demand systems to explore the implications of switching costs for pricing. For each of the categories, we compute the steady-state Markov perfect equilibrium prices corresponding to the demand estimates. We then examine the sensitivity of these steady-state price levels to specific parameter values.

To compute prices, we need to simplify the demand estimates to reduce the dimension of the state space of the model to a feasible range. For the orange juice data with 355 consumer “types” and 6 products, one would literally need to solve a dynamic programming problem with a $355 \cdot 5 = 1,775$ dimensional state space. We simplify the problem as follows. For the orange juice category, we focus only on 64 oz Tropicana and Minute Maid. We also take each household’s posterior mean taste vector and cluster them into 5 consumer “types.” Then our state space is $5 \cdot 1 = 5$ dimensional. Similarly,

in the margarine category we focus on all 4 products, and we cluster consumers into 2 “types.” This clustering reduces the state space to $2 \cdot 3 = 6$ dimensions.

Results from the clustering are reported in Table 6 for each of the categories. Recall that the flexible distribution of consumer tastes was critical during estimation to ensure we did not confound the empirical identification of switching costs with unobserved taste heterogeneity. While the current simplifications eliminate some of the richness of the true demand system, they should not detract from our main objective, which is to examine the pricing implications of the estimated switching costs.⁷

Only one segment displayed in Table 6 has a negative loyalty coefficient. Some have interpreted a negative coefficient on lagged choice as evidence of variety-seeking. However, this segment is small and will have little effect on our equilibrium pricing computations. This finding is consistent with some of the empirical literature on state dependence (c.f. Chintagunta (1999), Seetharaman, Ainslie and Chintagunta (1999)) that finds no evidence of variety seeking. Consideration of the implications of non-trivial segments of variety-seekers for equilibrium prices is a subject for future research.

We calculate steady-state prices as follows. First we compute the equilibrium pricing strategies of each firm. We then choose an arbitrary initial state for period $t=0$, and calculate the corresponding sequence of equilibrium price levels and state vectors for periods $t=1,2,\dots$. We stop this process once convergence of the state vector and corresponding equilibrium prices to fixed values occurs. Although we cannot prove uniqueness of the “steady-state”, we found convergence to a unique steady state for all our parameter values and all initial starting values. In the markets used to calibrate our

demand models, there exist temporary price changes or deals. Our goal is to understand the implications of switching costs for long-run or regular shelf pricing. We are not attempting to explain short-run variation in prices.

In Table 7, we report our results relating steady state price and profits levels to the magnitude of the switching costs. We compute equilibrium prices for a range of switching costs achieved by scaling the distribution of cluster parameters. That is, we multiply the switching cost parameter, γ , in each cluster by a scale factor reported in the left margin of Table 7. To isolate the impact of switching costs on inter-brand switching behavior (i.e. not on the outside good), we use the adjusted comparative static discussed previously and outlined in Appendix B. We see that prices decline as the switching cost increases from zero. At the estimated switching costs, prices fall by six per cent for Promise and ICBINB margarine. For orange juice, prices fall about three per cent at estimated switching cost levels. We compute equilibrium prices not only for the level of switching costs found in our data, but also for much higher levels corresponding to scale factors greater than one. We find that even with switching costs levels twice those revealed in our data, equilibrium prices are lower in the presence of switching costs. Only at scale factors of 3 do we begin to see a small number of the product prices rising again and, at a scale factor of 4, only one of the products' prices (Shedd's) returns roughly to the zero-switching costs level. Moreover, at a scale factor of 4, prices for the margarine brands Promise and ICBINB decline by over 15%.

Even more striking are the profit implications documented in Table 7. As we raise the switching costs from zero to the estimated levels (i.e. scale factor of 1), profits

⁷ In a companion paper (Dube et al (2006), we consider the monopoly problem and exploit the fact that Euler equations can be used to characterize the solution for the single agent problem. This allows us to use more

for most of the brands fall. In the case of Promise and Parkay, profits fall by more than 10%. At a scale factor of 4, only Shedd's experiences profit levels that exceed those of the zero-switching costs regime. The price and profit results generally indicate that, well within the range of switching costs levels we estimate empirically, switching costs intensify price competition.

Conclusions

We have demonstrated that equilibrium prices fall as switching costs increase for a realistic model. In some instances, prices fall by over 15% and profits by over 10%. This finding holds for a wide range of switching costs centered on those obtained from consumer panel data. Very high levels of switching costs must prevail in order to obtain results similar to those conjectured by Klemperer, i.e. that switching costs make markets less competitive and provide additional profits. Our switching cost estimates are based on consumer panel data for two categories of consumer products, margarine and orange juice. These switching costs are important from a statistical point of view in the sense that models with switching costs account for observed behavior better than those without. Our switching costs distribution puts mass on switching costs in the range of 15 to 60 per cent of purchase price. In addition, we have scaled this distribution up by a factor of four and still observe lower prices with switching costs. This means that our basic result applies to situations where switching costs are more than double the purchase price.

Our results can be reversed if switching costs reach very high levels or if, indeed, they are infinite as assumed in some of the theoretical literature on switching costs. In a

consumer types, something not possible if the solution to the dynamic game is desired.

world with the levels of switching costs envisaged by much of the theoretical literature, we would not see consumers switching brands very often. The empirical fact that consumers are observed to switch brands in many product categories supports the relevance of our result of declining prices.

