Quant Mark Econon
DOI 10.1007/s11129-017-9187-8

Implications of parent brand inertia for multiproduct pricing

Polykarpos Pavlidis[1] · **Paul B. Ellickson**[2]

Received: 18 August 2015 / Accepted: 3 July 2017

Abstract This paper explores and quantifies the importance of parent brand state dependence to forward looking pricing outcomes in the area of umbrella branding and multi-product firms. We show through numerical simulations that loyalty (inertia) to the parent brand can decrease prices and reduce profits, as well as mitigate or even reverse the benefits of joint profit maximization relative to sub-brand profit maximization. These effects are mediated by brand asymmetries and the relative magnitude of sub-brand state dependence effects. Empirically, we focus on the Yogurt category, where we consider parent brands with several sub-brands. Using household level scanner data, we estimate the parameters that characterize consumer demand while flexibly accounting for consumer heterogeneity. We also estimate unobserved product costs based on a forward looking price setting game. Through counterfactual analysis, we study the overall effect of parent brand state dependence on prices and profits, as well as the empirical impact of joint profit maximization and changes in firms' beliefs regarding consumer inertia. Our findings have implications for markets where demand is likely characterized by parent brand dynamics.

Keywords State dependence · Multiproduct firms · Umbrella branding · Forward looking prices

JEL Classification M2 · M3 · L1

✉ Paul B. Ellickson
paul.ellickson@simon.rochester.edu

Polykarpos Pavlidis
polis.pavlidis@gmail.com

[1] Nielsen, Evanston, IL, USA

[2] University of Rochester, Rochester, NY, USA

Published online: 14 July 2017

1 Introduction

Consumer packaged goods are frequently marketed under a brand name that encompasses a wide variety of both physical (size, package type) and taste attributes (low fat, enriched with fruit, etc.). The practice of umbrella branding is observed across many consumer products including durables (e.g. automobiles, electronics, computers) and services (e.g. TV shows, hotels). The usual interpretation of a brand name is one of product identity: brands are identifiable in consumers' minds and they are associated with unique images that are both comparable and subject to preference rankings. Past exposure to a brand's performance should guide a consumer's purchase decision the next time they shop in a given product category.

Several studies have established that a consumer's *most recent* experience with a brand is an important factor in their next purchase decision (Seetharaman et al. 1999; Seetharaman 2004; Anand and Shachar 2004; Horsky and Pavlidis 2011). This type of persistence in demand (often described as state dependence or switching costs) has unique and complex implications for the pricing behavior of firms. While demand-side state dependence has been studied in academic research extensively over the last two decades, the pricing side of the problem has received considerably less attention. In general, state dependence in the demand for frequently purchased goods has been found to create two countervailing forces on the equilibrium prices that forward looking firms may charge (Farrell and Klemperer 2007; Dubé et al. 2009). On one hand, it places an upward pressure on prices because consumers' past choices partially "lock them in", making them less likely to switch. On the other hand, this "lock-in" also creates a threat of lost future sales should today's prices be set too high, yielding a downward force.

The goal of this paper is to study the effect of state dependence on a multiproduct firm's forward looking pricing problem and its impact on profitability. In particular, we focus on the implications of parent brand state dependence under umbrella branding. Given that the existence of switching costs is often motivated by brand loyalty (e.g. Klemperer 1995), this is an important question that has not been fully addressed in the extant literature. There are two issues to address in distinguishing the forward looking pricing behavior of single versus multiproduct firms. First, it is important to identify whether current demand for a specific product depends on past experience with that product alone or if it also depends on past experience with other products offered by the same firm, effectively rendering them dynamic complements. This is a measurement issue. Second, when there is state dependence to the parent firm, the two countervailing forces described above become more complex, since a firm that maximizes the joint profits of a portfolio of products must account for the impact of each product's price on the future demand of its full set of offerings. This is a structural question.

It is well known that optimal, myopic, multi-product pricing increases market power by allowing the firm to internalize contemporaneous cross-cannibalization within its product set, thereby better leveraging its "local" monopoly power. The existence of firm state dependence has non-trivial implications for the magnitude of the benefit that joint pricing provides because, for a forward looking multiproduct firm, two additional countervailing forces come into play. In addition to the "harvest-invest" dilemma that single product price managers face, there is also a cross-product "harvest-invest" dilemma. Consider the simple case of two product offerings, under the same brand name, where the relevant pricing manager maximizes joint profits and consumer demand is state dependent at both the sub- and parent-brand level. Deciding on the price of each sub-brand, a forward looking manager must carefully weigh two options: 1) charge a little more for each sub-brand since part of the lost demand will inter-temporally accrue to the other sub-brand through state dependence to the parent brand, or 2) lower the price of each sub-brand since this will create greater future demand for the other sub-brand as well (the dynamic complement effect). The total impact of any parent brand loyalty that past experiences create is naturally determined by the balance of these two countervailing forces.

To illustrate this important tension, we first develop a set of theoretical results to explore the impact of parent brand state dependence on firm prices and profits. Our approach draws heavily on the framework developed by Dubé, Hitsch and Rossi (DHR 2009), which we extend here to the case of multiproduct firms with multiple layers of state dependence. Analyzing the demand for parent brands of yogurt that offer a portfolio of sub-brands, we model state dependence (inertia) that applies both to the overall parent brand (PBSD) and to all constituent sub-brands (SBSD), as well as to the product type. Solving our computational model under a variety of conditions, we first show that PBSD can lead to both lower prices and profits in equilibrium, so long as the levels of state dependence are relatively moderate. For these moderate levels of PBSD, the incentive to invest outweighs the incentive to harvest, leaving prices lower and firms worse off in equilibrium (than a setting in which PBSD effects are absent).

However, the presence of both PBSD and SBSD also make the firm's multi-product pricing problem more complex. In particular, we find that for certain regions of the parameter space, a forward looking firm can be actually better off operating its sub-brands as independent profit centers, as this allocation of decision rights breaks some of the connections between current actions and future profits, thereby softening the dynamic competition that state dependence fosters. We then further demonstrate that this countervailing force is mitigated by greater levels of SBSD (which does not exhibit this tension), rendering the overall effect an empirical question. To heighten the connection to real-world settings, we introduce brand asymmetries to the theoretical model, demonstrating that the equilibrium impact of PBSD also depends on the brand's relative strength in the product market. In particular, we find that PBSD makes the strongest firm stronger, at the expense of its weaker rival. These asymmetries are critical to understanding the patterns revealed in our empirical application.

Turning to this application, we study the empirical role of firm state dependence on dynamic multiproduct pricing in the CPG yogurt category. We first establish the empirical significance of PBSD (in addition to SBSD) by estimating a structural model of individual-level demand using household level scanner data. Following DHR (2009), we employ a rich distribution of consumer heterogeneity to account for possible confounds between state dependence and unobserved tastes. We also use this demand system to recover estimates of product-level costs, employing the forward simulation approach developed by Bajari et al. (2007). With these cost estimates in hand, we then present a suite of counterfactuals that evaluate and build upon the insights provided by our theoretical analysis.

We first show that, as suggested by the theoretical model and consistent with DHR, higher levels of PBSD do generally lead to lower prices. However, the impact on individual firms depends on their relative position in the market. In particular, the U-shaped pattern of pricing applies only to the dominant brand - the prices of the smaller brands decrease uniformly as PBSD increases. Interestingly, we also find that, while the dominant firm sets prices that are lower than if PBSD were eliminated, increasing the level of PBSD from its estimated value would lead it to increase its price and earn even higher profits, at the expense of its weaker rivals. As the effect of PBSD becomes stronger, the market leader is able to capitalize on the dynamic complementarities of its strong sub-brands even though the market as a whole becomes more competitive. We also show that the leading parent brand would earn lower profits if, at the estimated levels of PBSD, it lost its dynamic complementarities through marketing one or more of its sub-brands separately and not under a common umbrella brand. In this case, the market leader would increase its price, but earn lower profits (due to more intense competition with its now relatively strengthened rivals who, in this exercise, would still enjoy their dynamic complementarities). Turning next to the question of multiproduct pricing, we show that, in our empirical setting, joint profit maximization actually yields higher prices and profits for all firms, as the standard intuition would suggest. We conclude that, in the context studied here, the sub-brand effects outweigh those of the parent brand with regards to the tension between dynamic competition and the harvest incentive. Finally, we ask whether forward looking firms would actually be better off if they were unaware of the existence of either sub-brand or parent brand state dependence, and priced accordingly. We find that they in fact would be, as this would serve to soften price competition. However, all firms would need to be equally naive in their beliefs for this to occur. If only one firm followed this limited information behavior, it would be at a disadvantage compared to its more sophisticated rivals.

This paper makes several contributions to the literature. First, we establish the empirical relevance of parent brand state dependence by estimating a corresponding structural model of state dependent demand. We also provide a method for using these estimates to recover inherently unobserved product level costs. Second, we provide new theoretical insights into the impact of PBSD on firm pricing behavior, highlighting new implications regarding multiproduct pricing, the moderating influence of SBSD, and the role of brand asymmetries. Third, we explore the practical impact of PBSD on firms in the yogurt market, providing additional support for some of our theoretical insights and identifying the limits of others.

The structure of the paper is as follows. In Section 2 we discuss the related literature. Section 3 presents our model of state dependent demand and forward looking prices, as well as our theoretical simulations. In Section 4, we provide an overview of the empirical setting and describe the data used in our application. Section 5 contains the results of estimation, along with the full set of counterfactual exercises. Section 6 concludes.

2 Related literature

2.1 State dependence and pricing implications

The implications of state dependence for firm behavior, market structure, and consumer welfare have been analyzed extensively in both the economics and marketing literatures. Farrell and Klemperer (2007) provide a comprehensive survey, covering both empirical and theoretical results. Of particular relevance to our study is the impact on product market competition. Klemperer (1995) expresses the traditional view that switching costs make markets less competitive, as the lock-in effect is likely to dominate the incentive to harvest. Viard (2007) empirically tests the impact of portability switching costs on prices for toll-free services and finds a positive impact - prices decreased after the introduction of portability, which reduced customer switching costs.

DHR (2009) challenged the conventional wisdom, showing that switching costs can result in pricing equilibria which are more competitive - characterized by both lower manufacturer prices and profits. The main intuition behind this result is that if switching costs arising from state dependence are relatively low, the strategic effect of attracting and retaining customers in a competitive market can outweigh the harvest incentive. They demonstrate that switching costs for the categories of refrigerated orange juice and margarine, estimated using consumer transaction data, are well within the range that increases competition. In related work, Dubé et al. (2008) establish a similar result for category managers setting a portfolio of prices in a product category: the prices of higher quality products decline relative to lower quality substitutes in the presence of greater loyalty. Cabral (2008, 2016) showed that the theoretical results of DHR (2009) generalize to a much broader class of models.

These new empirical findings have sparked additional theoretical research examining the impact of switching costs on equilibrium pricing. Arie and Grieco (2014) demonstrate theoretically the existence of an additional "compensating" effect that pushes prices downwards when switching costs increase from zero to moderate levels. Based on this effect, single product firms may reduce prices to compensate marginal consumers who are loyal to rivals. In their setting, the investing effect of attracting future loyal customers decreases prices when switching costs increase, provided that no "very" dominant firm is present in the market. Similar theoretical results, with prices decreasing in the presence of relatively "low" switching costs, are reported by Doganoglou (2010), who further proves the existence of an MPE that supports switching in equilibrium.

2.2 Multiproduct firms / umbrella branding

In the context of multiproduct firms, Erdem (1998) shows the existence of cross-category brand spillover effects. Consumers' choices in the categories of toothpaste and toothbrush reveal behavior consistent with the predictions of signaling theory. A positive experience with a brand in one category reduces the perceived uncertainty of the brand's quality in a different but related category. Erdem and Sun (2002) extend the above framework to include advertising. For the same product categories, they show that advertising, as well as experience, reduce uncertainty about quality and also affect preferences. Under umbrella branding firms enjoy marketing mix synergies that go beyond advertising. Draganska and Jain (2005) examine empirically the category of yogurt and find that consumers perceive product lines to be different, that they are willing to pay for quality, but that they consider flavors belonging to the same product line to be of comparable quality. Miklos-Thal (2012) shows that forward looking firms have incentives to employ umbrella branding for new products only if their existing products are of high quality. Anand and Shachar (2004) test and confirm that consumer loyalty to multiproduct firms is also driven by information-related benefits. Anderson and De Palma (2006) outline how firms tend to restrict their product ranges in order to relax price competition but that in turn generates relatively high profits which attract more entrants to the market.

