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 2006

Keywords: switching costs, dynamic model, brand loyalty, state dependence
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 find that steady-state equilibrium prices may fall as switching costs are introduced into a simple model of dynamic price competition that allows for differentiated products and imperfect lock-in. To assess whether this finding is of empirical relevance, we consider a more general model with heterogeneous consumers. We calibrate this model with data from a frequently purchased packaged goods market where consumers exhibit inertia in their brand choices, a behavior consistent with switching costs. We estimate the level of switching costs from the brand choice behavior in this data. At switching costs of the order of magnitude found in our data, prices are lower than without switching costs.

All correspondence may be addressed to the authors at the Graduate School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637; or via e-mail: jdube@gsb.uchicago.edu, guenter.hitsch@chicagogsb.edu, or peter.rossi@chicagogsb.edu.

We are grateful for comments and suggestions from Alan Sorenson, Matt Gentzkow, Ariel Pakes, Peter Reiss, Miguel Villas-Boas, and seminar participants at the University of Chicago structural IO lunch, Erasmus University in Rotterdam, Northwestern University, the University of Minnesota, Yale University, the 2006 NBER summer I.O. meetings, and the 2006 SED conference. Support from the Kilts Center for Marketing, Graduate School of Business, University of Chicago is gratefully acknowledged.

1. Introduction

A large theoretical literature has studied the impact of consumer switching costs on price competition (c.f. survey by Farrell and Klemperer 2005). In general, markets with switching costs exhibit two forces that have opposite effects on equilibrium prices. First, firms have an incentive to invest in their market share, which induces them to lower their prices. Second, firms want to "harvest" their base of customers for whom switching is costly by raising prices. An important branch of the extant literature has found that, in infinite-horizon games, switching costs typically make markets less competitive in the sense that prices are higher in equilibrium with switching costs than without (Farrell and Shapiro 1988, Beggs and Klemperer 1992, Padilla 1995, and Anderson et al. 2004). Accordingly, the harvesting motive is found to overshadow the investment motive in long-run equilibrium. Klemperer (1995) conjectures that this result is likely to hold in general whenever firms are unable to price discriminate between existing and new customers.[1] A small body of empirical literature has documented a positive correlation between switching costs and prices in specific case studies (Stango 2002, Viard 2005, and Shi et al. 2006). These findings have led to the conventional wisdom that switching costs soften price competition.

In this paper we challenge this view. We analyze a class of switching cost models that can be taken directly to data and use numerical techniques to compute equilibrium prices. For realistic parameter values, we find that equilibrium prices are lower in the presence of switching costs. Although the basic intuition of competition with switching costs can be illustrated in two-period models (Klemperer 1987a and 1987b), we focus on infinite-horizon games, which more closely correspond to real-world markets in which

[1] See Section 4, entitled "Do Switching Costs Make Markets Less Competitive?" in Klemperer (1995).

trading does not end at some known point in time. Our model departs in empirically relevant aspects from the previous literature that has analyzed infinite-horizon games. First, in contrast to Farrell and Shapiro (1988) and Padilla (1995), we allow for differentiated products. Second, we allow for imperfect lock-in, such that consumers can switch in equilibrium. 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. For example, consumers switch between cell phone plans and even computer operating systems. The previous literature (Beggs and Klemperer 1992) assumes that there is perfect lock-in, or at least that switching costs are so large that consumers never switch among products.

We find that steady-state prices often fall under a wide range of plausible switching costs levels. Hence, switching costs can make markets more competitive. We first illustrate this effect in a simple model that allows us to explore the economic forces that drive this result. The main source of our finding is imperfect lock-in, introduced into our model through a random component in consumer's utility and the assumption of finite switching costs.[2] Imperfect lock-in stimulates the investment motive. If customers can be induced to switch between products, firms have an incentive to compete for one-another's loyal customers and, at the same time, to defend their own loyal customer base. This dimension of competition puts downward pressure on prices. In our model, we find that as switching costs increase from zero, the investment motive overshadows the harvesting motive and leads to lower equilibrium prices. However, for sufficiently high switching costs, the incentive to harvest the base of loyal customers overwhelms the investment motive, and equilibrium prices begin to rise.

Our theoretical results imply that whether switching costs soften or toughen price competition is an empirical question about the magnitude of switching costs consistent with actual consumer behavior. Switching costs are typically not directly observed. Instead, the analyst must infer their magnitude from the observed switching behavior of consumers. Our approach is to use panel data with a reasonably long time dimension and considerable price variation coupled with a semi-parametric model of consumer heterogeneity. This allows us to separate heterogeneity from state dependence in demand and obtain reasonably precise estimates of the distribution of switching costs across consumers. 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 equilibrium.

Orange juice is an example of a branded, frequently purchased product. A large literature in marketing has shown that consumer choices for such products exhibit inertia, often termed "loyalty" or "state dependence". Empirical models used to capture this purchase inertia (e.g. Erdem 1996, Keane 1997 and Shum 2004) are very similar to the theoretical models of switching costs. Since both switching costs models and loyalty models give rise to inertia in consumer choices, Klemperer (1995) and Farrell and Klemperer (2005) refer to "loyalty" as an example of switching costs. Brand loyalty may have "psychological" sources or it may be the rational result of "shopping costs," where consumers do not re-optimize the set of products considered and bought at each shopping occasion. In this paper, we are not concerned with the exact source of brand loyalty, especially given the long precedent for using such models in the context of package goods demand. Rather, we focus on whether it in fact exists and can be identified from observed purchasing behavior.

---

[2] Viard (2003) and Doganolu (2005) obtain similar results for specialized model formulations which cannot be readily extended to form the basis for an empirical specification.

Our estimated switching costs are on the order of 15 to 60 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 several times those inferred from our data, we still find that prices decline in the presence of switching costs.

2. A Simple Model

In this section, we develop a model of price competition in markets with switching costs. For this simplified model, we can establish the existence of a Markov perfect equilibrium. The model also allows us to explore the economic forces that determine equilibrium prices and the robustness of these forces to several model extensions. 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 a model where both of these features are present, we find that equilibrium prices can be lower in the presence of switching costs. This result arises from the inherent dynamics of the game whereby firms use current prices to invest in future locked-in customers. In section 3, we will extend the model for a much richer empirical application with many heterogeneous consumers.

Model Details

We consider a market with J competing firms. Each firm sells one product. Time is discrete, $t = 0,1,\ldots$. There is exactly one consumer in the market, who chooses among the J products and an outside option in each period.

In each period, the consumer is loyal to one product, $j = s_t$. The loyalty variable $s_t \in X = \{1,...,J\}$ summarizes all current-period payoff-relevant information, and describes the state of the market. Demand is derived from a discrete choice model. Conditional on price p_{jt} and her current loyalty state s_t, the consumer's utility index from the choice of j is

$$U_{jt} = \delta_j + \alpha p_{jt} + \gamma I\{s_t = j\} + \lambda \varepsilon_{jt}. \tag{1}$$

As is common in much of the empirical literature on demand estimation, we assume that the random utility component ε_{jt} is i.i.d. Type I Extreme Value distributed. λ determines the scale of the utility shock, and thus the degree of horizontal product differentiation between the products. In the limiting case of $\lambda = 0$, product differentiation is purely vertical. 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. This specification is typically termed the "state dependent" demand model.

Let $U(j, s_t, p_t)$ denote the deterministic component of the utility index, such that $U_{jt} = U(j, s_t, p_t) + \lambda \varepsilon_{jt}$. The utility from the outside alternative is $U_{0t} = U(0, s_t, p_t) =$ $= \delta_0 + \varepsilon_{0t}$. If $\lambda > 0$, demand is given by the logit choice probabilities

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

If there is no horizontal product differentiation ($\lambda = 0$), the consumer buys the product with the highest utility index. If there is more than one product that maximizes utility, the consumer chooses the product to which she is loyal if it is among the utility-maximizing options, and randomizes among the utility-maximizing products otherwise.