In our model, the source of switching costs is psychological. It is well-known that the mere purchase/consumption of a product can create a form of inertia or brand loyalty which has psychological origins. Psychological switching costs are well recognized in the switching cost literature as important (c.f. the survey by Farrell and Klemperer (2006)). Moreover, psychological switching costs are present in any product category for which there are distinctly identified products (e.g. brands). This makes psychological switching costs more broadly applicable than a more narrow definition which is restricted to monetary switching fees or product adoption costs.

In our empirical work, we take a structural interpretation of the state dependence model as accommodating a form of switching costs. It is also possible that state dependence might arise as the result of other behavioral processes such as learning or consumer search. In a companion working paper (Dubé et al. 2008), we attempt to discriminate between these processes and show that there is evidence in favor of a switching cost interpretation. We emphasize that our categories of products include long-standing brands and households who have been active in the category for years. Under these conditions, it is hard to envision much additional learning taking place.

In our empirical model, consumers are not forward-looking. As consumers are unlikely to be consciously aware of the existence of psychological switching costs, the assumption of myopic behavior seems appropriate. In other markets, for example, the

market for operating systems there are non-trivial time costs of switching which consumers are aware of. For these markets, the assumption of myopic consumer behavior would not be appropriate. The impact of switching costs on a market with both forward-looking firms and consumers in any realistic empirical model would present a formidable computational challenge which we leave to future work.

References

- Allenby, G. and P. E. Rossi (1999), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics* 89, 57-78.
- Beggs, A. and P. Klemperer (1992), "Multi-Period Competition with Switching Costs," *Econometrica* 60, 651-666.
- Besanko, D., U. Doraszelski, Y. Kryukov, and M. Satterthwaite (2007), "Learning-by-Doing, Organizational Forgetting, and Industry Dynamics," manuscript, Harvard University.
- Cabral, L.M.B. and M. Villas-Boas (2005), "Bertrand Supertraps," *Management Science*, 51, 599-613.
- Cabral L.M.B. (2007), "Small Switching Costs Lead to Lower Prices," working paper, Stern School of Business, NYU.
- Che, H., K. Sudhir, P. B. Seetharaman (2006), "Pricing behavior in markets with state dependence in demand," *Journal Marketing Research*, forthcoming.
- Doganoglu, T. (2005), "Switching Costs, Experience Goods and Dynamic Price Competition," Working Paper, University of Munich.
- Doraszelski, U. and M. Satterthwaite (2005), "Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity," working paper, Harvard University.
- Dubé, J. P., G. Hitsch, P. E. Rossi, M. Vitorino (2006), "Category Management with State Dependent Utility," forthcoming *Marketing Science*.
- Dubé, J. P., G. Hitsch, P. E. Rossi (2008), "Identifying and Measuring Inertia in Consumer Choices," working paper, University of Chicago.
- Erdem, T. (1996), "A Dynamic Analysis of Market Structure Based on Panel Data." *Marketing Science*, 15, 359-378.
- Farrell, J. and P. Klemperer (2004), "Coordination and Lock-In: Competition with Switching Costs and Network Effects," forthcoming, *Handbook of Industrial Organization*.
- J. Farrell and C. Shapiro (1988), Dynamic competition with switching costs, *RAND Journal of Economics*, 19, 123-137.

- Fudenberg, Drew, and J. Miguel Villas-Boas (2006): "Behavior-Based Price Discrimination and Customer Recognition" *Handbook on Economics and Information Systems*, (T.J. Hendershott, Ed.), pp. 377-436.
- Heckman, J. J. (1981), "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating A Discrete Time-Discrete Data Stochastic Process and Some Monte Carlo Evidence" in C. Manski and D. McFadden (eds.) *Structural Analysis of Discrete Data With Econometric Applications*, Cambridge: MIT Press.
- Keane, M.P. (1997), "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *Journal of Business and Economic Statistics*, 15, 310-27.
- Klemperer, P. (1987), "Markets with Consumer Switching Costs," *Quarterly Journal of Economics*, 102, 375-394.
- Klemperer, Paul (1995), "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies*, 62 (4), 515-539
- Padilla, A. J. (1992), "Revisiting Dynamic Duopoly with Consumer Switching Costs," *JET*, 67, 520-530.
- Rossi, P. E., G. M. Allenby and R. McCulloch (2005), *Bayesian Statistics and Marketing*, New York: John Wiley and Sons.
- Roy, R., P. K. Chintagunta, and S. Haldar (1996), "A Framework for Investigating Habits, 'The Hand of the Past,' and Heterogeneity in Dynamic Brand Choice," *Marketing Science*, 15 (3), 280-299.
- Rust, J. (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55 (5), 999-1033.
- Seetharaman, P.B., A. Ainslie and P.K. Chintagunta (1999), "Investigating Household State Dependence Effects Across Categories," *Journal of Marketing Research*, 36, 4, 488-500.
- Seetharaman, P. B. (2004), "Modeling Multiple Sources of State Dependence in Random Utility Models: a distributed lag approach," *Marketing Science* 23, 263-271.
- Viard, B. (2003), "Do Switching Costs Make Markets More or Less Competitive?: the case of 800-number Portability," working paper, Stanford University.