3 Model setting

3.1 Demand utility function

We model demand using a discrete choice framework. We assume consumers are in the market for yogurt every week they visit a grocery store or supermarket, denoting the purchase choice set by J and reserving the subscript 0 for the outside option of not buying yogurt. An individual consumer h's conditional indirect utility from consuming product j is composed of a linear index, U, and an additive Type I Extreme Value demand shock, ε, yielding the well known logit probabilities for purchase of each sub-brand.

$$\begin{aligned} u_{hjt} &= \beta_{hj} - \beta_{hp} P_{hjt} + \beta_{hPB} PBL_{hjt} + \beta_{hSB} SBL_{hjt} + \beta_{hT} TL_{hjt} + \varepsilon_{hjt} \\ &= U_{hjt} + \varepsilon_{hjt}, \quad j \in \{1, ..., J\} \end{aligned} \tag{1}$$

$$u_{h0t} = \varepsilon_{h0t} \tag{2}$$

$$Pr(C_{hjt} = 1) = Pr\left(u_{hjt} > u_{hkt} \, \forall k \in J, \; k \neq j, \; \& \; \max u_{hkt} > u_{h0t}, \; \forall k \in J\right) \tag{3}$$

$$Pr(C_{hjt} = 1) = \frac{e^{U_{hjt}}}{1 + \sum_{k \in J} e^{U_{hkt}}} \tag{4}$$

In the utility functions (1), β_{hj} denotes the intrinsic preference for each choice alternative j, while P_{hjt} represents the alternative's net price at week t. The choice set is defined over sub-brands of popular parent brands that may cover one or more different alternatives. The different levels of state dependence, for the *Parent Brand* and/or the specific *Sub-brand*, are captured with the variables PBL_{hjt} and SBL_{hjt} respectively. These are both indicator variables, the former taking the value one if the household's last purchase was for the parent brand of the choice alternative j

and the latter being equal to one if last purchase was of the same sub-brand.[1] For a given alternative j there are three possibilities with respect to the brand related loyalty variables: a) both PBL_{hjt} and SBL_{hjt} could be equal to one if the same product line was bought last time, b) both PBL_{hjt} and SBL_{hjt} could be zero if an alternative of a different brand than j was bought or c) PBL_{hjt} could be equal to one and SBL_{hjt} could be equal to zero if another product line of the same parent brand was last chosen by the consumer. Coefficients β_{hPB} and β_{hSB} thus capture the effect of the previously purchased brand on the current period's choice. Significant positive values for both would imply the existence of demand state dependence associated with both the specific sub-brand as well as the parent brand. If consumers tend to repeat their choices only with respect to specific alternatives, but there is no purchase reinforcement from the parent brand, then β_{hPB} should not be statistically different from zero.

In addition to brand related state dependence, we further allow for state dependence with respect to the type of yogurt. The indicator variable TL_{hjt} takes the value one for sub-brands that are of the same type (i.e. *light* or *regular*) as the last sub-brand purchased by consumer h. By including this additional state variable in the utility function, we ensure that the brand related variables do not confound sub-brand/brand state dependence with state dependence to yogurt type. Finally, we note that concerns regarding the potential endogeneity of price are mitigated by the inclusion of brand intercepts. As is the case in many scanner data settings, the bulk of the variation in prices occurs across brands, presumably due to highly persistent differences in ingredient quality or brand capital (both unobserved to the researcher), which will be absorbed by these intercepts.

Dubé et al. (2010) highlight an important confound between state dependence and unobserved preference heterogeneity: absent a rich representation of consumer preferences, measured state dependence could simply be an artifact of unobserved taste variation. To address this concern, we follow their approach, which controls for heterogeneity by approximating the distribution of household coefficients across the population with a flexible mixture of Normal distributions (Rossi et al. 2005).

$$\beta_h \sim N(\bar{\beta}_{l_h}, \Sigma_{l_h}) \tag{5}$$

$$l_h \sim multinomial\,(\pi) \tag{6}$$

Including this flexible heterogeneity distribution should account for any persistence in choice that is driven by consumer tastes. To further address any possible confounds, we also provide robustness tests assessing possible misspecification of the heterogeneity distribution or the presence of consumer learning.

[1]In our empirical application, a small number of household Yogurt trips (roughly 3.4%) involve the purchase of more than one sub-brand. For our empirical analysis, we treat these outcomes as separate (contemporaneous) purchases, effectively allowing these consumers to be loyal to more than one product (or brand).

3.2 Demand estimation

Following Dubé et al. (2010), we estimate the demand model using a Bayesian MCMC framework. Given diffuse standard priors for the parameters, we take draws through a mixed Gibbs sampler with a Metropolis-Hastings step. To assess the convergence of the posterior parameters, we visually inspect the series of draws and experiment with adding more draws to ensure that the estimates are stable. We compare different model specifications based on the Decision Information Criterion (DIC).[2]

$$DIC(y|M) = 2 \times \hat{D}_{Avg}(y|M) - D(y, \hat{\theta}|M) \tag{7}$$

where $D(y, \hat{\theta}|M) = -2 \times \log\left[l(\hat{\theta}|M)\right]$ and $\hat{D}_{Avg}(y|M) = \frac{1}{R}\sum_{r=1}^{R} D(y, \theta^r|M)$. Lower values of DIC indicate a preferred model.

3.3 Total demand & evolution of states

While we use a very flexible heterogeneity distribution for demand estimation, we use a single consumer type to approximate total demand and the evolution of states that lead to future discounted profits. While it is straightforward in theory to add more than one consumer type, the large number of choice alternatives considered here quickly renders the pricing game intractable in practice.[3] To mitigate this dimensionality problem we restrict attention to a single consumer type for purposes of the pricing model.[4] Note that this does have implications for the firms' incentives to price discriminate, which could be lessened with fewer types.

Returning to the demand setup, at any period t, a fraction s_{kt} of all consumers will have chosen a particular alternative k for their last purchase. The fraction of consumers loyal to each sub-brand is the key state variable driving the firm's dynamic

[2]DIC (defined in Eq. 7) is based on the Deviance of the model ($D(y, \hat{\theta}|M) = -2 \times \log[l(\hat{\theta}|M)]$) which reflects discrepancy between the model and data. DIC is a more complete model selection metric compared to the simple Deviance because it reflects coefficient uncertainty and penalizes model complexity (Gelman et al. 2004; Gamerman and Lopes 2006).

[3]For example, in a market of one type and twelve sub-brands belonging to less than twelve parent brands, the state space has dimension of eleven ((12-1)x1). If one wants to add a consumer type, the dimensionality increases to twenty-two ((12-1)x2). Additional details regarding the overall size of the state space (and dimensions of the associated grid) and the clear trade-off between consumer types and included brands are provided in the Appendix.

[4]In a previous version of the paper, we instead allowed for two consumer types, but a smaller set of sub-brands. While the main implications were qualitatively similar (results available upon request), the smaller set of products is insufficient to highlight the strategic impact of multi-product pricing. A potential advantage of specifying the pricing game with a single consumer type is that it makes it easier for firms to track the current state when applying our methodology in practice. It is significantly more costly to track the loyalty of different consumer types in conjunction with each type's preference and price sensitivity parameters.

pricing problem. The probability of observing choice *j* conditional on *k* having been chosen last is denoted by $Pr(C_{jt} = 1|k)$. Since each product is associated with one parent brand only, tracking past sub-brand choices is sufficient to characterize the vector of state variables at both the brand and sub-brand level. That is, both PBL_{hjt} and SBL_{hjt} are uniquely defined by the last product line chosen. Total demand for product j is then given by

$$D_{jt} = \sum_{k=1}^{J} s_{kt} \times Pr(C_{jt} = 1|k) \tag{8}$$

where $\sum_{k=1}^{J} s_{kt} = 1$. The state vector $s_t = (s_{1t}^{,}..., s_{Jt})'$ evolves deterministically over time based on the choices made by consumers the previous period, following a Markovian transition matrix Q. The element in the j_{th} row and k_{th} column of this transition matrix is denoted by Q_{jk}, and is equal to the conditional probability that a consumer chooses sub-brand *j*, given that he or she is loyal to sub-brand *k*.

$$s_{t+1} = g(P_t, s_t) = Q(P_t) \times s_t \tag{9}$$

$$Q_{jk}(P_t) = \begin{cases} Pr(C_{jt} = 1|k) + Pr(C_{0t} = 1|k) & if \quad j = k \\ Pr(C_{jt} = 1|k) & if \quad j \neq k \end{cases} \tag{10}$$

The transition matrix is naturally a function of the demand parameters and prices of all sub-brands. Given the choice probabilities (4) and the transition function (9), the share at time *t* of a specific alternative *j* belonging to brand *B* depends on its share at *t-1* and also on the share at *t-1* of the other alternatives that belong to parent *B*. The dependence is operationalized through i) the state vector at *t* which determines the distribution of loyalty values across consumers for the particular period, and ii) the three state dependence variables, PBL_{hjt}, SBL_{hjt}, TL_{hjt}, and their associated coefficients in the utility function.

3.4 Pricing behavior

We now specify a model of forward looking pricing based on the demand system described above. There are K firms in the market, each of whom may carry several sub-brands $j \in f, \; j = 1, .., J_f, \; f = 1, ...K$. Time is discrete and the profits of each firm *f* in any period *t* are a function of the state vector and all market prices: $\pi_{ft}(s_t, P_t) = \sum_{j \in f} D_{jt} \times M \times (P_{jt} - c_j)$ where s_t is the complete state vector, c_j is the marginal cost of sub-brand *j, M* is the total market size, and *D* and *P* denote demand and prices respectively. Having the state vector enter the profit function directly is equivalent to assuming that firms are fully aware of the state dependence dynamics and observe the fractions of the market loyal to each sub-brand, setting prices accordingly. The wide availability of rich scanner data makes this assumption plausible, especially for established brand manufacturers with long presence in their respective markets, as is the case in our application. In equilibrium, firms have best response pricing strategies that maximize their current and future profits conditional on the payoff relevant information captured in the state vector. Focusing on pure strategy Markov Perfect Equilibria, the associated Markov strategies are then functions mapping the current state into the continuum of possible prices, $\sigma_f : S \rightarrow R$.

3.4.1 *Firm-level profit maximization*

Firms maximize discounted total profits for all their products over an infinite horizon. The current and future payoffs of each firm at any point in the state space (e.g., a specific allocation of consumer loyalties across sub-brands) are described by the Bellman equation

$$V_f(s) = max_{p_j \geq 0,\, j \in f} \left\{ \pi_f(s, P) + \beta V_f\left[g(P, s)\right] \right\} \quad \forall s \in S \tag{11}$$

in which $g(\cdot)$ is the transition kernel in Eq. 9. The value function V_f in Eq. 11 captures all payoffs, given a specific strategy profile for firm *f*, σ_f. We use the notion of Markov Perfect Equilibrium to compute equilibrium prices in the form of pure strategies. At equilibrium each firm has an optimal strategy that prescribes the best response to all rivals, at all possible states. Denoting the strategy profiles of competitors by σ_{-f}, the optimal strategy for *f*, σ_f^*, satisfies the following Bellman equation

$$V_f(s) = max_{p_j \geq 0,\, j \in f} \left\{ \pi_f\left[s, P, \sigma^*_{-f}(s)\right] + \beta V_f\left[g(P, s, \sigma^*_{-f}(s))\right] \right\} \forall s \in S,\ \forall f \in K \tag{12}$$

3.4.2 *Sub-brand profit maximization*

With minor adjustments, the game and solution algorithm can be adapted to a setting in which prices are set to optimize product level profits, rather than firm level profits. In particular, our conception of this alternative pricing regime involves firms treating their sub-brands as isolated profit centers, so that pricing decisions for sub-brand j, for example, would not take into account the impact of product j's price on the profits accruing to any other sub-brand $k \neq j$ owned by that same firm. However, we also assume that these pricing decisions are made with full knowledge of the true impact of both parent- and sub-brand state dependence (i.e., the true demand system). The idea is that by treating the sub-brands as profit centers, the sub-brand managers would face no incentive to either 1) avoid cannibalizing the sales of other sub-brands, or 2) account for possible increases in their sales due to dynamic complementarities. It is the latter effect that can lead to higher overall profits (by softening competition), as we will demonstrate below. By comparing the optimal prices in two different scenarios, one where firms maximize sub-brand profits jointly and one where each sub-brand is its own separate profit center, we can evaluate the impact of full firm-level profit maximization on profitability.

3.4.3 Pricing equilibrium

Establishing the existence and uniqueness of a Bertrand Nash equilibrium between multiproduct firms under logit demand is a challenging theoretical problem, even in the simple static case where firms compete by maximizing current period profits. While there are existence results for a variety of special cases, they are limited to either static pricing assuming standard logit demand (without heterogeneity) or dynamic pricing by single-product firms.[5] We follow DHR (2009) in using a numerical approach to computing pure strategy equilibrium for this fully dynamic, multi-product game. Since the game may admit more than one equilibrium, we are effectively assuming that our algorithm finds the correct one. While we do employ a variety of different starting values for the algorithm, there are no known methods for finding all equilibrium and thus no guarantee that we will find the correct one.