The current loyalty state of the consumer, s_t, evolves as follows. If the consumer buys product $k \neq j$ in period *t*, then $s_{t+1} = k$, i.e. she becomes loyal to *k*. If the consumer buys product $j = s_t$ or chooses the outside option, then $s_{t+1} = s_t$, i.e. her loyalty remains unchanged. The firms cannot observe the random utility component. Hence, conditional on a product price vector, p_t, the state variable follows a Markov process from the firm's point of view:

$$\Pr\{s_{t+1} = j \mid s_t, p_t\} = \begin{cases} P_j(s_t, p_t) + P_0(s_t, p_t) & \text{if } j = s_t, \\ P_j(s_t, p_t) & \text{if } j \neq s_t. \end{cases} \tag{2}$$

Below, we discuss the extension to forward-looking consumer behavior.

Conditional on all product prices and the state of the market, firm *j* receives the expected current-period profit $\pi_j(s_t, p_t) = P_j(s_t, p_t) \cdot (p_{jt} - c_j)$. c_j is the marginal cost of production, which does not vary over time. Firms compete in prices, and choose Markovian pricing strategies that depend on the current payoff-relevant information, summarized by s_t. This assumption rules out behavior that conditions current prices also on the history of past play, and thus collusive strategies in particular. We denote firm *j*'s strategy by $\sigma_j : X \rightarrow \mathbb{R}$. Firms discount the future using the common factor β, $0 \leq \beta < 1$. For a given profile of strategies, $\sigma = (\sigma_1, ..., \sigma_J)$, the expected PDV of profits, $\sum_{t=0}^{\infty} \beta^t \pi_j(s_t, \sigma(s_t))$, 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

$$V_j(s) = \max_{p_j \geq 0} \left\{ \pi_j(s, p) + \beta \left(\sum_{k=1}^{J} P_k(s, p) V_j(k) + P_0(s, p) V_j(s) \right) \right\} \qquad \forall s \in X. \tag{3}$$

In this equation, the price vector consists of firm *j*'s price and the prices prescribed by the competitor's strategies, $p = (\sigma_1(s),...,\sigma_{j-1}(s), p_j, \sigma_{j+1}(s),...,\sigma_J(s))$. Therefore, the Bellman equation (4) depends on the pricing strategies chosen by the competitors. Note that the expectation of the firm's future value is taken with respect to the transition probabilities of s_t, which are directly related to the current choice probabilities.

We use Markov perfection as our solution concept. For this pricing game, there always exists a Markov perfect equilibrium in pure strategies. The proof for the case of horizontal product differentiation, $\lambda > 0$, relies on the quasi-concavity of a logit-based objective function, which is well known for the static case and also holds for the case of dynamic competition considered here. Quasi-concavity ensures that each player has a unique best response. The proof is presented in Appendix A. While we can show the existence of a pure-strategy equilibrium, we cannot characterize the equilibrium policies analytically. Instead, we solve the game numerically for different parameter values.

In the case of no horizontal product differentiation, $\lambda = 0$, the equilibrium can be characterized analytically. We focus on the case of symmetry across players, where all firms have the same utility intercepts and costs. We assume that $\delta > c \geq 0$.

Proposition. *Let* υ *be such that* $0 \leq \upsilon \leq (1-\beta)\gamma$ *and* $c + \upsilon \leq \delta + \gamma$. *Then under the assumptions stated above there is a symmetric Markov perfect equilibrium with pricing strategies* $\sigma_j^*(j) = c + \upsilon$ *and* $\sigma_j^*(k) = c + \upsilon - \gamma$ *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 $p_j = c + \upsilon = p_k + \gamma$, 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 $0 \le \delta + \gamma - (c + \upsilon)$, she will not choose the outside option. The value from this strategy is $V_j(j) = (1-\beta)^{-1}\upsilon$ 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 loose its loyal customer and receive a payoff of zero now and in future. Hence, $p_j = c + \upsilon$ is a best response to p_k. Competitor k needs to offer a price $p_k = c + \upsilon - \gamma - \varepsilon,\ \varepsilon > 0$, in order to acquire the customer. Because $\upsilon \le (1-\beta)\gamma$, the present value from this one-period deviation is negative:

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

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

Equilibrium Price Computations

We now explore the predictions of the pricing model developed above. 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 $\sigma_1^*(2) = \sigma_1^*(1)$. We therefore only need to know firm 1's pricing policy to characterize the market equilibrium.

We first consider the case of homogenous products $(\lambda = 0)$. The proposition above states 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 $\upsilon = (1-\beta)\gamma$. υ 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 $(\lambda > 0)$. In the case of homogenous products, the customer never switches in equilibrium, and hence the realized transaction price is the price that the customer pays for the product to which she is loyal. 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:

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

That is, 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 $s_t = 2$.

Figure 1 shows the relationship between the level of switching costs and the average transaction price for the case of $\delta_j = 1$, $c_j = 0.5$, $\alpha = 1$, and $\lambda = 1$. 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. For a switching cost level of $\gamma = 3$, despite the fact that the probability of staying loyal is 0.77, the average prices are lower than without switching costs[3]. Table 1 displays the average transaction price and the individual prices set by firms 1 and 2. The table also shows the purchase probabilities for each product and the probability that the customer stays loyal. 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 (i.e. the variance of the random component of utility). 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 simple, yet nonetheless the foundation 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. 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

[3] This pattern of declining-then-rising equilibrium prices is robust to other parameter values. We have not been able to find an example where average transaction price always—even for small switching cost levels—

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 our main comparative static result 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). Although not reported, the latter effect makes both firms better off. Since our main interest lies in the role of the switching cost, we construct an additional comparative static exercise in which γ does not influence the outside share and leaves the total market size constant. We provide technical details for this exercise in Appendix B. Table 2 indicates that we still observe the main result of decreasing-then-increasing prices. We also report results from a scenario in which we also hold the prices fixed at the levels obtained under zero switching costs $(\gamma = 0)$. 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. This outcome is an instance of a "Bertrand supertrap," as analyzed in Cabral and Villas-Boas (2005) for finite-horizon games. Cabral and Villas-Boas

---

increases.

show that Bertrand supertraps can arise under demand complementarities, which include but are not limited to switching costs.

Some Variations on the Simple Model

We now briefly discuss three variants of the simple model to check the robustness of our results to other model features typically considered in the theoretical literature. In particular, we will separately consider the switching costs specification commonly used in the applied theory literature, as well as the impact of forward-looking consumer behavior and overlapping generations of consumers on equilibrium prices. We discuss results briefly below, with a more complete analysis of the latter two cases in the Appendices C and D.

Switching cost model. The applied theory literature typically uses a slightly different specification of the utility index which, accordingly, we term "the pure switching costs model":

$$U_{jt} = \delta_j + \alpha p_{jt} - \gamma I\{s_t \neq j\} + \lambda \varepsilon_{jt}. \tag{4}$$

If there is no outside alternative, the "state-dependent demand model" considered above, (1), is equivalent to the switching cost formulation in (2). Thus, our results for the cases $\delta_0 = -\infty$ and for the restricted outside share, above, are robust to this specification. With an outside alternative, there are some subtle differences due to the exact way that the switching cost enters the utility index: in model (1), the customer "gains" additional utility from the product to which she is loyal, while in (2) she pays a monetary or utility cost when she switches to another product. The empirical brand choice literature has routinely used model (1) to capture observed inertia in household brand choices, c.f. Erdem (1996), Roy, Chintagunta, and Haldar (1996), Keane (1997), Seetharaman, Ainslie, and Chintagunta

(1999), Seetharaman (2004) and Shum (2004). However, both specifications generate inertia in consumer choices over time and, accordingly, theorists consider the state-dependent model as an example of switching costs (c.f. Klemperer 1995 and Farrell and Klemperer 2006). In Figure 2, we show that our main finding of decreasing-then-increasing prices is occurs in the pure switching costs formulation as well as in the state dependent model.