Whitt, W. (1980), "Representation and Approximation of Non-cooperative Sequential Games," *SIAM Journal of Control and Optimization* 18, 33-48.

Appendix A.1 Existence of Equilibrium in the Simple Model

The existence of a Markov perfect equilibrium in our model follows from arguments given in Whitt (1980) and Doraszelski and Satterthwaite (2005). In order to show that the equilibrium is in pure strategies, we need to show that the best-reply correspondence is single-valued. Our strategy is to show that in the case of one consumer with logit demand, the right-hand side of the Bellman equation is strictly quasi-concave, and hence has a unique maximizer. This strategy has been employed previously by Besanko et al. (2007).

The Bellman equation in the simple model is given by

$$V_j(s) = \max_{p_j \geq 0} \left\{ \pi_j(s, p) + \beta \left(\sum_{k=1}^J P_k(s, p) V_j(k) + P_0(s, p) V_j(s) \right) \right\} \quad \forall s \in X.$$

Denote the right-hand side of this functional equation by $\Psi_j(s, p_j, p_{-j})$, such that

$$V_j(s) = \max_{p_j \geq 0} \Psi_j(s, p_j, p_{-j}), \quad \forall s \in X.$$

This maximization problem has the following first-order condition:

$$\begin{aligned} \frac{\partial \Psi_j}{\partial p_j} &= \alpha P_j (1 - P_j) (p_j - c_j) + P_j + \beta \left(\sum_{k=1}^J (-\alpha) P_j P_k V_j(k) + \alpha P_j V_j(j) - \alpha P_0 P_j V_j(s) \right) \\ &= \alpha P_j \left(-\Psi_j + (p_j - c_j) + \frac{1}{\alpha} + \beta V_j(j) \right). \end{aligned}$$

Here, P_j is shorthand for $P_j(s, p_j, p_{-j})$. Evaluating the second-order condition at a price where $\partial \Psi_j / \partial p_j = 0$, we find that

$$\begin{aligned} \frac{\partial^2 \Psi_j}{\partial p_j^2} &= \alpha^2 P_j (1 - P_j) \left(-\Psi_j + (p_j - c_j) + \frac{1}{\alpha} + \beta V_j(j) \right) + \alpha P_j \left(-\frac{\partial \Psi_j}{\partial p_j} + 1 \right) \\ &= \alpha (1 - P_j) \frac{\partial \Psi_j}{\partial p_j} + \alpha P_j \left(-\frac{\partial \Psi_j}{\partial p_j} + 1 \right) \\ &= \alpha P_j < 0. \end{aligned}$$

Hence, Ψ_j is strictly quasi-concave in p_j , and hence it follows that there is a unique price that maximizes the right-hand side of the Bellman equation for any state s and price profile $p_{-j} = \sigma_{-j}(s)$.

Appendix A.2 Simplified Model without Logit Error

If we remove the random utility component in (1), ε_{jt} , the equilibrium can be characterized analytically. We focus on the case of symmetry across players, where all firms have the same utility intercepts (i.e. no vertical differentiation) and costs. Removing ε_{jt} is analogous to eliminating the (horizontal) product differentiation and, hence, we are left with a model of homogeneous products. We assume that the common utility intercept, $\delta > c \geq 0$.

Proposition. *Let v be such that $0 \leq v \leq (1 - \beta)\gamma$ and $c + v \leq \delta + \gamma$. Then under the assumptions stated above there is a symmetric Markov perfect equilibrium with pricing strategies $\sigma_j^*(j) = c + v$ and $\sigma_j^*(k) = c + v - \gamma$ for all $k \in X$, $k \neq j$.*

Proof. j denotes the product to which the customer is loyal and k denotes any other product. Because $p_j = c + v = p_k + \gamma$, the customer's utility index is the same for all products. Therefore, by assumption she will not switch from product j to k , and because $0 \leq \delta + \gamma - (c + v)$, she will not choose the outside option. The value from this strategy is $V_j(j) = (1 - \beta)^{-1}v$ and $V_j(k) = 0$. In order to assess whether the proposed strategies constitute a best response for each player, we only need to consider one-period deviations. If firm j reduces its price, it will reduce its current-period profit and leave its future value unchanged. If firm j raises its price, it will lose its loyal customer and receive a payoff of zero now and in future. Hence, $p_j = c + v$ is a best response to p_k . Competitor k needs to offer a price $p_k = c + v - \gamma - \varepsilon$, $\varepsilon > 0$, in order to acquire the

customer. Because $v \leq (1 - \beta)\gamma$, the present value from this one-period deviation is negative:

$$v - \gamma - \varepsilon + \beta \frac{v}{1 - \beta} = \frac{v}{1 - \beta} - \gamma - \varepsilon < 0.$$

Alternatively, firm k cannot improve on its current outcome by raising its price, and hence, $p_k = c + v - \gamma$ is a best response to p_j .