3.5 Equilibrium computation

To study the implications of parent brand state dependence for firm pricing decisions, we solve for equilibrium steady state prices under different counterfactual assumptions regarding the role of parent brand state dependence or the underlying game structure.[6] Before obtaining steady state prices, we compute the optimal competitive policies for each firm for each unique point in the state space. Given the formulation of the game, the optimal price for each firm (or sub-brand, depending on the scenario) depends on the state of the market, namely how many consumers are loyal to the sub-brands of the respective firm at each point in time. The policy functions are infinite-dimensional functions. The value function in Eq. 11 is also of infinite dimension since it is defined uniquely for each point in the state space. To circumvent the intractable problem of solving for infinite-dimensional functions, we approximate the solution to the dynamic game specified by Eqs. 11 and 12 by discretizing the state space and interpolating the values for points outside this discrete grid. This general approach is described in Judd (1998) and has been implemented successfully by

[5]Morrow and Skerlos (2010) prove and characterize the existence of Bertrand Nash equilibrium between multi-product firms for a very general logit based demand system. Their framework imposes relatively week restrictions on the specification of utility functions and price effects. To overcome the problem that the logit-based profit functions of multi-product firms are not quasi-concave, they base their approach on fixed-point equations. Although they generalize their numerical fixed-point approach for equilibrium price computations to the Mixed logit context, they do not prove existence there. In forward looking settings, there are proofs for existence of equilibrium between single product firms under standard logit demand. A relatively common approach is to prove the quasi-concavity of the profit function (i.e. current period profits plus the value function for a given state vector) which in turn guarantees the existence of a unique profit maximizing price. Besanko et al. (2010), and DHR (2009) for a special case of their model, prove the existence of equilibrium through quasi-concavity. Unfortunately, the same approach cannot be followed in the case of multi-product firms as the profit function is no longer quasi-concave.

[6]Computing optimal prices requires knowing each firm's costs. We first assume (for purposes of conducting a variety of theoretical exercises) that costs are known to the researcher. We will later describe a method for recovering cost estimates that are subsequently used in the counterfactual exercises provided post-estimation.

DHR (2009) and others. Details regarding our implementation of the algorithm can be found in the Appendix.

3.6 Theoretical implications of PBSD

3.6.1 Symmetric brands

In this sub-section, we use numerical simulations to explore the theoretical impact of parent brand state dependence on equilibrium prices and profits. Our results are based on a simplified model that will help trace out these implications for a range of state dependence parameter values. In our streamlined model, there are just two firms and a single consumer. The first firm carries sub-brands 1 and 2, while the second firm carries sub-brands 3 and 4. The utility function is a simplified version of the one presented in Section 3.1, with the following symmetric parameter values: sub-brand intercepts $\beta_j = 1\ \forall j$, price coefficient $\beta_p = -1$ and costs $c_j = 0.5\ \forall j \in 1, .., 4$. The parameters for parent brand (β_{PB}) and sub-brand (β_{SB}) state dependence vary across different simulation scenarios. Following DHR (2009) we adjust β_j as we increase the state dependence parameters so that the size of the outside good remains constant (note that increasing the level of state dependence reduces the share of the outside good, thereby complicating the relevant comparative static). This allows us to isolate the effect of the state dependence magnitude from market size expansion and provide a ceteris paribus accounting of its impact on prices and profits.

Returning to the setup, brands (or sub-brands depending on the scenario) compete each period for the consumer's purchase while maximizing their discounted net future profits. State dependence at time t is based on the consumer's choice in the previous period (e.g. state 1 refers to sub-brand 1 being chosen the previous period) and it induces imperfect lock-in as the consumer always has positive probability of switching and choosing any of the remaining sub-brands.

In each of Tables 1, 2 and 3 we report the average transaction price ($\overline{p} = \sum_{j=1}^{J} \frac{\sigma_j(s=1) \times Pr(j|s=1)}{\sum_{k=1}^{J} Pr(k|s=1)}$) of the pricing game for state s=1. This is the average price that the consumer is expected to pay if, in the previous period, she chose sub-brand 1. Due to symmetry, the average transaction price is the same across all states. We also report the optimal pricing policy and the value function of sub-brand 1 for each possible state. Since sub-brands are symmetric in our numerical analysis, the policies and value functions of all sub-brands are identical up to the value of the state variable. Table 1 reports results for various values of parent brand state dependence (β_{PB}), but no sub-brand state dependence ($\beta_{SB} = 0$) (Case 1). Tables 2 and 3 then report results for positive values of sub-brand state dependence (Case 2 for $\beta_{SB} = 0.5$ and Case 3 for $\beta_{SB} = 1$ respectively). All three tables include two sets of results, one corresponding to firms maximizing joint profits over their sub-brands and one for sub-brand profit maximization. We note that the value function in both cases includes the profit flow from both sub-brands (i.e., we report profits for firm 1, who owns sub-brands 1 and 2).

The first conclusion we draw from this numerical analysis is that equilibrium pricing policies have a U-shaped relationship to parent brand state dependence in the fully

Table 1 Theoretical implications of PBSD - case 1: SBSD = 0

	Joint profit maximization					Sub-brand profit maximization				
Average transaction price ($\overline{p}$) and equilibrium policies (σ_1)										
PBSD	$\overline{p}$	σ_1(s = 1)	σ_1(s = 2)	σ_1(s = 3)	σ_1(s = 4)	$\overline{p}$	σ_1(s = 1)	σ_1(s = 2)	σ_1 (s = 3)	σ_1(s = 4)
0.0	1.94	1.94	1.94	1.94	1.94	1.70	1.70	1.70	1.70	1.70
0.5	1.77	1.91	1.91	1.61	1.61	1.62	1.69	1.69	1.53	1.53
1.0	1.62	1.88	1.88	1.25	1.25	1.56	1.68	1.68	1.33	1.33
1.5	1.49	1.84	1.84	0.88	0.88	1.51	1.68	1.68	1.09	1.09
2.0	1.37	1.81	1.81	0.50	0.50	1.47	1.67	1.67	0.83	0.83
2.5	1.27	1.78	1.78	0.11	0.11	1.44	1.67	1.67	0.53	0.53
3.0	1.39	1.81	1.81	0	0	1.32	1.67	1.67	0	0
4.0	1.68	1.89	1.89	0	0	1.53	1.68	1.68	0	0
5.0	1.85	1.93	1.93	0	0	1.64	1.70	1.70	0	0
Net discounted profits (value function)										
PBSD		V_1(s = 1)	V_1(s = 2)	V_1(s = 3)	V_1(s = 4)		V_1(s = 1)	V_1(s = 2)	V_1(s = 3)	V_1(s = 4)
0.0		219.4	219.4	219.4	219.4		199.5	199.5	199.5	199.5
0.5		205.9	205.9	205.6	205.6		191.7	191.7	191.4	191.4
1.0		188.9	188.9	188.3	188.3		184.1	184.1	183.4	183.4
1.5		171.2	171.2	170.3	170.3		177.2	177.2	176.0	176.0
2.0		153.7	153.7	152.4	152.4		171.1	171.1	169.4	169.4
2.5		138.0	138.0	136.3	136.3		166.0	166.0	163.8	163.8
3.0		153.9	153.9	151.5	151.5		147.3	147.3	144.8	144.8
4.0		193.7	193.7	187.8	187.8		179.7	179.7	173.5	173.5
5.0		216.6	216.6	201.4	201.4		199.2	199.2	183.3	183.3

dynamic pricing model. Prices initially decrease with the magnitude of PBSD, reach a minimum, and then start increasing thereafter. This result holds for all three values of SBSD considered, and matches the findings of DHR (2009) (for product-level state dependence). The intuition is that, for moderate levels of PBSD, firms price more aggressively to compete for future business across their portfolio of sub-brands. As shown in Fig. 1, the U-shaped pattern of equilibrium pricing policies is driven by dynamic incentives to invest in future sales, as it does not occur in the equivalent myopic game. Figure 2 depicts the corresponding difference in profits between the forward looking and myopic case. Similar to prices, profits also decrease for moderate levels of PBSD, due to the enhanced dynamic competition. In the forward looking case, as the magnitude of PBSD increases brands either charge relatively higher prices to a loyal consumer (e.g. the policy at state 1 in the top left panel of Table 1) or offer deep discounts to attract a non-loyal consumer (e.g. the policy at states 3 or 4). For high values of PBSD, the loyal consumer is quite unlikely to switch away and the firm must price very aggressively to attract their business. While these

Table 2 Theoretical implications of PBSD - case 2: SBSD = 0.5

	Joint profit maximization					Sub-brand profit maximization				
Average transaction price ($\overline{p}$) and equilibrium policies (σ_1)										
PBSD	$\overline{p}$	σ_1(s = 1)	σ_1(s = 2)	σ_1(s = 3)	σ_1(s = 4)	$\overline{p}$	σ_1(s = 1)	σ_1(s = 2)	σ_1 (s = 3)	σ_1(s = 4)
0.0	1.85	1.92	1.92	1.76	1.76	1.58	1.69	1.53	1.53	1.53
0.5	1.69	1.89	1.89	1.41	1.41	1.50	1.68	1.50	1.34	1.34
1.0	1.55	1.86	1.86	1.05	1.05	1.44	1.67	1.47	1.12	1.12
1.5	1.42	1.82	1.82	0.67	0.67	1.39	1.66	1.45	0.87	0.87
2.0	1.32	1.79	1.79	0.28	0.28	1.35	1.66	1.43	0.59	0.59
2.5	1.30	1.78	1.78	0	0	1.32	1.65	1.42	0.29	0.29
3.0	1.49	1.84	1.84	0	0	1.31	1.65	1.41	0	0
4.0	1.74	1.90	1.90	0	0	1.46	1.67	1.41	0	0
5.0	1.88	1.94	1.94	0	0	1.53	1.69	1.41	0	0
Net discounted profits (value function)										
PBSD		V_1(s = 1)	V_1(s = 2)	V_1(s = 3)	V_1(s = 4)		V_1(s = 1)	V_1(s = 2)	V_1(s = 3)	V_1(s = 4)
0.0		212.4	212.4	212.2	212.2		186.3	186.3	186.1	186.1
0.5		196.7	196.7	196.3	196.3		176.9	176.9	176.4	176.4
1.0		179.1	179.1	178.3	178.3		168.5	168.5	167.6	167.6
1.5		161.3	161.3	160.2	160.2		161.2	161.2	159.8	159.8
2.0		144.7	144.7	143.2	143.2		155.3	155.3	153.4	153.4
2.5		141.9	141.9	139.9	139.9		150.5	150.5	148.0	148.0
3.0		167.5	167.5	164.4	164.4		148.3	148.3	145.1	145.1
4.0		201.2	201.2	193.5	193.5		173.1	173.1	165.0	165.0
5.0		221.9	221.9	202.2	202.2		189.4	189.4	168.6	168.6

implications closely mirror the results of DHR (2009), there are additional (novel) implications regarding the impact of multi-product pricing and firm asymmetries that we turn to next.

The second robust pattern that emerges from this theoretical analysis is that PBSD can actually decrease, and in some cases reverse, the benefit of centralized pricing. Starting with Table 1 we see that, in most cases, the value function for state 1 (V_1(s = 1)) is greater under joint profit maximization, as would generally be expected. However, at moderate levels of PBSD, the opposite is true. Figure 3 summarizes the difference in net discounted profits under joint and sub-brand profit maximization for the whole range of parameters we analyze. Note that there is a sizable range of values for the PBSD parameter where forward-looking profits under joint profit maximization are actually lower than forward looking profits under sub-brand profit maximization. This is a striking result in which centralized pricing leads to lower profits than de-centralized pricing. Our interpretation of this outcome is that, for this range of parameter values, the future benefit of locking in consumers outweighs the harvest incentive (under joint profit maximization) and leads firms to compete harder

Table 3 Theoretical implications of PBSD - case 3: SBSD = 1

	Joint profit maximization					Sub-brand profit maximization				
Average transaction price ($\overline{p}$) and equilibrium policies (σ_1)										
PBSD	$\overline{p}$	σ_1(s = 1)	σ_1(s = 2)	σ_1(s = 3)	σ_1(s = 4)	$\overline{p}$	σ_1(s = 1)	σ_1(s = 2)	σ_1 (s = 3)	σ_1(s = 4)
0.0	1.74	1.90	1.90	1.52	1.52	1.45	1.67	1.32	1.32	1.32
0.5	1.59	1.87	1.87	1.17	1.17	1.38	1.66	1.26	1.11	1.11
1.0	1.46	1.83	1.83	0.79	0.79	1.33	1.65	1.22	0.87	0.87
1.5	1.35	1.80	1.80	0.41	0.41	1.28	1.65	1.19	0.61	0.61
2.0	1.25	1.77	1.77	0.02	0.02	1.25	1.64	1.16	0.33	0.33
2.5	1.43	1.82	1.82	0	0	1.21	1.64	1.14	0	0
3.0	1.59	1.86	1.86	0	0	1.29	1.65	1.12	0	0
4.0	1.80	1.92	1.92	0	0	1.38	1.66	1.10	0	0
5.0	1.91	1.96	1.96	0	0	1.42	1.68	1.10	0	0
Net discounted profits (value function)										
PBSD		V_1(s = 1)	V_1(s = 2)	V_1(s = 3)	V_1(s = 4)		V_1(s = 1)	V_1(s = 2)	V_1(s = 3)	V_1(s = 4)
0.0		202.1	202.1	201.7	201.7		170.7	170.7	170.3	170.3
0.5		184.9	184.9	184.2	184.2		160.4	160.4	159.7	159.7
1.0		166.9	166.9	165.9	165.9		151.7	151.7	150.6	150.6
1.5		149.9	149.9	148.5	148.5		144.7	144.7	143	143
2.0		134.4	134.4	132.7	132.7		139.1	139.1	136.9	136.9
2.5		160.0	160.0	157.3	157.3		133.4	133.4	130.7	130.7
3.0		181.3	181.3	177.1	177.1		146.8	146.8	142.6	142.6
4.0		209.0	209.0	198.4	198.4		164.0	164.0	153.2	153.2
5.0		228.5	228.5	201.6	201.6		178.6	178.6	151.2	151.2

for the consumer's business today. Notice that, as we show in Fig. 4, joint profit maximization always generates higher profits when firms are not forward looking. However, with dynamic incentives, the increased probability of future profit streams across the sub-brand portfolio leads firms to lower prices today to invest in their clientele tomorrow. By muting these dynamic incentives, firms can actually earn higher profits by operating their sub-brands as separate profits centers, even after accounting for the increased cannibalization of contemporaneous sales. It is important to note that this is an equilibrium effect. When only one of the two firms employs sub-brand profit maximization (while the other maximizes joint profits), that firm earns a lower PDV over the whole range of parameter values.[7] Thus, it is competition between sub-brand portfolios that produces lower prices and profits. Note that the levels of PBSD where we observe this behavior corresponds to the range where the average transaction price is lowest overall. We see a similar result in Table 2 where, at PBSD

[7]These results are available from the authors upon request.