Forward-looking Consumers. The demand side of the model (1) can be extended to allow for forward-looking consumer behavior, where the customer anticipates the consequences of becoming loyal to a particular product. Demand is then similar to Rust (1987); the technical details can be found in Appendix C. In the particular case of symmetric competition discussed above, demand and hence equilibrium prices are the same under both myopic and forward-looking consumer behavior. Due to symmetry, the customer's current and future payoffs are identical regardless of the identity of the product to which she is loyal. If she is currently loyal to product 1, for example, she faces the identical choice situation next period regardless of whether she switches to product 2 today or remains loyal to product 1. This argument depends on the assumption that the customer is always loyal to one of the products, and not to the outside option. More importantly, our pricing results are robust to a forward-looking consumer.

Overlapping Generations. In the theory literature on switching costs, competition in an infinite horizon setting (Beggs and Klemperer 1992 and Padilla 1995, for example) is examined using overlapping generations (OLG) models. In Appendix D, we develop an OLG version of our simple state dependence model. In each period, a new customer is born with no loyalty affiliations and lives for two periods. Hence, the market always consists of one "young" and one "old" consumer. Shortening the lifetime of a customer reduces the incentive to invest in customer loyalty and prices might therefore be higher compared to an infinite horizon setting. We find that unless switching cost levels are sufficiently large, both the young and the old customer pay a lower price, on average, than in the case without switching costs. Thus, our main conclusion that switching costs do not necessarily lead to higher prices is robust to this different model formulation as well as a wide range of parameter values.

3. Empirical Model

In the simple model, we observed that equilibrium prices are *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. 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 as the panel length is long relative to the average inter-purchase times and there is extensive price variation. If households are observed to forego large utility increases afforded by a temporary price cut or sale, we can infer that there must be a relatively high level of switching cost.

To infer switching costs from consumer panel data, we must enrich our model to consider multiple differentiated products as well as multiple consumer types or consumer heterogeneity. It is well documented (c.f. Allenby and Rossi 1999) that consumers exhibit a very high degree of heterogeneity with differing product preferences (intercepts) as well as price sensitivities (price coefficient). It is also entirely possible that households will exhibit differing degrees of switching costs. Our approach will be to specify a very flexible model of consumer heterogeneity around a standard logit specification.

Extending the Simple Model

For these reasons, we consider a market populated by many heterogeneous customers. We allow for N different types and assume that there is a continuum of consumers with mass μ_n for each type *n*. The latter assumption is for convenience. As we will see below, it makes the evolution of the state vector deterministic. Demand at the consumer level is identical to the simple model described previously, but the utility parameters are now type-specific. We thus index the choice probability of a consumer by her type. The probability of buying product *j* by a consumer of type *n* in state s_t, for example, is denoted by $P_j(s_t, p_t; n)$.

To summarize the overall state of the market, we need to know the distribution of consumers of different types over loyalty states. Let x_{jt}^n be the fraction of consumers of type *n* who are loyal to product *j*. The vector $x_t^n = (x_{1t}^n, ..., x_{Jt}^n)$ summarizes the distribution over loyalty states for all consumers of type *n*, and $x_t = (x_t^1, ..., x_t^N)$ summarizes the state of the whole market. As before, we denote the state space by X. We write each state as an NJ dimensional vector. Note, however, that by definition, $\sum_{j=1}^{J} x_{jt}^n = 1$ for all types *n*. Hence, the information contained in the state can be described more parsimoniously by a vector that has only $N(J\text{-}1)$ dimensions. This is not theoretically important, but allows us to simplify the solution algorithm for the model on a computer.

Aggregate demand is obtained by summing household level demand over consumer types and loyalty states:

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

In contrast to the simple model, demand is now deterministic. This is a consequence of the assumption that there is a continuum of consumers for each type.

In the discussion of the simple model, we described the law of motion of the state variable at the individual level. The transition of the aggregate state can be easily derived from the transition probabilities of the individual states, as shown in (3). Conditional on a price vector p_t, we can define a Markov transition matrix $Q(p_t; n)$ with elements

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

$Q_{jk}(p_t; n)$ denotes the probability that a household 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

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

Households 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*, $x_{t+1} = f(x_t, p_t)$.

Firm *j*'s current-period profit function is $\pi_j(x_t, p_t) = D_j(x_t, p_t) \cdot (p_j - c_j)$. As in the case of the simple model, firms compete in Markovian strategies, $\sigma_j : X \to \mathbb{R}$. The best response to a profile of competitor's strategies, σ_{-j}, is found from the Bellman equation:

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

Here, $p = (\sigma_1(x),...,\sigma_{j-1}(x), p_j, \sigma_{j+1}(x),...,\sigma_J(x))$.

As in the simple model, we again use Markov perfection as our solution concept. In contrast to the simple model, we cannot prove that a pure strategy equilibrium generally exists. 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. In our empirical application, we can therefore only establish the existence of a pure strategy equilibrium computationally on a case-by-case approach. In Appendix E, we briefly describe the numerical algorithm used to compute the price equilibrium.

Econometric Specification

We have extended the simple model by using a standard multinomial logit model conditional on consumer/household type. The probability that household h chooses alternative j given loyalty to product k is given by

$$P\left(j \mid s=k;\theta^h\right) = \frac{\exp\left(\delta_j^h + \alpha^h p_j + \gamma^h I\{s=j\}\right)}{1+\sum_{k=1}^{J}\exp\left(\delta_j^h + \alpha^h p_j + \gamma^h I\{s=k\}\right)}.$$

The coefficient on consumer loyalty, γ^h, reflects the switching cost. Note that loyalty differs from brand taste heterogeneity. Loyalty reflects inertia in consumer choices whereas persistent preference for a specific product is reflected by the brand intercept, δ_j^h. One might argue that the switching cost could be modeled using a much more complex function of a household'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.

To accommodate differences across households, we use a potentially large number of household types and a continuum of households 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 household in our data will be viewed as "representative" of a type. We will use MCMC methods to construct a Bayes estimate of each household's coefficient vector.

It is well known (c.f. Heckman 1981 and Keane 1997) that state dependence and heterogeneity can be confounded in the sense that mis-specified tightly parametric models of heterogeneity can lead to spurious findings of state dependence. 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 households. For example, there may be sub-populations of households 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 household *h*. The mixture of normals model specifies the distribution of θ^h across households as follows:

$$\theta^h \sim N\left(\mu_{ind}, \Sigma_{ind}\right)$$
$$ind \sim \text{multinomial}\left(\pi\right)$$

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:

$$\pi \sim Dirichlet\left(\alpha\right)$$
$$\mu_k \left| \Sigma_k \sim N\left(\bar{\mu}, \Sigma_k \times a_\mu^{-1}\right)\right.$$
$$\Sigma_k \sim IW\left(\upsilon, V\right)$$
$$\left\{\mu_k, \Sigma_k\right\} \text{ independent}$$

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 household. Conditional on the draws of θ^h, we use an unconstrained Gibbs sampler. Since our goal is to estimate the distribution of model parameters over households, 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.

---

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

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 household parameters. We can average these densities to provide a Bayes estimate of the household 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 household distribution.

Some might argue that you do not have a truly non-parametric method unless you can claim that your procedure consistently recovers the true density of parameters in the population of all possible households. In the mixture of normals model, this requires that the number of mixture components (*K*) increases with the sample size. There are several ways to achieve this. One could put a prior over models with differing numbers of mixture components and use a reversible jump MCMC algorithm to navigate this space of models. However, there are no reliable reversible jump MCMC methods for multivariate mixtures of normals. Our approach is to fit models with successively larger numbers of components and gauge 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

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 354 households to estimate orange juice, and 444 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. This means that the outside good share is very large as is typical in many product categories and analyses of scanner data. In addition, households who adopt a pattern of 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.