Appendix B. Comparative Statics Holding the Market Size Constant

In order to eliminate the difference between the state dependence and the pure switching cost models, which is due to the differential impact on the outside good market share under changes in γ , we define the outside good intercept as a function of the switching cost. If $\gamma = 0$, we choose some arbitrary intercept, such as $\delta_0 = 0$. We denote the resulting equilibrium prices in state s by $p^0(s)$, and let the corresponding outside good share be P_0^0 . For $\gamma > 0$, choose δ_0 such that $P_0(s, p^0(s); \delta_0) = P_0^0$, and note that due to symmetry, the left-hand side of this equation and hence the choice of δ_0 is the same for either state, $s = 1, 2$. Under this choice of δ_0 , if the firms do not change their prices compared to the case without switching costs, the outside good market share at any $\gamma > 0$ will remain constant at P_0^0 , its level under $\gamma = 0$. However, even though the total market size does not change, the customer is more likely to purchase the good to which she is loyal to for larger values of γ . Note that this technique of defining the outside good intercept as a function of the switching cost eliminates any difference between the state-dependent and the pure switching cost model. In particular, at any given price vector and δ_0 as defined above, the predicted demand from either model formulation is identical.

Appendix C. Forward-Looking Consumers in the Simple Model

We now extend the model to allow for forward-looking consumers who anticipate the consequences of becoming loyal to product j . In general, the presence of forward-looking consumers can complicate the computation of an equilibrium. For example, Anderson, Kumar and Rajiv (2004) show the equilibrium proposed by Padilla (1995) does not in fact constitute a Markov perfect equilibrium under forward-looking consumer behavior.

As before, the current-period utility from choosing product j is $U_j = U(j, s, p) + \lambda \varepsilon_j$. But, now consumers maximize the PDV of current and future utilities. For simplicity, we assume that consumers discount future utilities at the same rate as firms, β . Define the state transition function $s' = \phi(s, j) = j$ if $j \neq 0$ and $s' = \phi(s, 0) = s$. The value function of the consumer given state s and idiosyncratic utility draws $\varepsilon = (\varepsilon_0, \dots, \varepsilon_j)$ is

$$v(s, \varepsilon) = \max_{j=0, \dots, J} \left\{ U(s, \sigma(s), j) + \lambda \varepsilon_j + \beta \int v(\phi(s, j), \varepsilon') f(\varepsilon') d\varepsilon' \right\}. \quad (\text{C.1})$$

Note that this value function depends on the consumer's expectation that the firms choose prices according to $p_j = \sigma_j(s)$. Following arguments given in Rust (1987), the consumer's decision problem can be reformulated in the following way. Let the expected future value from choosing alternative j in state s be

$$W(s, j) = \int \max_{k=0, \dots, J} \left\{ U(s', \sigma(s'), k) + \lambda \varepsilon_k + \beta W(s', k) \right\} f(\varepsilon) d\varepsilon,$$

where $s' = \phi(s, j)$. Since ε has the Type I extreme value distribution, $W(s, j)$ has the closed form expression

$$W(s, j) = \lambda \left(\gamma + \log \left[\sum_{k=0}^J \exp \left(\frac{1}{\lambda} (U(s', \sigma(s'), k) + \beta W(s', k)) \right) \right] \right). \quad (\text{C.2})$$

Here, $\gamma \approx 0.57722$ is Euler's constant. The consumer then chooses the alternative $j = 0, \dots, J$ that yields the highest utility index

$$U(s, \sigma(s), j) + \beta W(s, \phi(s, j)) + \lambda \varepsilon_j.$$

Conditional on the consumer's choice behavior, which is now also described by the consumer's value function, W , the firm's problem remains the same under forward-looking consumer behavior. A Markov perfect equilibrium now consists of pricing strategies and value functions for each firm j and the consumer's consumption strategy, which is fully described by the value function W , such that (i) each firm's pricing strategy is optimal given the consumer's strategy and given the competitors' strategies, and (ii) given the firms' pricing strategies, the consumer's value function satisfies equation (C.2).

In Section 2, we explored the predictions of the simple model for the symmetric case with a symmetric equilibrium. In this case, myopic and forward-looking consumer behavior is identical. This can be seen from equation (C.2): W actually depends only on $s' = \phi(s, j)$, the product that the consumer is loyal to in the next period. Due to symmetry, the identity of this product does not matter. Therefore, W is exactly the same for all $s' \in X$, and therefore adds the same constant to each utility index. Thus, the choice probabilities are not affected by the presence of W . Therefore, our pricing results are robust to a forward-looking consumer.

Table 1

Equilibrium Prices under Different Switching Cost Levels

Switching Cost	p_1	p_2	p^a	Purchase Prob. 1	Purchase Prob. 2	Prob. Stay Loyal
0.00	1.808	1.808	1.808	0.236	0.236	0.764
0.25	1.802	1.658	1.734	0.258	0.232	0.768
0.50	1.794	1.500	1.662	0.279	0.227	0.773
0.75	1.784	1.335	1.593	0.298	0.221	0.779
1.00	1.773	1.165	1.528	0.317	0.214	0.786
1.25	1.762	0.991	1.467	0.334	0.207	0.793
1.50	1.750	0.813	1.410	0.349	0.199	0.801
1.75	1.738	0.631	1.356	0.363	0.191	0.809
2.00	1.727	0.445	1.307	0.376	0.183	0.817
3.00	1.732	0.000	1.352	0.421	0.119	0.881
4.00	1.782	0.000	1.607	0.450	0.049	0.951
5.00	1.812	0.000	1.740	0.460	0.019	0.981
6.00	1.844	0.000	1.816	0.458	0.007	0.993
7.00	1.896	0.000	1.885	0.448	0.003	0.997
8.00	1.972	0.000	1.967	0.430	0.001	0.999

For positive switching cost levels, δ_i is adjusted such the total market size remains constant; see Appendix B for the details. The results were calculated for product intercepts = 1.0 and a price coefficient = 1.0, and outside good intercept at 0. The discount factor is $\beta = 0.998$.