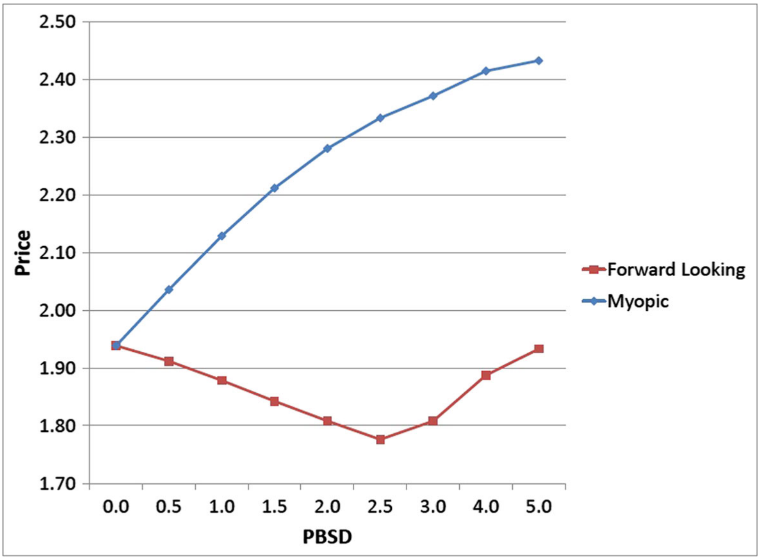

Fig. 1 Equilibrium policy: forward looking vs myopic

= 2.5, the value function for state 1 under joint profit maximization is 144.7, while under sub-brand profit maximization it is 155.3. The difference in profits between sub-brand profit maximization and joint profit maximization is smaller in this case but still positive; centralized pricing leads to stronger competition and lower profits due to parent brand state dependence.

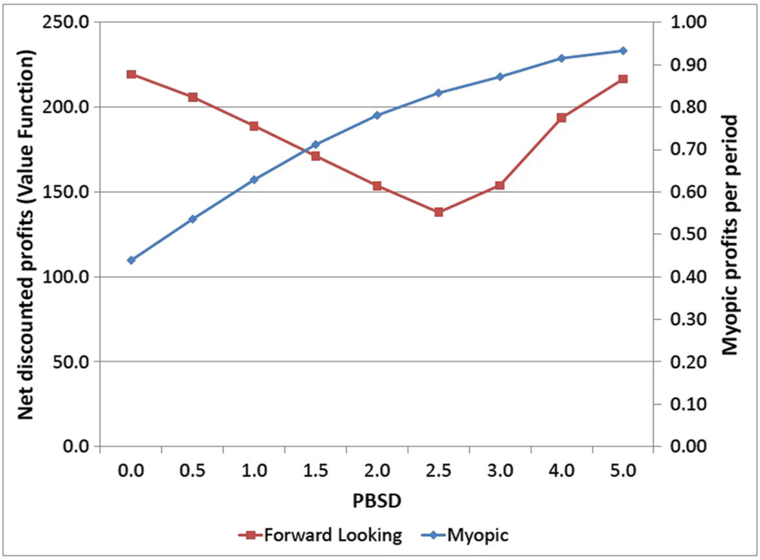

Fig. 2 Equilibrium profits: forward looking vs myopic

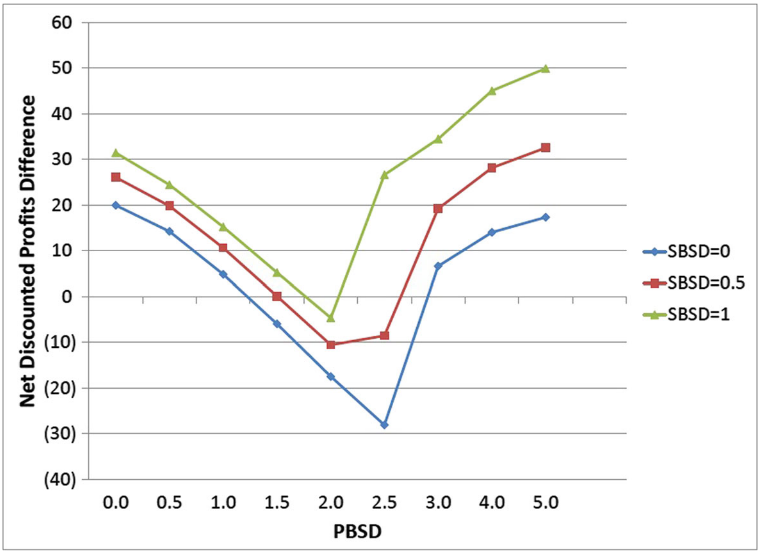

Fig. 3 Benefit from joint profit maximization - forward looking

Interestingly, the negative effect of PBSD on the benefits of joint profit maximization is mitigated by greater values of sub-brand state dependence, for which these countervailing pricing effects do not exist. As Fig. 3 shows, when SBSD is higher, joint profit maximization yields higher profits for a wider range of PBSD values. In other words, when sub-brand state dependence is high, the incentive to avoid cannibalizing the other sub-brands overwhelms the increased competitive pressure

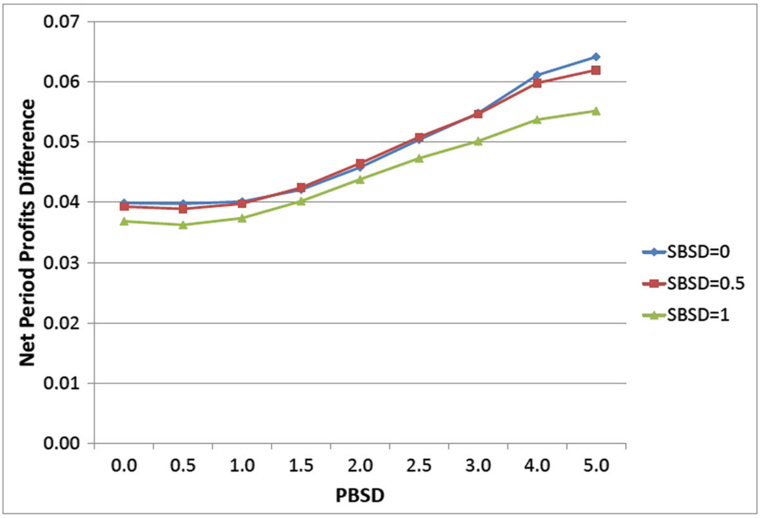

Fig. 4 Benefit from joint profit maximization - myopic

between firms that PBSD induces. This is reflected more clearly in the equilibrium policies of sub-brand 1 when the consumer is loyal to sub-brand 2 (which belongs to the same parent), $\sigma_1(\mathrm{s} = 2)$. The difference in these policies between joint profit maximization and sub-brand profit maximization, shown in Fig. 5, tends to increase with SBSD. As SBSD increases, sub-brand 1 prices more and more aggressively to win back the consumer and maximize its own profits. Centralized pricing internalizes this effect, leading to higher benefits compared to de-centralized pricing. In light of this intuitive tension, we conclude that it is important to measure existing levels of both PBSD and SBSD when evaluating centralized pricing.

The numerical analysis of the simplified model also yields insights into the relative pricing of sub-brands and how this is affected by PBSD and centralized pricing. Starting with Table 1, we see that under both joint and sub-brand profit maximization, the pricing policy of sub-brand 1 prescribes a very similar price for states 1 and 2. With only parent brand state dependence and under joint profit maximization, firm 1 views loyalty to sub-brand 1 or 2 as inter-changeable, and sets prices that mitigate within-brand competition. Optimal prices of sub-brand 1 for states 1 and 2 are also very similar under sub-brand profit maximization. Since we assume that firms have full information on the state variable and the demand system, the manager pricing sub-brand 1 knows that, in this game, state 1 is as good as state 2 and prices accordingly. However, when sub-brand state dependence is introduced in addition to PBSD, we see a different picture (Tables 2 and 3). Under joint profit maximization, firms still price their sub-brands quite similarly when the consumer is loyal to one of their products. For example, firm 1 sets almost exactly the same price in states 1 and 2, reflecting the internalizing of cross-product cannibalization effects. It also prices similarly when the consumer is loyal to the rival firm, irrespective of which rival

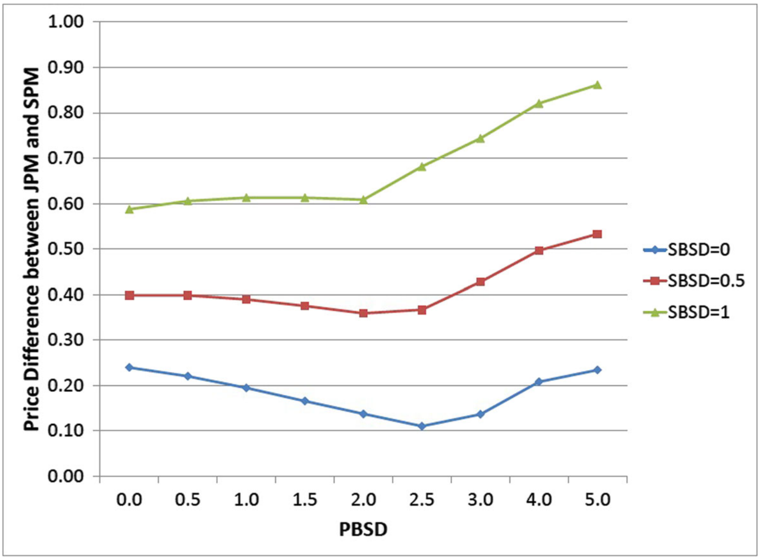

Fig. 5 Difference in $\sigma_1(\mathrm{s} = 2)$ between joint profit maximization and sub-brand profit maximization

sub-brand the consumer purchased last. To be clear, this result is partly driven by symmetry, but nevertheless contrasts in an informative way with the corresponding pricing policies under sub-brand profit maximization. The same symmetrical sub-brands have markedly different pricing policies depending on which one is favored by the consumer's state. Under state 2, where the consumer is loyal to sub-brand 2, sub-brand 1 charges a lower price compared to state 1 albeit still higher than the rival firm's sub-brands which do not benefit from PBSD. In this sense, we see that PBSD may lead firms more to price more uniformly compared to SBSD, to the extent this is supported by sufficient symmetry in sub-brand preferences. Intuitively, when consumers tend to repeat their choices within the brand portfolio and there is no state dependence to the sub-brand, even self-interested sub-brand managers will not compete with sub-brands of the same parent but only with the rival parent brand. From a forward looking perspective, consumers in this case would be equally likely to buy sub-brand 1 next time irrespective of whether they chose sub-brand 1 or 2 before. What is important therefore is a good understanding of the loyalty dynamics across sub-brands, even if each sub-brand maximizes its own profits.