In our econometric specification, we have been careful to control for heterogeneity as flexibly as possible to avoid confounding state dependence with unobserved heterogeneity. Even with these controls in place, it is still important to ask which patterns in our consumer shopping panel give rise to the identification of a "switching cost." Table 4 (a) indicates that for each of the brands in the two categories, the marginal purchase probability is considerably smaller than the re-purchase probability. While this evidence is consistent with state dependence, it could also be a reflection of heterogeneity in consumer tastes for brands. Identification of state dependence in our context relies on the frequent temporary price changes typically observed in supermarket scanner data. If there is sufficient price variation, we will observe consumers switching away from their preferred products. The detection of state dependence relies on spells during which the consumer purchases these less-preferred alternatives on successive visits, even after prices return to their "typical" levels.

We use the orange juice category to illustrate the sources of identification of state dependence in our data. First, we observe spells during which a household repeat-purchases the same product. Conditional on a purchase, we observe 1889 such repeat-purchases out of our total 3328 purchases in the category. Second, we observe numerous instances where a spell is initiated by a discount price. We classify each product's weekly prices as either "regular" or "discount," where the latter implies a temporary price decrease of at least 5%. Focusing on non-favorite products (i.e. products that are not the most frequently purchased by a household), nearly 60% of the purchases are for products offering a temporary price discount. In Table 4 (b), we report the repeat-purchase rate for spells initiated by a price discount (i.e. a household repeat-buys a product that was on discount when they previously purchased it). As in Table 4 (a), the re-purchase rate continues to exceed the marginal purchase rate for each brand. While this simple analysis is still insufficient to test for state dependence versus unobserved taste heterogeneity, it nevertheless makes us optimistic that the data are capable of disentangling the two in the context of our choice model. We will revisit this discussion of identification during our discussion of the estimation results below.

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 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. However, the usual state dependence specification appears to generate a better fit than the pure switching costs specification. These results confirm the well-established belief that consumer demand for frequently-purchased CPG products exhibits state dependence[5]. For the remainder of this section, we will focus on the results from the state dependence specification. In the next section, we will contrast the equilibrium implications of state dependence versus switching costs to assess whether the choice of specification alters the substantive predictions we make for pricing.

An interesting finding is the extent to which flexibility in the heterogeneity distribution may be required to "fit" the data. In the orange juice category, a model with a single mixing component (the usual normal random coefficients model) performs relatively well. However, in the margarine category, 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

---

[5] Although not reported, our findings for state dependence are robust to the inclusion of promotional variables such as weekly product features and in-aisle displays.

each of the normal components, we instead plot the fitted marginal densities for several taste coefficients.

In Figures 3 and 4, we plot several fitted densities from the 1, 2 and 5 component mixture models for the margarine data under the state dependence specification. 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 3 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 bimodality implies that there are households who differ markedly in their quality perceptions for margarines (note: the outside good is purchased most often so that the intercepts for all margarine brands are typically negative). 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 4, the price coefficient (upper panel) for the 5-component model leads to a slightly asymmetric density with fat tails. 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 state dependence estimates (lower panel), the 1 component model has a higher mean and thinner tails than the 5-component model.

In Figure 5, we report fitted densities from the orange juice category. These plots illustrate why we do not get the same improvement from more mixing components as we did in the margarine data. In the upper plot (96 oz MM), the marginal densities from the 1 and 2 component models are completely contained within the credibility region around the

5-component model. For the orange juice data, the one component normal approximation seems adequate.

Figures 6 and 7 display the fitted densities of the state dependence premium (i.e. switching costs) 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 6 and 7 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 0.13 for margarine and 0.19 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 0.61 for margarine and 0.60 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.

To confirm the estimated state dependence is not simply spurious correlation between current and lagged choices, we re-estimate two variations of the 5-component state dependence model. We create a new dataset in which the sequence of observed choices for each household is randomized. Once we randomize the choice sequence, the inclusion of state dependence should not improve the model after controlling for heterogeneity if the state dependence correctly captures inertia in purchases (i.e. repeat-purchase spells). However, we would see an improvement in fit if the state dependence is in fact due to spuriously-correlated unobserved heterogeneity not captured by our model. At the bottom of Table 5, we report the results for both orange juice and margarine. After randomizing the choice order, we no longer see any additional improvement in fit from the state dependence in either category (i.e. compare the log marginal density of "w/o switching costs with 5-components" versus "randomized choice order with state dependence and 5 components").

In our second variation of the baseline model, we attempt to rule out serial-dependence in the unobserved taste shocks as an explanation for our state dependence findings. We partition the orange juice sample into two sub-samples. In the first sub-sample, we only include those trips for which the loyalty state was initiated by a price discount (i.e., the product to which the household is currently loyal to was purchased at a discount price level). In the second sub-sample, we only include those trips for which the loyalty state was initiated at a regular price. If choice inertia is driven by serially-dependent taste shocks, then we should only detect state dependence in the sub-sample in which loyalty is initiated without large price changes (i.e. a household switches products due to the random component of utility and then repeat-purchases due to the serial correlation). If the loyalty state is initiated by a price discount, we do not expect persistence in choice models with serially-correlated errors. Although not reported, we find comparable non-zero levels of state dependence in both sub-samples. This finding suggests that serial-dependence does not explain the estimated levels of state dependence.

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.

In Table 7, we report our results relating steady state price and profits levels to the magnitude of the state dependence premium/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 loyalty or state dependence parameter in each cluster by a scale factor reported in the left margin of Table 7. We see that prices decline as state dependence increases from the zero order case of zero state dependence. We compute equilibrium prices not only for the level of state dependence found in our data, but also for much higher levels corresponding to scale factors greater than one. We find that even with state dependence levels twice that revealed in our data, equilibrium prices are lower in the presence of state dependence. At scale factors of 3, OJ prices start to rise above the zero state dependence levels. State dependence is a source of additional profits to the firms as consumer utility increases for any fixed level of price.

We also compute the steady-state prices and profits for the pure switching cost model estimates in Table 8. In both the margarine and orange juice categories, prices fall as switching costs increase away from zero. By comparing the prices for scale factors of zero and one, we see that the estimated level of switching costs in this data result in a fall in equilibrium prices. As expected, this fall is larger in the margarine category than in the orange juice category as the estimated dollar value of switching costs relative to prices is smaller in the orange juice category. Even large scale factors of 4 do not reverse this finding. We still observe equilibrium prices below the levels without switching costs. The bottom half of Table 8 provides equilibrium profit calculations and shows that firms are worse off with switching costs than without.

5. Conclusions

We have demonstrated that equilibrium prices fall as switching costs increase for a variety of stylized and more realistic models. 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 a source of economic rent. 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. We would argue that many classic examples of switching costs such as cellular service carriers or airline frequent flyer programs have associated switching costs in this region.

Our results can be reversed if switching costs reach very high levels or if, indeed, they are infinite as assumed in Beggs and Klemperer. In a 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 suggests that our results are relevant in many situations.

References

Allenby, G. and P. E. Rossi (1999), "Marketing Models of Consumer Heterogeneity," *Journal of Econometrics* 89, 57-78.

Anderson, E.T., N. Kumar and S. Rajiv (2004), "A Comment on: 'Revisiting dynamic duopoly with consumer switching costs,'" *Journal of Economic Theory*, 116, 177-186.

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 (2005), "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.

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," 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.

Früwirth-Schnatter, S. (2001), "Markov Chain Monte Carlo Estimation of Classical and Dynamic Switching and Mixture Models," *JASA* 96, 194-209.

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.

Judd, K. L. (1998), *Numerical Methods in Economics*, 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. (1987a), "Markets with Consumer Switching Costs," *Quarterly Journal of Economics,* 102, 375-394.