Table 2
Firm Profits under Different Switching Cost Levels

Switching Cost	p^a	V_1	V_2	V_1^0	V_2^0
0.00	1.808	154.1	154.1	154.1	154.1
0.25	1.734	151.0	150.8	154.2	154.0
0.50	1.662	146.9	146.6	154.3	153.9
0.75	1.593	142.1	141.6	154.5	153.8
1.00	1.528	136.7	136.1	154.7	153.6
1.25	1.467	131.0	130.2	154.9	153.3
1.50	1.410	125.1	124.2	155.3	153.0
1.75	1.356	119.2	118.1	155.7	152.6
2.00	1.307	113.4	112.1	156.2	152.1
3.00	1.352	116.2	113.8	160.1	148.1
4.00	1.607	141.2	135.1	169.8	138.4
5.00	1.740	156.1	140.8	190.8	117.5
6.00	1.816	172.1	134.2	225.0	83.2
7.00	1.885	198.2	113.8	261.8	46.4
8.00	1.967	236.1	80.3	287.2	21.1

For positive switching cost levels, δ_i is adjusted such the total market size remains constant; see Appendix B for the details. The results were calculated for product intercepts = 1.0 and a price coefficient = 1.0. V_j denotes the value of firm j in state 1, while V_j^0 denotes the value of firm 1 in state s if prices remain constant at their no-switching-cost-level, $p_1 = p_2 = 1.808$.

Table 3
Description of Data

Refrigerated Orange Juice

Product	Retail Price	Wholesale Price	% trips
64 oz MM	2.21	1.36	11.1
premium 64oz MM	2.62	1.88	7.00
96 oz MM	3.41	2.12	14.7
premium 64oz TR	2.73	2.07	28.8
64 oz TR	2.26	1.29	6.76
premium 96 oz TR	4.27	2.73	7.99
no-purchase (% trips)	23.75		
# households	355		
# trips per household	12.3		
# purchases per household	9.37		

Margarine

Product	Retail Price	Wholesale Price	% trips
Promise	1.69	1.22	13.11
Parkay	1.63	1.02	4.98
Shedd's	1.07	0.83	12.66
ICBINB	1.55	1.11	23.51
no-purchase (% trips)	45.73		
# households	429		
# trips per household	18.25		
# purchases per household	9.90		

Table 4
(a) Purchase versus re-purchase rates

Category	Refrigerated Orange Juice		Tub Margarine				
	Brand	Minute Maid	Tropicana	Promise	Parkay	Shedd's	ICBINB
Sample purchase frequencies		0.429	0.570	0.241	0.091	0.233	0.433
Sample re-purchase frequencies		0.777	0.856	0.827	0.90	0.233	0.884

(b) Purchase versus re-purchase rates and price discounts

Category	Refrigerated Orange Juice		
	Brand	Minute Maid	Tropicana
Sample re-purchase frequencies after discount price		0.741	0.833

Table 5
Fit and the Role of Heterogeneity and State-dependence

		Log Marginal Density	
Model	# Components	Margarine	Orange Juice
w/o Switching Costs	No heterogeneity	-10750.93	-7612.014
	5-component	-5529.788	-4471.425
with Switching Costs	No heterogeneity	-8174.184	-6296.471
	1-component	-5548.706	-4473.165
	2-component	-5539.313	-4460.562
	5-component	-5469.324	-4425.97
	10-component	-5434.482	-4371.794

Table 6
Clusters Used In Equilibrium Pricing Computations

Refrigerated Orange Juice

segment	64 oz MM	premium 64oz MM	96 oz MM	premium 64oz TR	64 oz TR	premium 96 oz TR	price	loyalty	loyalty (\$)	size
1	-2.88	-2.57	-2.50	-0.25	-2.59	-0.31	-1.19	0.69	0.59	0.26
2	-2.62	-3.79	-1.79	-2.88	-3.72	-3.59	-0.91	1.23	1.36	0.25
3	-13.09	-12.20	-9.54	-1.22	-9.53	-3.19	-0.31	-0.03	-0.10	0.02
4	-0.37	0.32	0.01	1.53	-0.43	1.73	-2.08	0.23	0.11	0.18
5	-1.30	-1.59	-0.50	-0.71	-1.92	-0.82	-1.65	0.61	0.37	0.29

16 oz Tub Margarine

segment	Promise	Parkay	Shedd's	ICBINB	price	loyalty	loyalty (\$)	size
1	-1.95	-3.47	-1.22	-2.67	-2.46	0.17	0.07	0.50
2	-2.88	-6.87	-6.49	-2.97	-0.87	0.19	0.22	0.50