3.6.2 Asymmetric brands

To expand the applicability of our simplified model to more realistic scenarios, we also consider a setting where the sub-brands of one firm have stronger preferences than those of its rival. We keep most parameter values the same as in the symmetric case, changing only the sub-brand intercepts, which are now set at $\beta_j = 1.25$ for the "strong preferences" sub-brands of firm 1 and as $\beta_j = 0.75$ for the "weak preferences" sub-brands of firm 2. This configuration allows us to explore how the results change when tastes differ. Figure 6 shows equilibrium net discounted profits (i.e. value functions) for the two firms at different values of PBSD. For this particular example, we set the SBSD coefficient to 0.5 and compute the profits for state 1 (without loss of generality, since the pattern looks very similar for the other states as well). Interestingly, while the combined PDVs of the two firms decrease as the magnitude of PBSD increases, the profits of the strong brand actually increase at the expense of the weak brand; PBSD leads to increasing dominance. Intuitively, when brands have asymmetric preferences, they continue to compete more aggressively to win future sales as PBSD increases, but the stronger brand benefits while the weaker brand is hurt, a result that extends across most moderate levels of PBSD. Figure 7, which shows the share of profits earned by each firm, reveals that, for most of the range of PBSD coefficient values, the strong brand commands a larger share of total profits as PBSD increases. These additional results regarding asymmetric brands paint a more complete picture of the dynamics behind PBSD and its impact on forward looking profits.[8] On the one hand, they suggest that forward looking market leaders may have incentives to foster a certain level of parent brand loyalty, even if that would make the

[8]With regards to the effect of PBSD on the benefits of joint profit maximization, we found that brand asymmetry did not add any new insights to the findings reported above, so we do not repeat a similar discussion here.

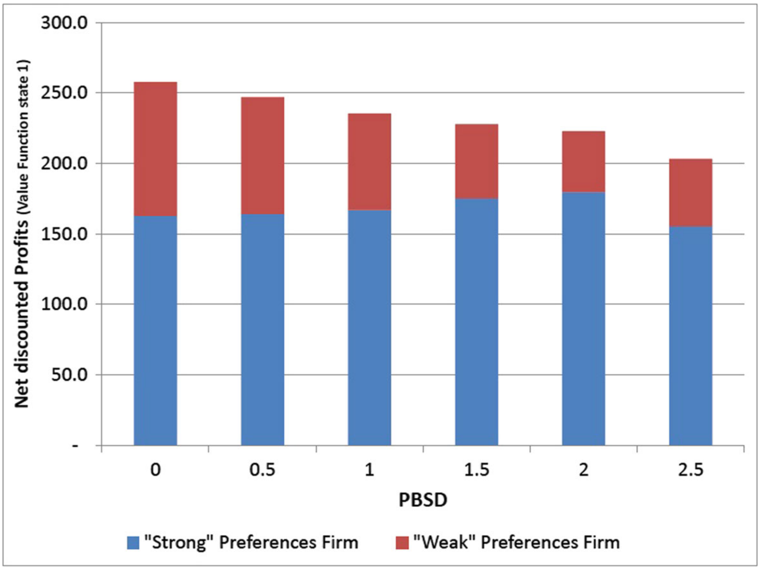

Fig. 6 Forward looking equilibrium profits for asymmetric brands

overall market more competitive, as they stand to benefit relative to their rivals. On the other hand, it indicates that the full impact of state dependence depends on subtle features of the market environment, rendering the observed implications an empirical question, which we turn to next.

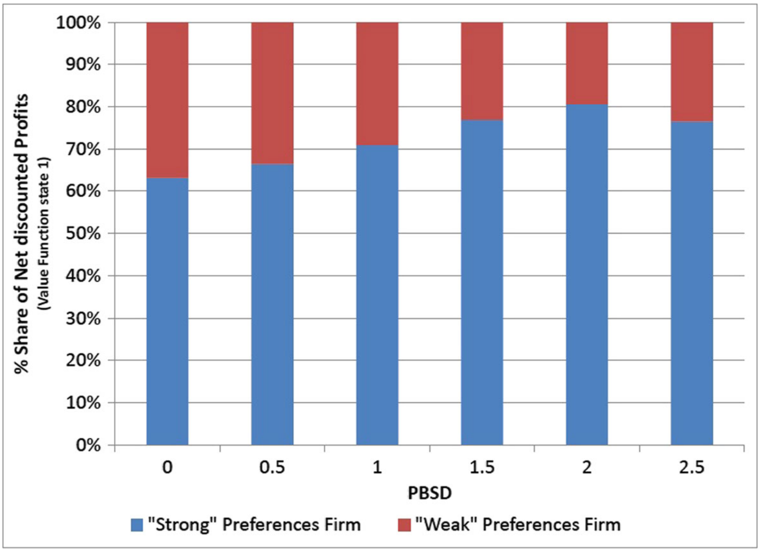

Fig. 7 Forward looking equilibrium profit shares for asymmetric brands

3.7 Cost estimation

Solving the equilibrium pricing game for the full empirical model (so as to conduct counterfactual analyses regarding changes in the underlying environment) requires knowing product level marginal costs. Since we do not observe costs (or even wholesale prices), we instead rely on the pricing policy implied by our framework (together with the estimated demand system) to infer their values. Our estimation approach is based on the method developed by Bajari et al. (2007) (BBL), adapted to the dynamic pricing model outlined in Section 3.4. Specifically, we recover costs that rationalize the prices observed in the data, *assuming* that firms are behaving optimally. Since joint profit maximization is a priori expected to generate superior profitability, we assume that this is how they actually behave (i.e. maximizing profits across all sub-brands, while taking into account both sub-brand and parent brand state dependence). The implied costs are then used to evaluate alternative (counterfactual) pricing strategies, such as decentralized pricing, by re-solving the full dynamic pricing game.

The BBL approach involves two steps. In the first step, we recover the firms' policy functions directly from the their observed actions and use these policies to forward simulate value functions for each firm (as a function of the unknown cost parameters). Along with the value functions corresponding to the firms' observed policies, we also compute parallel value functions corresponding to "counterfactual" or perturbed policies. In the second step, we use a minimum distance estimator to recover the structural parameters that rationalize the observed "optimal" behavior, by requiring that the observed policies outperform the perturbed ones. Neither step involves solving for equilibria, but does rely on a maintained assumption that the same equilibrium is played throughout the data frame. Details regarding the full estimation procedure are provided in the Appendix.

4 Empirical application

4.1 Data

Our empirical model of the Yogurt category uses scanner data drawn from the widely available IRI Marketing Dataset (Bronnenberg et al. 2008). The sample used for analysis includes household level shopping trips for groceries, purchases of well known yogurt products of the most popular size in the category (6 oz), and the respective prices for each sub-brand. The time period covered includes years 2002 and 2003 for a total length of 104 weeks. Additional purchase histories from year 2001 are used to construct the initial conditions of brand and product loyalty for each household. To ensure proper tracking of choices over time, we exclude households who do not satisfy the IRI criteria for regular reporting of information. The choice set used for the estimation covers five parent brands with twelve sub-brand lines in total, as shown below.

Sub-brands of the same brand name in this category are moderately differentiated in packaging, labelling, shelf space and ingredients. In all cases, the parent brand

Table 4 Marketplace descriptive statistics

Parent brand	Sub-brand	Total Flavors	New Flavors	Price/lb Average	Price/lb Std Dev.	Share
Dannon	Dannon creamy fruit blends	6	6	1.581	0.18	0.01
Dannon	Dannon fruit on the bottom	9	9	1.712	0.32	0.04
Dannon	Dannon light n fit	12	12	1.790	0.24	0.07
Dannon	Dannon light n fit creamy	6	6	1.674	0.10	0.03
Kemps	Kemps classic	9	0	1.261	0.12	0.04
Kemps	Kemps free	11	0	1.361	0.20	0.05
Old Home	Old home	8	1	1.383	0.19	0.01
Old Home	Old home 100 calories	9	1	1.341	0.18	0.04
Wells BB	Wells blue bunny lite 85	16	2	1.403	0.11	0.14
Yoplait	Yoplait light	15	1	1.671	0.16	0.23
Yoplait	Yoplait original	24	1	1.671	0.14	0.24
Yoplait	Yoplait thick and creamy	15	2	1.676	0.13	0.09
Total		140	40	1.551	0.24	1

name is prominent and clearly visible on the package of each sub-brand. Hence, consumers are able to identify or recall both the parent brand and the specific sub-brand name when shopping. In some cases the product lines of the same brand have very similar prices, like *Yoplait Light* and *Yoplait Original*. On the other hand, some brands use price as a differentiator between their various sub-brands; for example *Kemps Classic* is priced *10* cents lower than *Kemps Free* per pound. Table 4 reports the number of different flavor varieties per sub-brand, average price, standard deviation of price and category market share. *Yoplait* and *Wells Blue Bunny* carry the greatest number of flavors, consistent with their share leadership positions, which ensure ample shelf space and a large pool of loyal customers. Prices vary across brands, sub-brands and weeks/stores within any given sub-brand. *Kemps* is the lowest priced brand of the sample and carries the fewest flavors while *Dannon* has the two most expensive sub-brands with a moderate number of flavors but the largest number of sub-brands. The product category we are modeling is mature and relatively stable.[9]

Table 5 reports quartile, median and mean statistics for several metrics across households. The average number of shopping trips in our sample is 88.3, while the average number of trips with Yogurt purchases is 23.1. Only about 3.4 of those trips involve the purchase of multiple sub-brands during a single trip, mitigating somewhat

[9]Excluding Dannon, the rest of the brands collectively only had eight new flavor introductions over the two years covered by our sample. While Dannon did introduce several new varieties at this time (to build its position with new SKUs and refresh existing labels), it is important to note that Dannon enjoyed wide distribution in the US for many years prior to our sample period and is a well-established brand. Moreover, the new varieties mainly involved a re-positioning of *0.5* pound SKUs to match the *0.375* pound size of the market leader Yoplait.

Table 5 Choice summary statistics

	1st quartile	Median	Mean	3rd quartile
Shopping trips	82	91	88.3	97
Trips with purchase	13	20	23.1	29
Trips with multiple sub-brand purchase	1	2	3.4	4
Number of brands bought	2	3	3.2	4
Number of sub-brands bought	3	5	5.2	7
Switches within brand	3	6	7.9	11
Switches to other brand	1	4	6	9

concerns regarding variety seeking. To simplify estimation, we treat these multiple sub-brand purchases as separate observations, a practice typical in the scanner data literature. In particular, when a household is observed to buy more than one sub-brand on a given week, we track the household's state dependence to all purchased sub-brands (e.g., loyal to *Yoplait Original* and to *Dannon Light N Fit*). Rows four and five of the table report the number of different variants and different parent brands bought by each household in the sample; the average number of sub-brands a household purchased is 5.2 while the average number of brands is 3.2. This is evidence that there is switching observed in the data, which is clearly needed to identify switching costs. Rows three and four show a more specific pattern, namely that the number of sub-brand switches to a variant of the same parent name (*Switches within brand*) is higher than the number of switches to a variant of a different parent brand name (*Switches to other brand*). This pattern holds across the range of the household distribution (i.e. 1^{st} quartile, median, 3^{rd} quartile) with the respective averages being 7.9 versus 6.

5 Results & discussion

5.1 Demand estimation

We begin the discussion of our empirical results by reporting DIC values for model selection. In the first three rows of Table 6, we evaluate the importance of including state dependence to both the sub-brand and parent brand. Based on the fit measures, adding sub-brand and type state dependence improves the fit of the demand model to the data. This is evident by the DIC values decreasing from the first row of Table 6 to the second. Moreover, parent brand state dependence adds even more explanatory power, as DIC values decrease further in the third row of the table. Together, this is evidence that: i) consumers do tend to repeat their sub-brand (or type) choices over time (controlling for price effects and intrinsic preferences) and ii) they also tend to become loyal to parent brand names in cases where they switch their product line choice. In the next few rows of Table 6, we report results from model specifications where consumer heterogeneity is a mixture of Normal components. The preferred

Table 6 Model selection results

Mixtures	State dependence	DIC
1 component	None	68117.6
1 component	No brand SD (only sub-brand and type)	67215.4
1 component	Full state dependence	67114.3
2 components	Full state dependence	67157.9
3 components	No SD at all	67954.3
3 components	No brand SD (only sub-brand)	67201.4
3 components	Full state dependence	**67113.7**
4 components	Full state dependence	67150.0

heterogeneity specification, based on DIC, includes three Normal components (bolded in table). Furthermore, in the case of three components, incorporating sub-brand, type and parent brand state dependence improves the fit of the model. In the last row we explore whether adding more components to the heterogeneity mixture provides an even better fit, which it does not. Based on these results, we use average posterior mean household estimates from the three component mixture specification to estimate costs and compute market equilibrium and counterfactual scenarios.