Klemperer, P. (1987b), "The Competitiveness of Markets with Switching Costs," *RAND Journal of Economics,* 18, 138-150.

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.

Shi, M., J. Chiang, and B. Rhee (2006), "Price Competition with Reduced Consumer Switching Costs: The Case of "Wireless Number Portability" in the Cellular Phone Industry," *Management Science*.

Shum, M. (2004), "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast Cereals Market," *Journal of Economics and Management Strategy*, 13, 241-71.

Stango, V. (2002), "Pricing with Consumer Switching Costs: evidence from the credit card market," *The Journal of Industrial Economics*, 50, 475-92.

Stokey, N. L. and R. E. Lucas (1989), *Recursive Methods in Economics,* Cambridge: Harvard University Press.

Viard, B. (2003), "Do Switching Costs Make Markets More or Less Competitive?: the case of 800-number Portability," working paper, Stanford University.

Viard, B. (2006), "Do Switching Costs Make Markets More or Less Competitive?: the case of 800-number Portability," forthcoming *RAND Journal of Economics*.

Wernerfelt, B. (1991), "Brand Loyalty and Market Equilibrium," *Marketing Science*, 10(3), 229-245.

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

Appendix A. 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. (2005).

Recall the Bellman equation in the simple model:

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

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

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

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 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 $\varepsilon = (\varepsilon_0,...,\varepsilon_J)$ is

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

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

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

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

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

Here, $\gamma \approx 0.57722$ is Euler's constant. The consumer then chooses the alternative $j = 0,...,J$ 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 consumers 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*.

Appendix D. Overlapping Generations Version of the Simple Model

We now develop an OLG version of our simple state dependence model and examine the robustness of our previous finding that switching costs can lower equilibrium prices. In each period, a new customer is born. The customer lives for two periods and, hence, the market always consists of a "young" and an "old" customer. A customer can be loyal to one of the *J* products, or she can be unattached, i.e. loyal to the outside alternative. If a customer is loyal to the outside alternative, she does not incur a switching cost for any product choice. Otherwise, her demand is as in the model analyzed before. When the young customer is born, she is unattached. If she chooses the outside alternative, she stays unattached in the next period, when she is old. Otherwise, if she buys product *j* she becomes loyal to *j*. The state of the market is now described by $s_t \in \{0,1,...,J\}$, the choice that the currently old customer made in the previous period, $t-1$.

Table D.1 shows the average transaction prices paid by the young and the old customer for different switching cost levels. The model was solved with forward-looking consumers. Due to lock-in, the old customer always pays a higher average price than the young customer. Unless switching cost levels are sufficiently large, however, both the young and the old customer pay a lower price, on average, than in the case without switching costs. The young customer, in particular, generally pays a lower price. Thus, our main conclusion that switching costs do not necessarily lead to higher prices is robust to a different model formulation as well as a wide range of parameter values.

Table D.1

Equilibrium prices in the OLG model

Switching Cost	p^a_{young}	p^a_{old}
0.00	1.81	1.81
0.25	1.78	1.78
0.50	1.76	1.76
0.75	1.74	1.75
1.00	1.71	1.74
1.25	1.69	1.74
1.50	1.66	1.74
1.75	1.63	1.74
2.00	1.59	1.75
3.00	1.41	1.80
4.00	1.19	1.83
5.00	1.01	1.85
6.00	0.90	1.89
7.00	0.83	1.91
8.00	0.81	1.92

Appendix E. Numerical Solution to the Dynamic Program

We use numerical methods to solve for the equilibrium of the pricing game. We first discretize each axis of the state space using a finite number of points, $0 < x_{i0} < x_{i1} < ... < x_{iL} = 1$. We then form a grid representing the whole state space from the Cartesian product of these points. For each point in the grid, we store the value and policy functions of each competitor in the computer memory. For states outside the grid, we calculate the value and policy functions using bilinear interpolation. To solve for the equilibrium, we employ the following algorithm, which is an adaptation of policy iteration applied to the case of the games: start with some initial guess of the strategy profile, $\sigma^0 = (\sigma_1^0, ..., \sigma_J^0)$, and then proceed along the following steps:

1. For the strategy profile σ^n, calculate the corresponding value functions for each of the *J* firms. These value function are defined by the Bellman equation <equation reference>, where the right hand side of the Bellman equation is not maximized, but instead evaluated using the current strategy profile σ^n.
2. If $n > 0$, check whether the value functions and policy functions satisfy the convergence criteria, $\left\| V_j^n - V_j^{n-1} \right\| < \varepsilon_V$ and $\left\| \sigma_j^n - \sigma_j^{n-1} \right\| < \varepsilon_\sigma$ for all firms *j*. If so, stop.

Update each firm's strategy using the Bellman equation <reference>. In contrast to step 1, the maximization on the right hand side is now carried out. Denote the resulting new policies and value functions by σ_j^{n+1} and V_j^{n+1}, and return to step 1.

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	1.825	1.679	1.756	0.272	0.245	0.755
0.50	1.840	1.542	1.706	0.310	0.254	0.746
0.75	1.855	1.397	1.659	0.350	0.261	0.739
1.00	1.867	1.244	1.613	0.390	0.268	0.732
1.25	1.877	1.083	1.569	0.431	0.273	0.727
1.50	1.883	0.915	1.525	0.470	0.276	0.724
1.75	1.887	0.741	1.482	0.508	0.278	0.722
2.00	1.887	0.560	1.438	0.544	0.278	0.723
3.00	1.931	0.000	1.437	0.681	0.234	0.766
4.00	2.153	0.000	1.860	0.823	0.130	0.870
5.00	2.335	0.000	2.183	0.913	0.064	0.936
6.00	2.475	0.000	2.404	0.961	0.028	0.972
7.00	2.616	0.000	2.584	0.983	0.012	0.988
8.00	2.829	0.000	2.813	0.992	0.006	0.994

Note: The results were calculated for product intercepts = 1.0, price coefficient = 1.0, and mean outside good utility = 0.0. The discount factor is $\beta = 0.998$. The table shows the prices of firm 1 and firm 2 in state 1, and the average transaction price paid by the customer. The table also shows the purchase probabilities for the products in state 1, and the probability that the customer stays loyal.

Table 2

Equilibrium prices under different switching cost levels if the outside good intercept is adjust to keep the market size constant

Switching Cost	p_1	p_2	p^a	Purchase Prob. 1	Purchase Prob. 2	Prob. Stay Loyal	V_1	V_2	V_1^0	V_2^0
0.00	1.808	1.808	1.808	0.236	0.236	0.764	154.1	154.1	154.1	154.1
0.25	1.802	1.658	1.734	0.258	0.232	0.768	151.0	150.8	154.2	154.0
0.50	1.794	1.500	1.662	0.279	0.227	0.773	146.9	146.6	154.3	153.9
0.75	1.784	1.335	1.593	0.298	0.221	0.779	142.1	141.6	154.5	153.8
1.00	1.773	1.165	1.528	0.317	0.214	0.786	136.7	136.1	154.7	153.6
1.25	1.762	0.991	1.467	0.334	0.207	0.793	131.0	130.2	154.9	153.3
1.50	1.750	0.813	1.410	0.349	0.199	0.801	125.1	124.2	155.3	153.0
1.75	1.738	0.631	1.356	0.363	0.191	0.809	119.2	118.1	155.7	152.6
2.00	1.727	0.445	1.307	0.376	0.183	0.817	113.4	112.1	156.2	152.1
3.00	1.732	0.000	1.352	0.421	0.119	0.881	116.2	113.8	160.1	148.1
4.00	1.782	0.000	1.607	0.450	0.049	0.951	141.2	135.1	169.8	138.4
5.00	1.812	0.000	1.740	0.460	0.019	0.981	156.1	140.8	190.8	117.5
6.00	1.844	0.000	1.816	0.458	0.007	0.993	172.1	134.2	225.0	83.2
7.00	1.896	0.000	1.885	0.448	0.003	0.997	198.2	113.8	261.8	46.4
8.00	1.972	0.000	1.967	0.430	0.001	0.999	236.1	80.3	287.2	21.1