Table 7
Equilibrium Prices and Profits

Steady State Prices

Scale Factor	Prices					
	16-oz Tub Margarine				Refrigerated Orange Juice	
	Promise	Parkay	Shedd's	ICBINB	Minute Maid	Tropicana
0	1.887	0.732	0.704	1.838	1.517	1.996
1	1.773	0.720	0.698	1.728	1.472	1.935
2	1.680	0.708	0.696	1.646	1.451	1.895
3	1.607	0.698	0.698	1.588	1.461	1.879
4	1.549	0.694	0.704	1.548	1.494	1.901

Steady State per Period Profits

Scale Factor	Profits					
	16-oz Tub Margarine				Refrigerated Orange Juice	
	Promise	Parkay	Shedd's	ICBINB	Minute Maid	Tropicana
0	40.92	4.09	29.26	56.37	51.39	254.70
1	36.31	3.27	31.59	53.18	46.28	254.30
2	32.11	2.71	33.50	49.79	44.18	252.90
3	28.54	2.40	35.02	46.38	45.01	252.50
4	25.59	2.31	36.34	43.04	48.77	252.90

Figure 1
Average Transaction Price vs. State Dependence Level

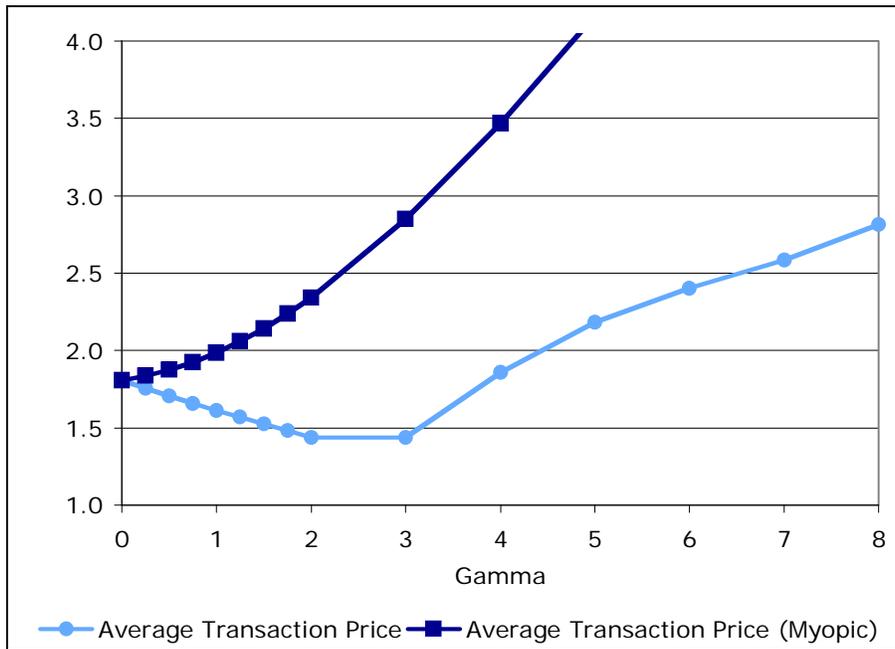


Figure 2
Fitted Densities for Shedd's and ICBINB Brand coefficients (Margarine)