The demand estimates used for the cost and the pricing model are reported in Table 7, which presents mean posterior estimates across the households in the estimation sample. The sub-brand specific preferences reflect the popularity of each

Table 7 Demand results

Sub-Brand	Posterior Estimates	
	Mean	[5, 95] pctl
Dannon creamy fruit blends	−2.896	[−7.99, 2.98]
Dannon fruit on the bottom	−2.690	[−8.17, 3.86]
Dannon light n fit	−1.083	[−6.42, 4.86]
Dannon light n fit creamy	−2.126	[−7.29, 3.26]
Kemps classic	−3.460	[−8.33, 2.03]
Kemps free	−2.860	[−7.82, 2.51]
Old home	−4.255	[−9.05, 1.01]
Old home 100 calories	−2.790	[−7.86, 2.47]
Wells blue bunny lite 85	−1.781	[−6.10, 2.88]
Yoplait light	−0.685	[−5.90, 4.19]
Yoplait original	−0.628	[−6.04, 4.52]
Yoplait thick and creamy	−2.077	[−7.33, 3.26]
Price coefficient	−2.186	[−5.20, 0.41]
State dependence		
Parent brand	0.286	[−0.56, 1.15]
Product line	0.729	[−0.27, 1.85]
Reg./ light type	0.313	[−0.45, 1.09]

Table 8 Identification testing on shuffled sample

Mixtures	State dependence	DIC
3 components	No SD at all	**68039.6**
3 components	No PBSD (only sub-brand and type)	68269.2
3 components	Only PBSD (no sub-brand and type)	68193.8
3 components	Full state dependence	68293.6

alternative. *Yoplait Original* and *Yoplait Light* are the market share leaders in this category and have the least negative taste coefficients.[10] The estimates of state dependence reveal that the effect of lagged purchase on utility due to purchasing the same sub-brand is a little over two and a half times that of the parent brand; the corresponding estimates are 0.729 and 0.286 respectively. State dependence based on fat content has comparable size to PBSD. The economic importance of these state dependence parameters is described below.

5.1.1 Robustness to confounds

As noted earlier, due to its inherently residual nature, structural state dependence is vulnerable to a potential confound with unobserved heterogeneity, hence the use of the mixture of Normals in estimating the demand system. To assess robustness to a misspecified heterogeneity distribution, we ran empirical tests similar to those of Dubé et al. (2010), which are based on the idea that structural state dependence depends on the actual sequence of purchases (i.e. what you actually purchased last), while unobserved heterogeneity should be more persistent (i.e. what you tend to purchase a lot). In particular, we randomly shuffled the order of observations within each household in our sample and then estimated versions of our model without state dependence, with only sub-brand state dependence and finally including parent-brand state dependence as well. The various model specifications are compared based on DIC values.

If our estimated measures of inertia were spurious and only captured unobserved heterogeneity, when we add them to the model and estimate on the shuffled observation data set, we should still find that they add predictive power to the model (e.g. DIC decreases). In contrast, we find that adding sub-brand state dependence to the model when observations are shuffled does not add significant information. Similarly, parent-brand state dependence also does not improve the fit when the purchase order is shuffled. As shown in Table 8, the preferred model in this series of tests is the one without any state dependence at all (bolded in table). This is strong evidence that what we estimate in our model with the correct order of purchases is structural inertia and not masked heterogeneity.

[10]Preference estimates are negative because they are estimated against the normalized outside option, which has the greatest share in the sample since most consumers do not purchase yogurt every time they go to a grocery store.

Table 9 Alternative explanation test - model selection

Mixtures	Test specification	DIC
3 components	Add brand experience only	66508.8
3 components	Add brand experience and interaction with state dependence	66568.0

A second potential confound, also noted and addressed in Dubé et al. (2010), concerns the role of consumer learning. In particular, if consumers are initially unsure about the match value of products (and engage in learning behaviors to identify their best match), then their choice histories might exhibit inertia patterns that are similar to state dependence. We note first that the category considered here has well established brands and does not involve large purchase quantities (or highly storable products). This makes it less likely that consumers are either learning about match values or are otherwise forward looking in their everyday decision-making process.

However, to further test whether our model captures state dependence as opposed to learning, we use a second test developed by Dubé et al. (2010). In particular, we re-estimate our model controlling for consumers' experience and how it interacts with state dependence. If our estimates reflect learning rather that structural state dependence, we should find that the interaction of cumulative brand purchases with sub-brand state dependence adds information to the model (increases explanatory power). In particular, as consumers gain more experience with the sub-brands they are buying, their state dependence should decrease to zero. The results, presented in Table 9, reveal that adding the interaction of brand experience with state dependence does not add meaningful information to the model. The DIC value of the model with the interaction (M2) is much higher than the one for the model without the interaction (M1).[11] This shows that inertia exists even controlling for consumers' cumulative experience, providing additional justification for our focus on state dependence.

5.2 Pricing implications

5.2.1 Cost estimates

Before turning to our counterfactual exercises, we first report the results from the cost estimation procedure laid out in Section 3.7. The results from this exercise are reported in the right side columns of Table 10. For comparison, we also report cost estimates from a simpler, static pricing-based procedure that instead implements a single period Bertrand pricing game evaluated at the average sub-brand prices

[11] We note that, similar to the approach used by Dubé et al. (2010), we only test for the interaction term by comparing to a model specification which includes brand experience but not the interaction. Comparing to our full base model would not be appropriate since the models evaluated for this sub-section contain additional information.

Table 10 Cost estimates

Sub-Brand	Static Game	Dynamic Game	
	Est.	Est.	90% CI
Dannon creamy fruit blends	1.117	1.175	[1.157, 1.204]
Dannon fruit on the bottom	1.256	1.243	[1.217, 1.267]
Dannon light n fit	1.325	1.338	[1.322, 1.385]
Dannon light n fit creamy	1.210	1.322	[1.292, 1.347]
Kemps classic	0.797	0.942	[0.879, 0.962]
Kemps free	0.904	1.052	[1.022, 1.083]
Old home	0.933	0.953	[0.924, 0.986]
Old home 100 calories	0.884	0.853	[0.839, 0.876]
Wells blue bunny lite 85	0.937	0.995	[0.982, 1.022]
Yoplait light	1.190	1.296	[1.275, 1.310]
Yoplait original	1.186	1.274	[1.228, 1.293]
Yoplait thick and creamy	1.194	1.281	[1.235, 1.330]

observed in the data (and the respective steady state that corresponds to them).[12] We note first that the cost estimates exhibit sensible patterns: the pricier and higher quality brands have higher production costs. Moreover, for 10 out of 12 sub-brands, the cost estimates from the dynamic game are higher than those of the static version, the only two exceptions being *Dannon Fruit on the Bottom* and *Old Home 100 Calories*.[13] Apart from these two exceptions, the dynamic cost estimates are higher than those implied by a static model. The difference is largest for Kemps (more than 15%) and Yoplait (more than 7%). We note that this result already suggests that, in our empirical application, the 'invest' motive outweighs the 'harvest' motive. This stems from the duality in the problems of obtaining costs given prices or obtaining prices given costs. If forward looking prices under state dependent demand are lower conditional on costs, it follows that costs from the same model will be higher conditional on prices. For the rest of the analysis we use the cost estimates corresponding to the dynamic game, in accordance with the forward looking view we take for the pricing problem.

[12]Note that, given our state dependent model, we always need to condition the firm's profit function to a state vector. While the forward looking costs are estimated based on a random sample of observed states, the myopic costs are only calibrated based on one state.

[13]The myopic cost estimate of *Dannon Fruit on the Bottom* is not statistically different from the forward looking one. The case of *Old Home 100 Calories* is different and we attribute its relatively low dynamic cost estimate to its strong preference coefficient in the demand model. For example, if we look at the average price of *Old Home 100 Calories*, it is only a little lower from that of *Old Home* (1.34 vs 1.38) while it's mean posterior preference parameter is much higher (-2.8 vs -4.3). The cost model infers that, since *Old Home 100 Calories* is priced comparably to *Old Home* despite having tastes so much in its favor, it must have a much lower cost. Note that, while the myopic estimate for the *100 Calories* sub-brand is also lower, the difference is more pronounced in the forward looking case.

5.2.2 *The impact of PBSD levels*

Turning to our counterfactual exercises, we first examine the equilibrium impact on prices and profits of changing the levels of PBSD in the now real-world context of our application. In particular, Table 11 contains steady state prices and per period profits for all the firms in the sample, evaluated at various levels of PBSD, ranging from zero to five times the estimated level. The importance of brand asymmetries is clear. For every brand but *Yoplait* (the dominant player), steady state prices steadily decrease as PBSD increases. For *Yoplait*, on the other hand, prices follow the U-shaped pattern documented earlier, though it is most pronounced for their smallest, and most horizontally differentiated sub-brand, *Thick and Creamy*. The other two sub-brands exhibit only moderate price changes. Turning next to per period profits, these are also uniformly decreasing for all brands but *Yoplait*, but steadily increasing for *Yoplait*, especially for the highest levels of PBSD. Clearly, increased PBSD amplifies the strength of the strongest brand. Our interpretation is in line with the theoretical exercises presented earlier. A strong investment incentive leads firms to lower prices to attract consumers and retain loyals. However, for the dominant brand, the harvest incentive eventually dominates and the retention effect is muted. Moreover,

Table 11 Effect of PBSD magnitude on prices and profits

PBSD scale factor	0	0.5	1	2	3	4
Steady state prices						
Dannon creamy fruit blends	1.648	1.635	1.622	1.595	1.565	1.535
Dannon fruit on the bottom	1.715	1.703	1.69	1.664	1.636	1.609
Dannon light n fit	1.760	1.747	1.733	1.705	1.669	1.625
Dannon light n fit creamy	1.775	1.763	1.751	1.727	1.70	1.674
Kemps classic	1.396	1.388	1.38	1.365	1.353	1.341
Kemps free	1.500	1.491	1.483	1.468	1.454	1.442
Old home	1.421	1.412	1.403	1.387	1.372	1.355
Old home 100 Calorie	1.289	1.28	1.271	1.253	1.236	1.219
Wells blue bunny lite 85	1.407	1.396	1.386	1.362	1.329	1.27
Yoplait light	1.727	1.719	1.713	1.71	1.713	1.718
Yoplait original	1.686	1.677	1.671	1.668	1.672	1.676
Yoplait thick and creamy	1.760	1.756	1.754	1.763	1.795	1.85
Steady state per period profits						
Dannon	799.7	779.9	754.1	683	586.4	473.5
Kemps	246.8	220.4	192.3	134.2	81.7	43.6
Old home	295.9	266.6	235.1	168.8	106.8	59.5
Wells blue bunny	494.6	477.9	456.5	397.7	318.5	228.4
Yoplait	1675.2	1776.8	1895	2181.3	2510.2	2823.8

at all levels of PBSD considered here, the dominant firm continues to benefit from increasing levels of PBSD.

5.2.3 The role of umbrella branding

For the next set of counterfactuals, we dig deeper into the benefits and costs of umbrella branding. In particular, we perform two experiments aimed at identifying the impact of PBSD on a individual firm's pricing strategy. Given the results of the previous experiments, we focus on the market leader, *Yoplait*, paying particular attention to the role of it's *Thick and Creamy* (*T&C*) sub-brand. In the first experiment, we eliminate the dynamic effects from *Yoplait Thick and Creamy* to *Yoplait Original* and *Yoplait Light*, and vice versa. This scenario is equivalent to a situation in which the *T&C* sub-brand is not co-branded under the *Yoplait* umbrella, but rather under a different name. Equivalently, it represents a setting in which consumers do not perceive such a connection, and exhibit no loyalty to the parent company. Equilibrium prices and per period profits in this counterfactual scenario (C2) are compared with prices and profits in a baseline scenario (C1) where the branding structure of *Yoplait*, alongside every other firm, does not change. The difference between the two scenarios is that, in the counterfactual, demand for *T&C* does not increase with loyalty to the other *Yoplait* sub-brands, and vice versa.

The results of these branding counterfactuals are reported in Table 12. The reader should focus attention on the bottom three rows, as the equilibrium impact on rival firms is negligible here. We draw two basic conclusions from this first set of results:

Table 12 The effect of yoplait PBSD on prices and profits

		Base case (C1)		Yoplait T&C PBSD = 0 (C2)		Yoplait all PBSD = 0 (C3)	
Brand	Sub-brand	Prices	Profit	Prices	Profit	Prices	Profit
Dannon	Creamy fruit blends	1.622		1.622		1.622	
Dannon	Fruit on the bottom	1.690		1.690		1.690	
Dannon	Light n fit	1.733		1.733		1.733	
Dannon	Light n fit creamy	1.751	754.1	1.751	755.6	1.751	757.4
Kemps	Classic	1.380		1.380		1.379	
Kemps	Free	1.483	192.3	1.483	192.5	1.483	193
Old home	Classic	1.403		1.403		1.403	
Old home	100 Calories	1.271	235.1	1.271	235.5	1.270	236
Wells B.B.	Lite 85	1.386	456.5	1.386	457.6	1.385	458.8
Yoplait	Light	1.713		1.712		1.717	
Yoplait	Original	1.671		1.671		1.674	
Yoplait	Thick and creamy	1.754	1895	1.797	1859.1	1.785	1791.9

Prices and per period profits are reported for forward simulated steady state in each scenario (C1, C2, C3)

1) the per pound price of *T&C* is about 4 cents higher when the PBSD effects are removed (while the prices of the two remaining *Yoplait* sub-brands remain essentially unchanged), and 2) *Yoplait* loses about 1.9% of its per period profits as a result of this change. Together, these results suggest that PBSD associated with the *T&C* sub-brand alone leads to a significantly lower price for *T&C*, but that this lower price also yields higher profits, due to substitution from rival firms and (more importantly) the outside good. In the presence of full PBSD effects, loyal *T&C* consumers enjoy the lower prices that *Yoplait* is using to attract new consumers to this product.