Note: In this example, the outside good intercept under no switching costs is $\delta_0 = 0$. For positive switching cost levels, δ_0 is adjust 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. The discount factor is $\beta = 0.998$. The table shows the prices of firm 1 and firm 2 in state 1, and the average transaction price paid by the customer. The table also shows the purchase probabilities for the products in state 1, and the probability that the customer stays loyal. 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	1.52
premium 64oz MM	2.62	1.88	0.96
96 oz MM	3.41	2.12	2.01
premium 64oz TR	2.73	2.07	3.96
64 oz TR	2.26	1.29	0.93
premium 96 oz TR	4.27	2.73	1.09
no-purchase (% trips)	89.53		
# households	355		
# trips per household	89.51		
# purchases per household	9.37		

Margarine

Product	Retail Price	Wholesale Price	% trips
Promise	1.69	1.22	2.93
Parkay	1.63	1.02	1.11
Shedd's	1.07	0.83	2.83
ICBINB	1.55	1.11	5.26
no-purchase (% trips)	87.86		
# households	429		
# trips per household	81.02		
# purchases per household	9.89		

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	-18302.52	-15888.78
	5-component	-13126.00	-11622.32
with pure Switching Costs	No heterogeneity	-16170.00	-15451.95
	1-component	-13237.49	-11600.43
	2-component	-13217.10	-11594.49
	5-component	-13107.69	-11574.06
	10-component	-13099.00	-11559.43
with State Dependence	No heterogeneity	-15735.04	-14330.43
	1-component	-13182.76	-11503.13
	2-component	-13155.41	-11510.86
	5-component	-13027.58	-11505.56
	10-component	-12994.88	-11478.65
Randomized Choice Sequence with State Dependence	5-component	-13144.81	-11618.54

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 for the State Dependence Model

Steady State Prices

Scale Factor	Prices					
	16-oz Tub Margarine				Refrigerated Orange Juice	
	Promise	Parkay	Shedd's	ICBINB	Minute Maid	Tropicana
0	1.44	0.91	0.82	1.42	1.51	1.79
1	1.42	0.91	0.81	1.39	1.49	1.77
2	1.41	0.91	0.81	1.37	1.48	1.79
3	1.39	0.91	0.81	1.35	1.50	1.84
4	1.38	0.90	0.81	1.34	1.54	1.94

Steady State per Period Profits

Scale Factor	Profits					
	16-oz Tub Margarine				Refrigerated Orange Juice	
	Promise	Parkay	Shedd's	ICBINB	Minute Maid	Tropicana
0	3.74	0.16	3.39	3.50	1.42	9.11
1	4.07	0.16	3.93	3.81	1.74	14.8
2	4.47	0.16	4.57	4.17	2.58	26.63
3	4.94	0.16	5.31	4.61	4.90	51.71
4	5.50	0.17	6.18	5.13	11.03	101.1

Table 8
Equilibrium Prices and Profits for the Pure Switching Cost Model

Steady State Prices

Scale Factor	Prices					
	16-oz Tub Margarine				Refrigerated Orange Juice	
	Promise	Parkay	Shedd's	ICBINB	Minute Maid	Tropicana
0	1.16	0.92	0.82	1.20	1.53	1.82
1	1.11	0.88	0.76	1.16	1.49	1.80
2	0.95	0.84	0.70	1.08	1.47	1.79
3	0.93	0.77	0.72	1.13	1.45	1.79
4	0.94	0.72	0.72	1.18	1.44	1.79

Steady State per Period Profits

Scale Factor	Profits					
	16-oz Tub Margarine				Refrigerated Orange Juice	
	Promise	Parkay	Shedd's	ICBINB	Minute Maid	Tropicana
0	4.23	0.38	2.90	10.09	1.66	15.48
1	2.31	0.18	1.91	8.08	1.33	15.01
2	1.56	0.08	1.29	6.88	1.12	14.68
3	1.36	0.05	1.13	6.84	0.98	14.45
4	1.31	0.03	1.08	7.01	0.89	14.29

Figure 1

Average Transaction Price vs. State Dependence Level

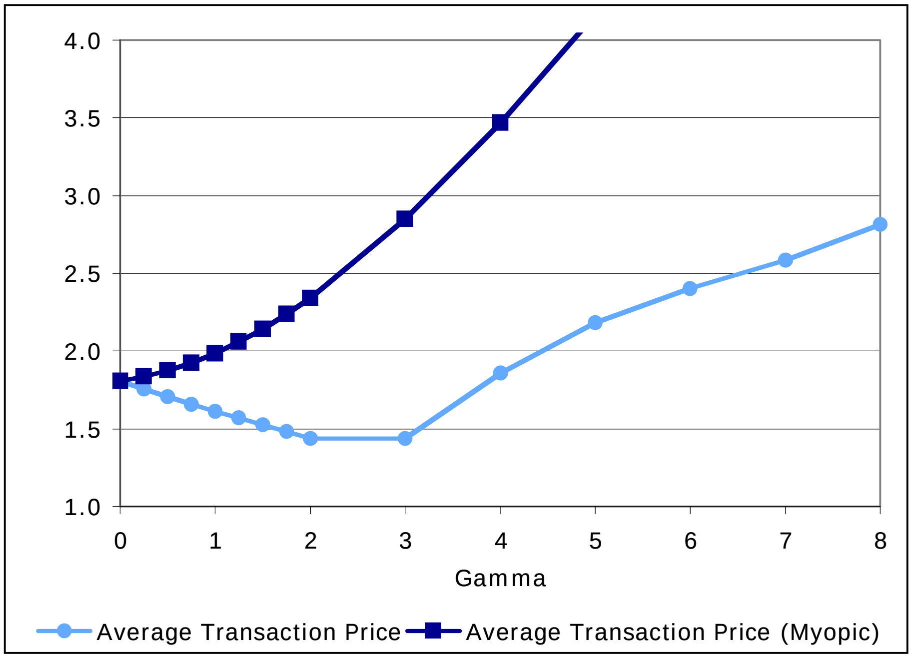


Figure 2
Average Transaction Price vs. Switching Costs

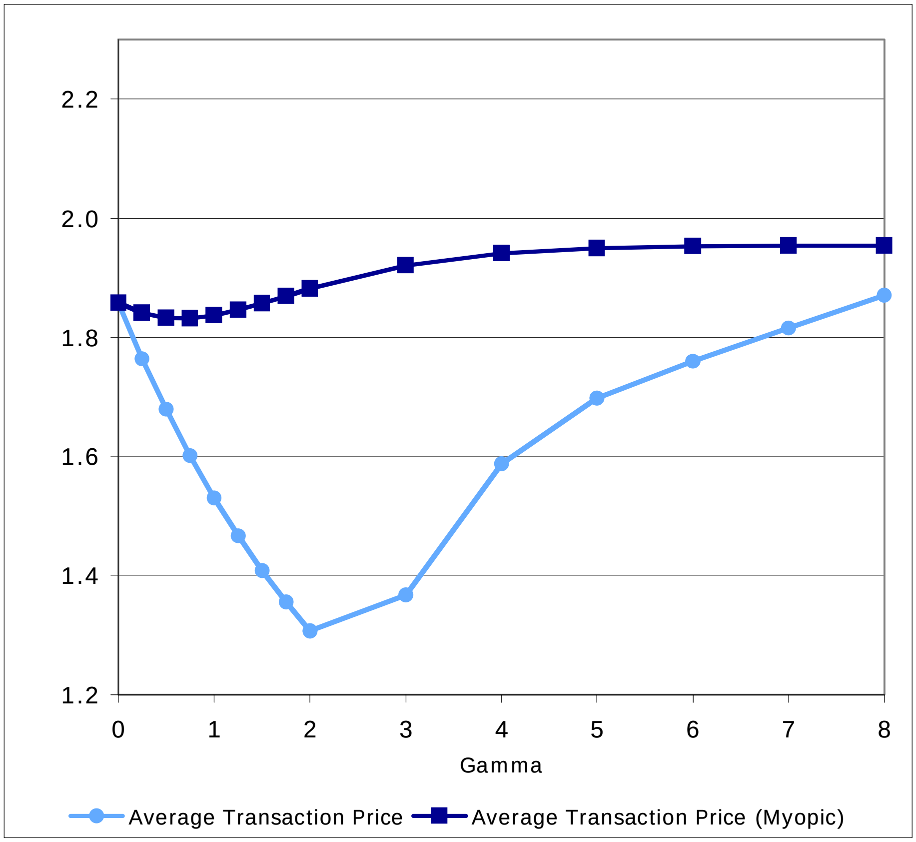


Figure 3
Fitted Densities for Shedd's and ICBINB Brand coefficients (Margarine)