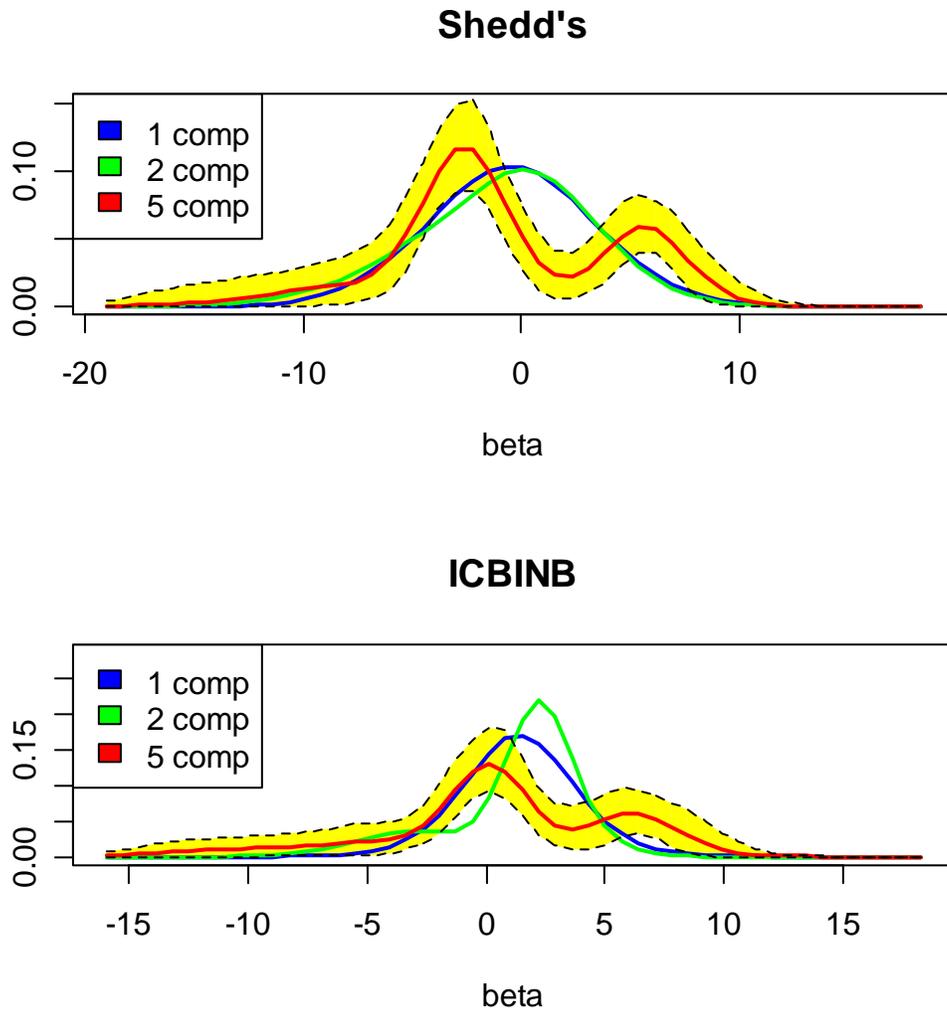


Figure 3
Fitted Densities for Price and State Dependence Coefficients (Margarine)

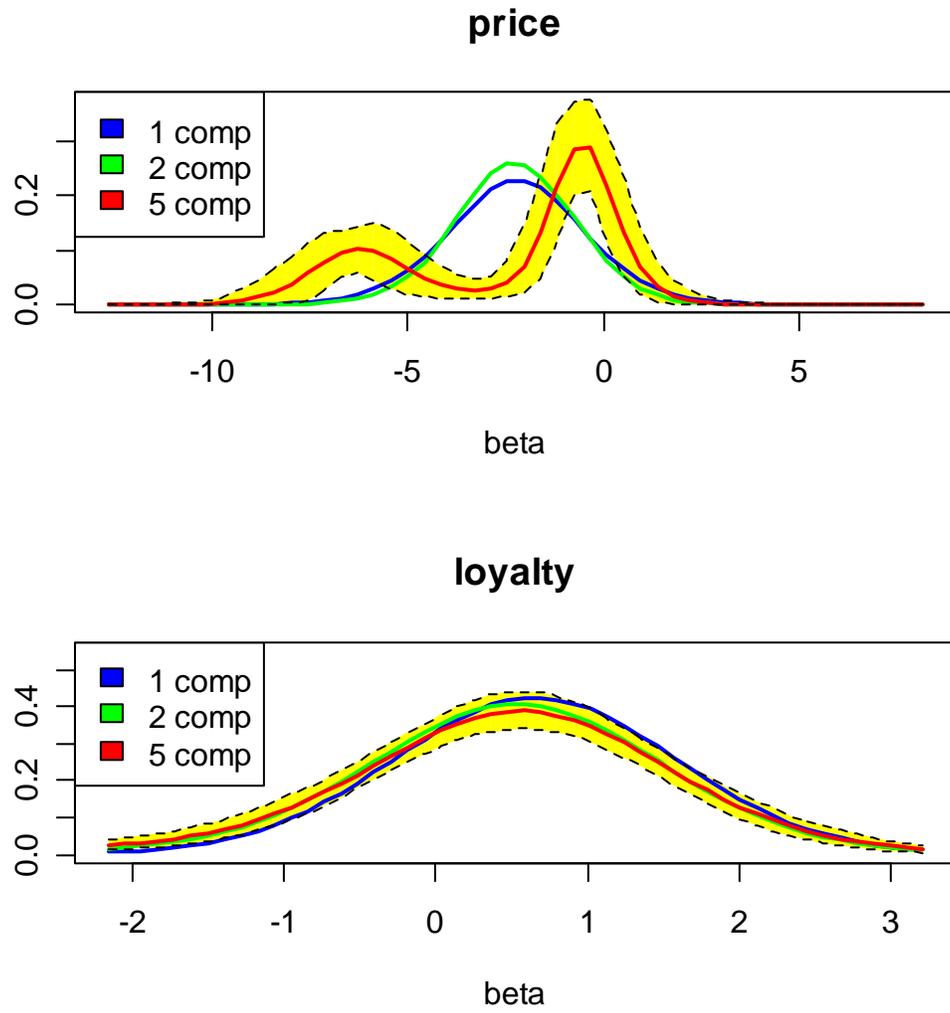


Figure 4
Fitted Densities for 96 oz Minute Maid and 64 oz Tropicana Brand coefficients
(Orange Juice)