Results from a second counterfactual are also reported in Table 12. In this case, for the alternative (scenario C3), we eliminate the parent brand state dependence effects for all three *Yoplait* sub-brands. This is equivalent to all three being marketed under different, unrelated names, without the common *Yoplait* banner. Similar to the first counterfactual, the own state dependence effects remain the same and only cross-sub-brand effects are eliminated. The comparison of this counterfactual with the baseline scenario (C1) reveals a price increase for all three *Yoplait* sub-brands, that is negligible for all but *T&C*. However, per period profits drop by 5.4% (much larger than the previous example), as *Yoplait* loses share to it's rivals and the outside good. This impact was muted in the previous counterfactual by their own-PBSD effects. Consistent with the intuition built earlier in the paper, we conclude that Yoplait can continue to profit from introducing additional layers of PBSD through successful sub-brands, even at the levels of PBSD estimated here (i.e. without increasing its coefficient).

5.2.4 The impact of firm-level profit maximization

For our third set of counterfactual exercises, we turn to the impact of firm-level profit maximization in this real-world context. Recall that, in our numerical analysis, we showed that under certain PBSD parameter values, firms could actually earn more profit by operating sub-brands as independent profit centers rather than maximizing profits jointly. This is a potential strategic effect that stems from shielding the sub-brands from the increased competition associated with high levels of PBSD. To evaluate whether this carries over to the empirical setting, we compute new optimal prices and profits when firms maximize sub-brand profits and compare them to the baseline scenario in which profits are maximized jointly. Any differences in prices and per period profits across the two scenarios can then be attributed to the centralization of the pricing decision alone. In particular, this experiment allows us to quantify the benefit that yogurt parent brands can derive by unifying the profit objective functions of all their sub-brands to account for both respective externalities (cross-product cannibalization and dynamic complementarity).

Table 13 contains steady state prices and per period profits for the base case (C1) (joint profit maximization) versus a counterfactual (C4) where all sub-brands instead maximize sub-brand profits. With respect to prices, we see that full multi-product pricing leads to higher prices across the board for all sub-brands. At the same time, profits for all parent brands are also higher in the joint profit maximization cases, reflecting an increase in market power. The benefits from avoiding cross-cannibalization clearly outweigh the effects of dynamic complementarity in the Yogurt category. However, the impact of the increased market power on profits

Table 13 Sub-brand profit maximization - effect on profits

		Base Case JPM (C1)		Sub-brand PM (C4)	
Brand	Sub-brand	Prices	Profit	Prices	Profit(%Δ from JPM)
Dannon	Creamy fruit blends	1.622		1.612	
Dannon	Fruit on the bottom	1.690		1.679	
Dannon	Light n fit	1.733		1.735	
Dannon	Light n fit creamy	1.751	754.1	1.753	747.0 (−0.9%)
Kemps	Classic	1.380		1.379	
Kemps	Free	1.483	192.3	1.484	191.2 (−0.6%)
Old home	Classic	1.403		1.401	
Old home	100 Calories	1.271	235.1	1.271	233.3 (−0.8%)
Wells B.B.	Lite 85	1.386	456.5	1.386	451.4 (−1.1%)
Yoplait	Light	1.713		1.682	
Yoplait	Original	1.671		1.649	
Yoplait	Thick and creamy	1.754	1895	1.708	1891.0 (−0.2%)

Prices and per period profits are reported for forward simulated steady state for each scenario

is quite modest; per period profits increases range from 0.2% for *Yoplait* to 1.1% for Wells Blue Bunny, which is a single sub-brand firm but loses some profits in the counterfactual because other sub-brands compete more aggressively.

It is clear that, in this particular empirical application, PBSD levels are not sufficiently large as to make sub-brand profit maximization more profitable than joint profit maximization. In fact, even if we repeat the exercise using higher values of PBSD, we still find that joint profit maximizing outperforms the alternative in all cases. The insights from our theoretical simulations suggest that this is likely due to the mitigating effects of SBSD (recall that the SBSD coefficient was 2.5 times larger than that of PBSD), which limits the scope of the potential gains from operating as profit centers instead of jointly.

Finally, the increase in profits we find when comparing the baseline scenario (C1) to the counterfactual (C4), can also be viewed as the incremental benefit to the firm of all parent brands fully internalizing the both the dynamic complement effect and the impact of cross-cannibalization. This quantification could prove useful to managers in making organizational decisions with respect to price setting. For example, if full coordination of prices across *Yoplait* sub-brands is costly (e.g. because of managers' workloads), and this cost is higher than 0.2% of per period profits, our simulation suggests that decentralized pricing might be preferable on balance.

5.2.5 The role of information

For our final set of counterfactuals, we turn to how a firm's knowledge of state dependence, or perhaps more accurately, its degree of sophistication, impacts its pricing and profits. In particular, we consider two scenarios in which firms are unaware of the true impact of state dependence. Our motivation lies in exploring whether firms could perhaps avoid the subtle aspects of dynamic competition induced by consumer inertia by simply being unaware that they exist. To address this possibility, we compute two final counterfactual scenarios in which firms set prices as if consumers do not exhibit particular aspects of state dependence, when they in fact do. This situation might occur if firms actually used empirical methods, such as those developed here, to estimate demand and compute optimal (equilibrium) prices, but failed to fully account for the role of state dependence. Note that this would not be a proper MPE, since firms would be both ignoring available payoff relevant information and failing to take advantage of profitable deviations (implying that their beliefs would not be correct on average, at least with regards to consumer behavior, and would therefore fail to satisfy rational expectations). To compute these outcomes, we first re-estimate the model, ignoring each type of state dependence (one at a time), solve for the "optimal" pricing policies given this new information set, and then compute profit outcomes based on the true (fully state dependent) demand system and these new (information restricted) policies.

The results of two such counterfactual experiments are presented in Table 14, alongside the original baseline results. In the first scenario, firms ignore PBSD (but account for the other two types of state dependence), while in the second they ignore SBSD (while accounting for PBSD and type SD). Note that, consistent with avoiding the full investment effect, when firms ignore PBSD in setting prices (C6), they in fact charge higher steady state prices and earn higher per period profits. For example, profits increase 0.72% for *Yoplait* and 0.56% for *Dannon*. This is a direct consequence of avoiding the increased competition that PBSD creates through its impact on future demand. It is important to note that this is a strategic effect. If, in the same scenario, we use a pricing policy that ignores PBSD for only a single deviant firm,

Table 14 Impact of ignoring state dependence in price setting

		Base case (C1)		Ignore PBSD (C6)		Ignore sub-brand SD (C7)	
Brand	Sub-brand	Prices	Profit	Prices	Profit	Prices	Profit
Dannon	Creamy fruit blends	1.622		1.645		1.667	
Dannon	Fruit on the bottom	1.690		1.712		1.735	
Dannon	Light n fit	1.733		1.755		1.796	
Dannon	Light n fit creamy	1.751	754.1	1.775	758.3	1.78	765.1
Kemps	Classic	1.380		1.392		1.409	
Kemps	Free	1.483	192.3	1.497	193.3	1.516	195.1
Old home	Classic	1.403		1.416		1.427	
Old home	100 calories	1.271	235.1	1.285	236.5	1.313	238.7
Wells B.B.	Lite 85	1.386	456.5	1.403	460.6	1.448	463.9
Yoplait	Light	1.713		1.73		1.778	
Yoplait	Original	1.671		1.686		1.736	
Yoplait	Thick and creamy	1.754	1895	1.785	1908.6	1.743	1915.3

this deviant firm earns lower per period profits, while its rivals (who price correctly), make more.[14] The second scenario (C7) considers the case where firms ignore sub-brand state dependence when setting prices (but still account for PBSD and type SD). Under this scenario, the steady state prices and per period profits also increase, similar to the effect from ignoring PBSD, but the impact is even greater. This likely reflects the relative magnitudes of the two types of state dependence (recall that SBSD was over twice as large). In particular, we find that *Yoplait*'s per period profits increase 1.1% while *Dannon*'s increase by 1.5%. As in the previous example, this is an outcome of softer competition and represents a strategic effect. Again, if only one firm ignores sub-brand SD while its rivals price correctly, it is made unambiguously worse off by its ignorance.

6 Conclusion

There is increasing academic interest in the long term impact of marketing policies and the perils of myopic behavior on the part of managers and decision makers. In this work, we take a forward looking perspective on the pricing decision of multiproduct firms and examine the effect of demand choice dynamics that reflect consumers' tendency to stay loyal to a firm when choosing between multiple sub-brands of a broader umbrella brand. We examine a product category, refrigerated yogurt, where consumers' choices depend not only on past choices of specific product alternatives but also on past choices of the parent brands. Such cross-sub-brand state dependence, for repeatedly purchased goods, generates meaningful dynamics with non-trivial implications for equilibrium prices and profitability.

Our numerical simulations reveal that state dependence to the parent brand can decrease prices and reduce profits, as well as mitigate or even reverse the benefits of joint profit maximization relative to sub-brand profit maximization. Our empirical analysis establishes the real world relevance of parent brand state dependence and confirms that moderate levels of inertia lead to both lower prices and reduced profits, though these effects are mediated and can be even reversed by asymmetries in the market positions of brands and the relative importance of sub-brand state dependence. In the yogurt category, joint profit maximization leads to higher prices and profits, relative to sub-brand profit maximization, but firms could benefit from being unaware of (or ignoring) consumer inertia as long as they all behaved similarly. If only one firm deviated towards "unawareness", it would do so at its peril. Expressed differently, in a "naive" market, there are profitable deviations to becoming more sophisticated.

In order to specify a tractable model and explore some of the dynamic implications we had to abstract from important factors like frequent price discounts and retailer involvement in the choice of shelf price. Nevertheless, studying equilibrium steady state prices in a dynamic game context reveals novel aspects of price

[14]Results not shown, but available upon request. Note that this situation is not a coherent equilibrium, since both sets of firms now have inconsistent beliefs regarding the other type's behavior (i.e. they are setting prices using policies consistent with different underlying games).

dynamics for multiproduct firms. A more complex game with the inclusion of the retailer/manufacturer game and/or the exploration of equilibrium pricing with frequent discounts is an interesting avenue for future research.

Overall, the findings reported in this work highlight the intricate role of demand choice dynamics for prices of forward looking multi-product firms and suggest new dimensions to explore. It seems reasonable to expect that the methodology and insights of this work are relevant for several other product categories where demand is likely characterized by parent brand dynamics.

Appendix: Additional details regarding computation and estimation

State space discretization

One of the challenges we had to overcome in order to compute optimal dynamic price policies was the infinite dimension of the state space. Following the literature we addressed this issue by approximating the infinite state space with a finite set of points and interpolating. The approximation begins by discretizing the state space on a multidimensional grid. Each axis of the grid corresponds to one dimension of the state space, namely the fraction of a particular consumer type loyal to a specific sub-brand. We discretize each axis of the grid with a finite number of points g, such that: $0 = g_{jt}^{n1} < ... < g_{jt}^{nL} = 1,\ \forall j,\ \forall n$. The grid is formed by the Cartesian product of all finite sets of points for each axis, such that $\sum_{k=1}^{J} g_{kt}^{nl} \leq 1,\ \forall l = 1, ..., L,\ \forall n$. Intuitively, this condition says that, for any consumer type n, the fraction of the market loyal to each sub-brand in the choice set should sum to one across all sub-brands. The main computational challenge in solving a discretized version of a dynamic programming problem with a continuous state space is caused by dimensionality. This is because the number of grid points at which one must solve for the value and policy functions increase exponentially with the dimensionality of the state space. For the game described in Section 3, the state space has dimension equal to $G = (J-1) \times N$, which is the number of choice alternatives minus one, times the number of consumer types. Due to the fact that loyalty states for each type sum to one across sub-brands, only *J-1* loyalty states need to be tracked for each type. For a regular grid with L points in each axis, the total number of points would be L^G. The grid for our problem is not rectangular, but rather triangular (if we think about it in two dimensions) due to the requirement that the states of loyalty of a consumer type to all sub-brands must sum to one. Even though this condition reduces the number of grid points significantly, as the dimensionality of the state space increases, the computational requirements of the grid become exorbitant. In practice we are using a grid consisting of six points in each dimension (0, 0.2, 0.4, 0.6, 0.8, 1). While the complete Cartesian product for such a grid would have 362 million points (6^{11}), the condition that the state shares for all brands should sum up to one limits our total number of grid points to 4368. To see this, consider the case when sub-brand A has a state vector of 1, then the only possible states for the other sub-brands is to have zero state share. Similarly, when sub-brands A and B have state shares of 0.4 and 0.6 respectively, the only possible states for the other sub-brands are zero state shares.

It is relatively easy to see that the dimensionality of the state space increases faster with the number of brands when the number of consumer types is higher. This is especially true because adding consumer types is not economizing as much on the condition that state shares sum up to one (this condition allowed us to work with 4368 grid points instead of 363 million grid points). An additional consumer type would imply another vector of sub-brand state shares (so another 4368 points to enumerate) and then the total state space would be the product of the two consumer type grids (4368^2). This creates a trade-off between using more consumer types or more choice alternatives. By settling with one consumer type, we were able to include all twelve sub-brands of the sample in the pricing model, thereby ensuring internal consistency.