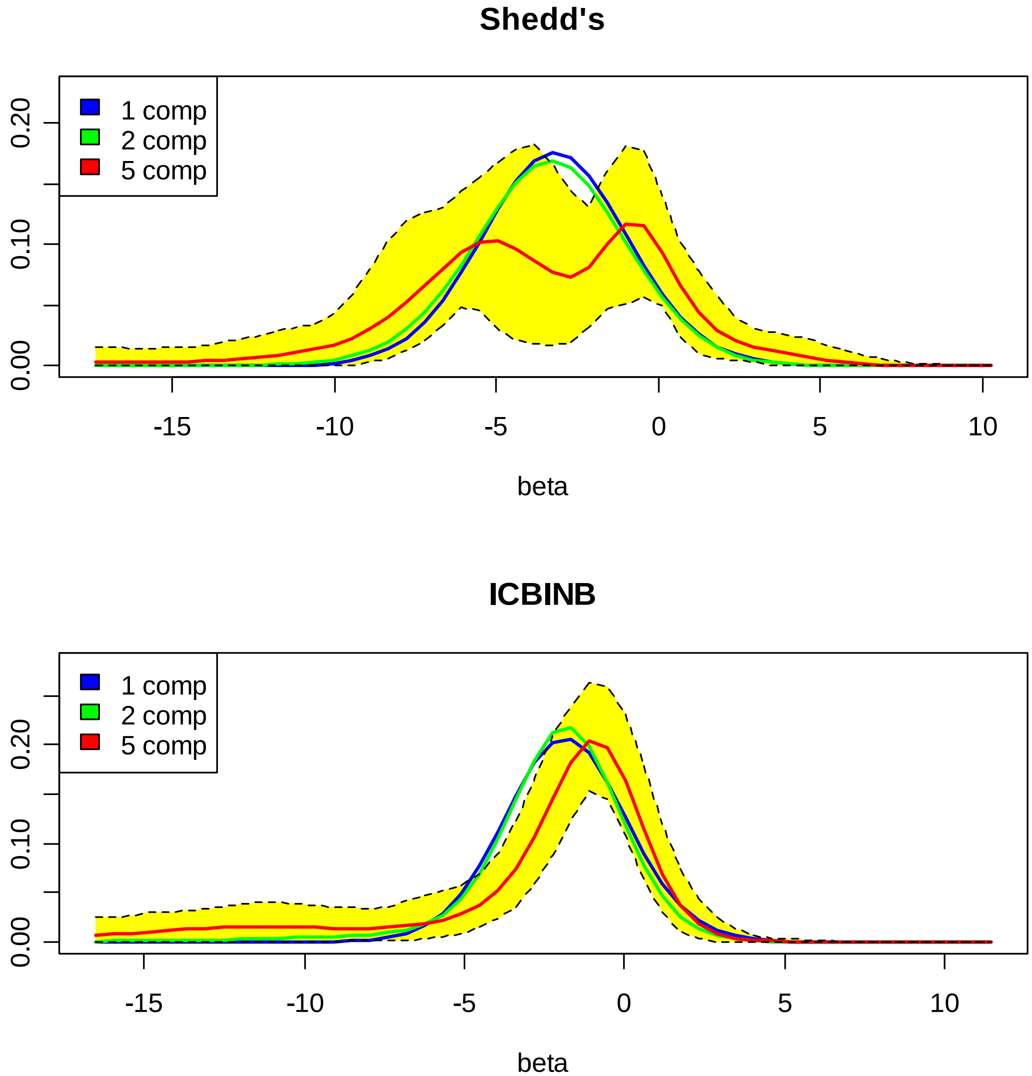
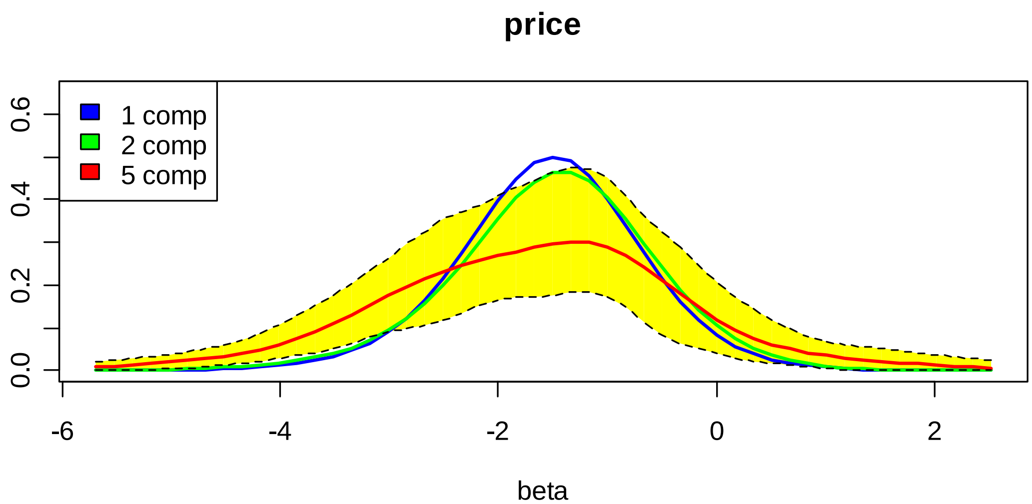


Figure 4
Fitted Densities for Price and State Dependence Coefficients (Margarine)