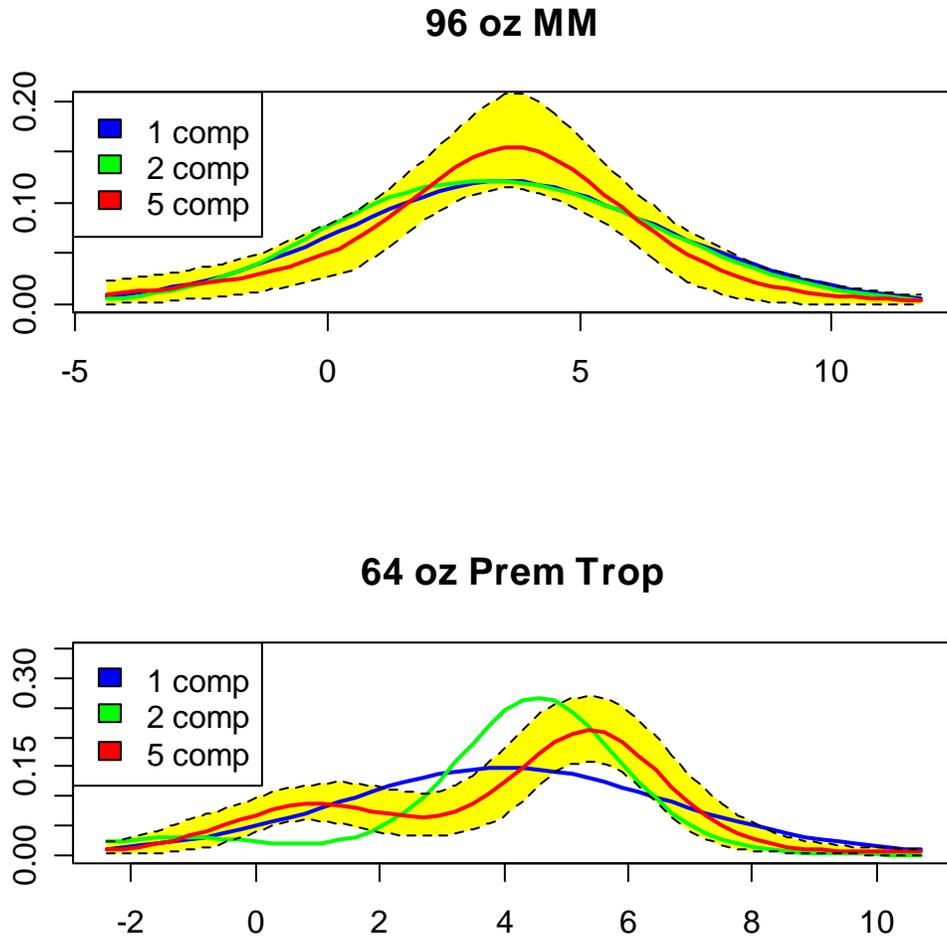


Figure 5
Fitted Densities and 95% Posterior Credibility Regions for the Money-metric State
Dependence Premium in dollars (Margarine)

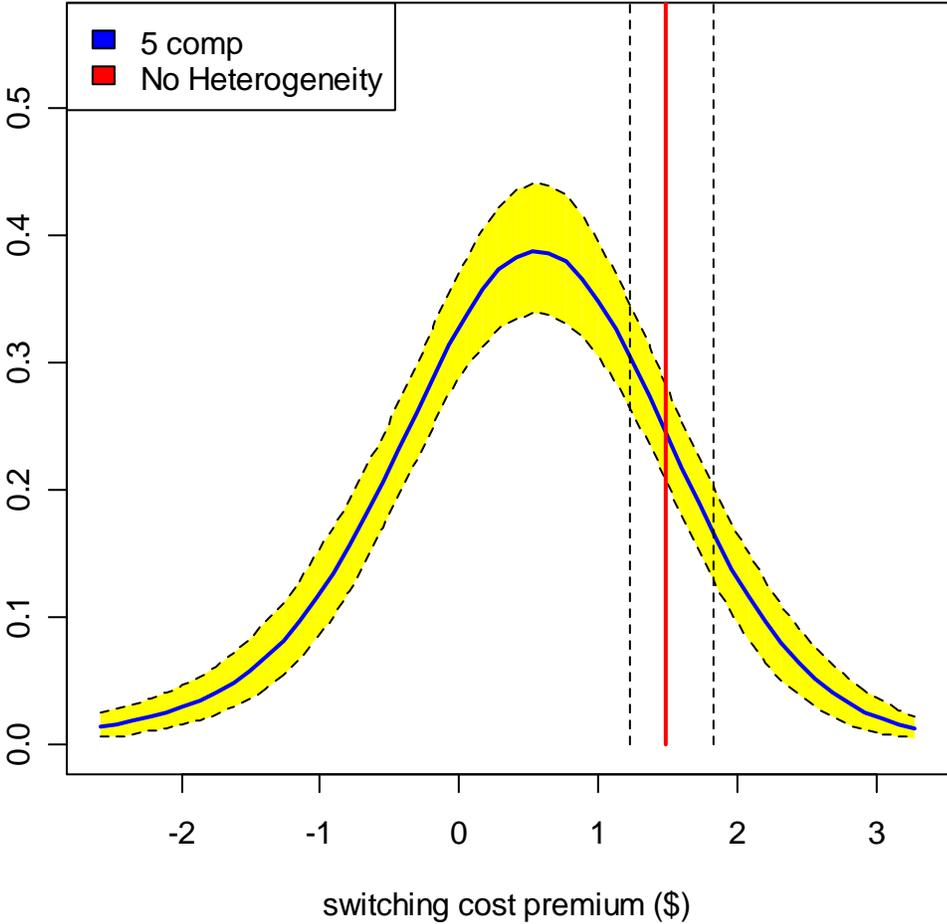


Figure 6
Fitted Densities and 95% Posterior Credibility Regions for the Money-metric State
Dependence Premium in dollars (Orange Juice)