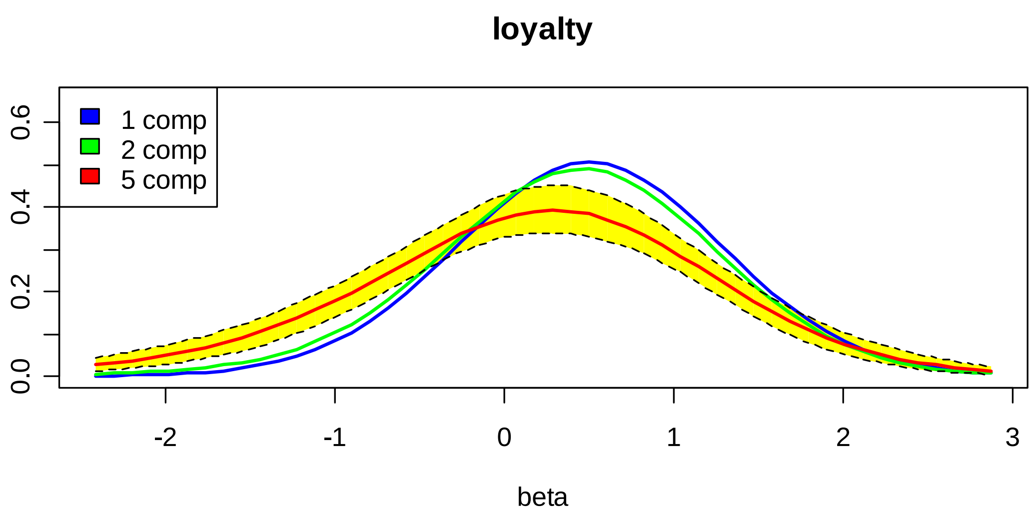
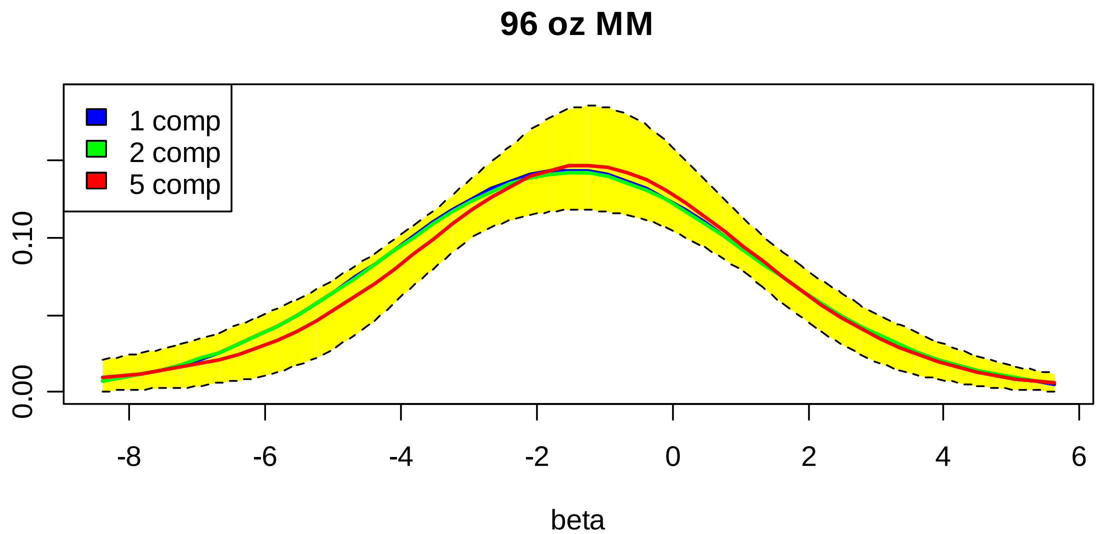


Figure 5
Fitted Densities for 96 oz Minute Maid and 64 oz Tropicana Brand coefficients (Orange Juice)



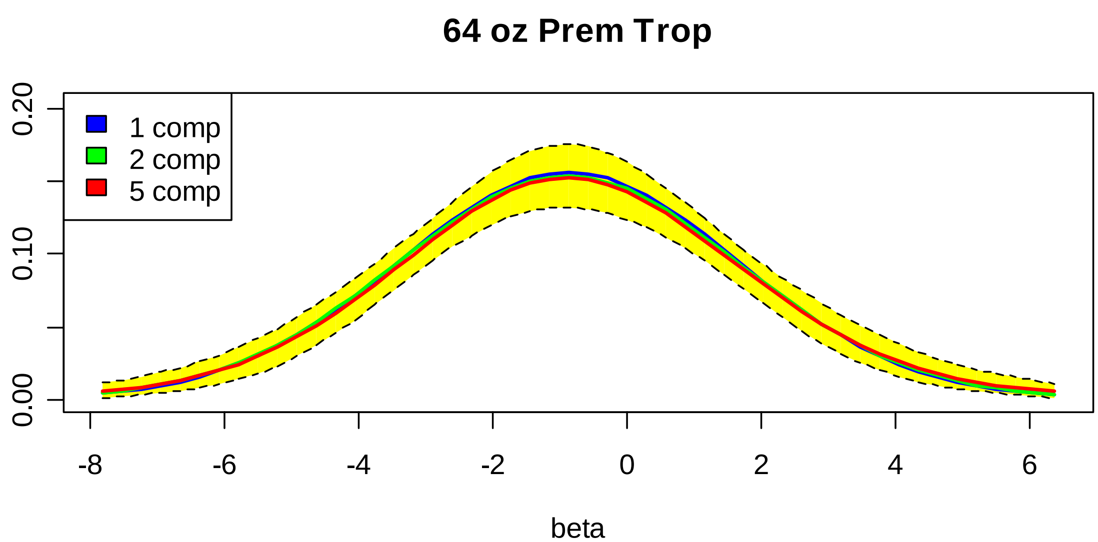


Figure 6

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

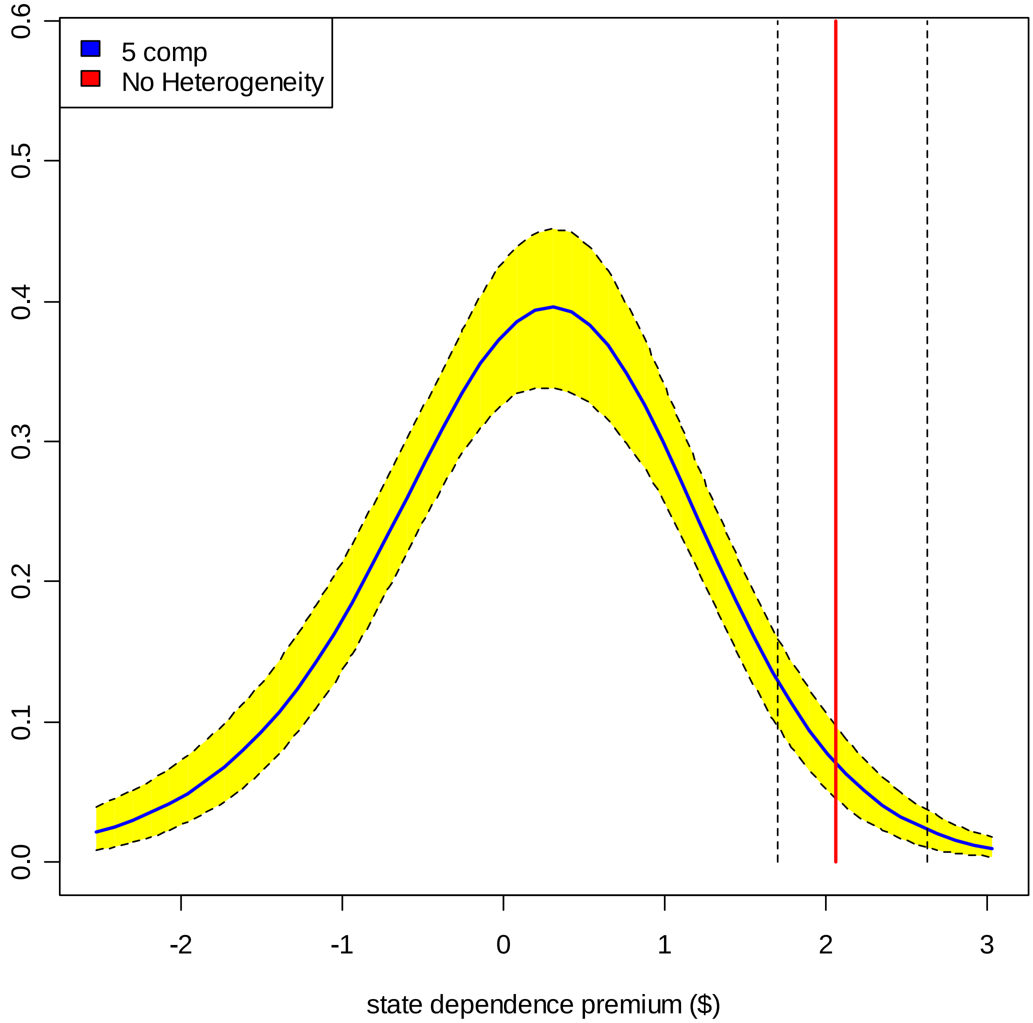


Figure 7

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