Interpolation

During computation, we use polynomial based interpolation for all cases where we need to compute the value or policy functions on state space points outside the grid. Our polynomial approximation function has the general form given below

$$\hat{y}_j(s) = \hat{a_0} + \sum_{n=1}^{N} \sum_{j=1}^{J-1} \sum_{m \in \{0.5,1,2,3\}} \hat{a_{njm}} \left(s_j^n\right)^m + \sum_{k=1}^{K} \hat{a_k} \prod_{l \in I_k} s_l^n \tag{13}$$

It includes all the first, second and third order terms of the state variables, their square root, and several sets of (two way, three way, four way, etc.) interactions between states of different brands for the same consumer type. In the implementation of the algorithm, the correlation between predicted polynomial values and actual values is about 0.99 while the average percent error of the prediction (MAPE) is at most 0.1%. This suggests that the approximation works well in practice.

Policy function iteration

The dynamic game analysis proceeds in two steps. In a first phase we compute optimal pricing policies for each point in the state space and in a next step we compute steady state prices and shares for all sub-brands in the sample. The steps for the policy function iteration are as follows:

1. Start with initial guesses of value functions and price strategies. For all reported results we use zero to initiate the value functions and optimal prices for static period profits to initiate the policy functions, for each point in the state space; $V_f^0(s) = 0$ and $\sigma_f^0(s) = max_{P_j} \pi_f \left[s, P, \sigma_{-f}^0(s)\right] \; \forall s \in S, \; \forall f$. The initial policies are iterated so that they are best responses for the static case.
   - We experiment with different initial guesses, for the value and policy functions, to examine whether the equilibrium policies are the same or change depending on the starting values; the latter would imply the existence of multiple equilibria. It is noteworthy that all trials generated the same equilibrium policies.

2. For each firm, given V_f^{iter-1} and σ^{iter-1}, compute V_f^{iter}; here superscripts refer to iterations of the algorithm. The computation of value functions involves the purchase probabilities that determine both static and future discounted profits.

   (a) Iterate the Bellman equation until convergence, $\frac{max|V_f^{iter}-V_f^{iter-1}|}{max|V_f^{iter}|} \leq \varepsilon_v, \ \ \forall f$. Then move to the next step. We set $\varepsilon_v = 0.0000001$.

3. For each firm, given V_f^{iter} and σ^{iter-1}, compute σ^{iter}. This involves finding the optimal price for f, for each point on the grid, given the right hand side of the Bellman equation and the rival's policy. In practice we iterate across firms, solving for optimal prices at all points of our state space grid, given rival prices from the previous iteration. Upon convergence, no firm can get a higher value function by changing its policy. Policies are then best responses for each state, conditional on value functions. To speed our price optimization sub-routine and to avoid local maxima issues we start with a simple global search to identify the region of the global maximum for prices. That is, we evaluate the right hand side of the Bellman equation in 20 cent intervals to find the neighborhood of the solution. We then bound our quasi-Newton optimization algorithm in a 40 cent neighborhood of the solution identified with the initial search.

   (a) If the computed policies converge for each firm , $\frac{max|\sigma_f^{iter}-\sigma_f^{iter-1}|}{max|\sigma_f^{iter}|} \leq \varepsilon_\sigma, \ \ \forall f$, stop the algorithm. We set $\varepsilon_\sigma = 0.00001$.
   (b) If not, update the policies and return to step 2.

Doraszelski and Satterthwaite (2010) give the following definition for a Markov-perfect equilibrium: "An equilibrium involves value and policy functions V and σ such that i) given σ_{-f}, V solves the Bellman equation for all f and ii) given $\sigma_{-f}(s)$ and V_f, $\sigma_f(s)$ solves the maximization problem on the right hand side of the Bellman equation for all s and all f." Markov-perfect equilibria are by definition sub-game perfect, meaning that firms follow optimal strategies at each possible state. Upon convergence, the algorithm described above satisfies these general conditions and thus computes a Markov-perfect equilibrium.

To ensure that our solution is not a local minimum, but rather a global maximum, as well as a best response for all profit maximizing firms, we augmented our algorithm with an additional step that checks our equilibrium policies with a global optimizer that combines genetic algorithms with derivative based search. Indeed, when we used this global optimizer for our final policy iteration, equilibrium policies did not change, evidence that our original approach was suitably robust. Details about this optimizer can be found in Mebane W. and Sekhon J., 2011, "Genetic Optimization Using Derivatives: The rgenoud package for R.", Journal of Statistical Software, 42(11).

To complete the game specification we also need to make assumptions regarding parameters that are part of the model but are unobserved, namely the discount factor and the total size of the market. The parametrization used in the algorithm is: $\beta = 0.998, \ \ MS = 100000$.

Steady state computation

To compute the steady state of a pricing game specification, we start from some initial state vector and, given the best response price policies, we first find the optimal price of each firm for the given period. Next, we compute the states that prevail in the market under the prices of the last iteration and use them as the next period's state vector. We repeat the process until both the state variables and the prices of each specific product alternative converge. To verify if the obtained steady state is unique, we repeat the computations several times starting from different initial states. Throughout all trials, the algorithm always converged to the same unique steady state.

BBL estimation procedure

Our implementation of the BBL estimator involves the following steps:

1. We estimate policies in a first stage using spline-based regression. Policies for each sub-brand are estimated as very flexible functions of the loyalty shares of the various sub-brands (i.e. observed state vectors)
2. We define a set of inequalities H based on deviations from the optimal estimated policies by varying amounts (e.g. $+/-$ 0.005, $+/-$ 0.01, $+/-$ 0.02) separately for each sub-brand.
3. Using our first stage policies, we forward simulate "correct" and "perturbed" value functions for each inequality in H in the following way:
   (a) Draw a random state vector s^0 and use it as a starting point.
   (b) Simulate current period profit flow $\pi_{ft}(s_t, \hat{\sigma}(s_t, v_t))$ using the current state vector and estimated policies. The estimated policies $\hat{\sigma}(s_t, v_t)$ also include i.i.d. random shocks drawn from the residuals of the first stage estimation.
   (c) Calculate next period's state vector based on choice probabilities that correspond to the demand model, the policies and the state vector from the previous step. This is specified explicitly in our pricing game and is described in Section 3.3 and Eq. 9, $s^n_{t+1} = g(P_t, s^n_t) = Q^n(P_t) \times s^n_t$.
   (d) Continue forward simulation until discounted period profit flows are negligible.[15]
   (e) Repeat for R different starting state vectors and average the results. This step averages out both different starting state vectors and the random shocks of the policies.

$$\hat{V}_f(s;\sigma;\theta) = \frac{1}{R}\sum_{r=1}^{R}\left[\sum_{t=0}^{T}\beta^t \pi_{ft}(s_t, \hat{\sigma}(s_t, v_t);\theta)\right] \tag{14}$$

4. We use the approach described above to obtain "correct" and "perturbed" value functions for a set of inequalities that constitute the equilibrium conditions of

[15]We use 48 years or 2500 weeks - we have also tested a horizon of 3000 and 3500 weeks and results do not change substantially, evidence that any simulation error is already too small when using 2500 weeks.

our model. The final step of our approach recovers cost parameters through a minimum distance estimator that penalizes violations of the equilibrium conditions. Denoting the correct and perturbed value functions by $V_f(s;\sigma_j,\sigma_{-j};\theta)$ and $V_f(s;\sigma_j',\sigma_{-j};\theta)$ respectively, the second stage estimator minimizes the following objective function.

$$Q(\theta)=\frac{1}{H}\sum_{h=1}^{H}\left(min\left\{V_f(s;\sigma_f,\sigma_{-f};\theta)-V_f\left(s;\sigma_f',\sigma_{-f};\theta\right),0\right\}\right)^2 \quad (15)$$

Applying the methodology is facilitated by the fact that the value functions for the pricing model of this study can be written as a linear function of the structural parameters which, in this case, are costs. This allows for significant computational economies in that the forward simulation need only be done once, before the estimation, and not for every trial parameter vector of the estimation routine. For demonstration purposes, we describe the linear form of the model below.

$$\begin{aligned} V_f(s;\sigma;\theta) &= \sum\nolimits_{t=0}^{\infty}\beta^t\left[\sum\nolimits_{j\in f}(P_{jt}-c_j)\times D_{jt}\times M\right] \\ &= \sum\nolimits_{t=0}^{\infty}\beta^t\left[\sum\nolimits_{j\in f}P_{jt}\times D_{jt}\times M\right]-\sum\nolimits_{j\in f}\left[\sum\nolimits_{t=0}^{\infty}\beta^t(D_{jt}\times M)\right]\times c_j \end{aligned} \quad (16)$$

The linearity of the value function allows us to construct empirical analogs of both the "true" and perturbed value functions using forward simulated estimates of D_{jt}, pricing policies, market sizes and any trial values for the cost parameter vector.

The demand parameters we use when solving for the cost parameters are Bayesian posterior estimates that correspond to the average household. We note that consistent estimation of the dynamic parameters requires consistent estimates of the demand parameters as inputs. Our Bayesian estimates are asymptotically equivalent to MLE estimates and therefore satisfy this requirement, so long as the demand model is correctly specified.

References

Anand, B., & Shachar, R. (2004). Brands as beacons: a new source of loyalty to multiproduct firms. *Journal of Marketing Research*, *16*, 135.

Anderson, S., & De Palma, A. (2006). Market performance with multiproduct firms. *The Journal of Industrial Economics*, *54*(1), 95.

Arie, G., & Grieco, P. (2014). Who pays for switching costs? *Quantitative Marketing and Economics*, *12*(4), 379.

Bajari, P., Benkard, L., & Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, *75*(5), 1331.

Besanko, D., Doraszelski, U., Kryukov, Y., & Satterthwaite, M. (2010). Learning-by-doing, organizational forgetting, and industry dynamics. *Econometrica*, *78*(2), 453.

Bronnenberg, B., Kruger, M., & Mela, C. (2008). Database paper: The IRI marketing data set. *Marketing Science*, *27*(4), 745.

Cabral, L. (2008). Small switching costs lead to lower prices. Working Paper, New York University.

Cabral, L. (2016). Dynamic pricing in customer markets with switching costs. *Review of Economic Dynamics*, *20*, 43.

Doganoglou, T. (2010). Switching costs, experience goods and dynamic price competition. *Quantitative Marketing and Economics*, *8*, 167.

Doraszelski, U., & Satterthwaite, M. (2010). Computable Markov-perfect industry dynamics. *RAND Journal of Economics*, *41*(2), 215.

Draganska, M., & Jain, D.C. (2005). Consumer preferences and product-line pricing strategies: an empirical analysis. *Marketing Science*, *25*(2), 164.

Dubé, J.P., Hitsch, G., Rossi, P., & Vitorino, M.A. (2008). Category pricing with state dependent utility. *Marketing Science*, *27*(3), 417.

Dubé, J.P., Hitsch, G., & Rossi, P. (2009). Do switching costs make markets less competitive? *Journal of Marketing Research*, *46*(4), 435.

Dubé, J.P., Hitsch, G., & Rossi, P. (2010). State dependence and alternative explanations for consumer inertia. *RAND Journal of Economics*, *41*(3), 417.

Erdem, T. (1998). An empirical analysis of umbrella branding. *Journal of Marketing Research*, *35*, 339.

Erdem, T., & Sun, B. (2002). An empirical investigation of the spillover effects of advertising and sales promotions in umbrella branding. *Journal of Marketing Research*, *39*, 408.

Farrell, J., & Klemperer, P. (2007). Coordination and lock-in: competition with switching costs and network effects. In *Handbook of industrial organization,* Vol. 3. Elsevier.

Gamerman, D., & Lopes, H. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*, 2nd edn. Chapman & Hall/CRC.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2004). *Bayesian data analysis,* 2nd edn. Chapman & Hall/CRC.

Horsky, D., & Pavlidis, P. (2011). *Brand loyalty induced price promotions: an empirical investigation*. Working paper, University of Rochester.

Judd, K. (1998). *Numerical methods in economics*. Cambridge, Massachusetts: The MIT Press.

Klemperer, P. (1995). Competition when consumers have switching costs: an overview with applications to industrial organization, macroeconomics and international trade. *The Review Of Economic Studies*, *62*(4), 515.

Miklos-Thal, J. (2012). Linking reputations through umbrella branding. *Quantitative Marketing and Economics*, *10*(3), 335.

Morrow, R., & Skerlos, S. (2010). On the existence of Bertrand-Nash equilibrium prices under logit demand. Working Paper. University of Michigan.

Rossi, P., Allenby, G., & McCulloch, R. (2005). *Bayesian statistics and marketing*. New York: Wiley.

Seetharaman, P.B. (2004). Modelling multiple sources of state dependence in random utility models: a distributed lag approach. *Marketing Science*, *23*(2), 263.

Seetharaman, P.B., Ainslie, A., & Chintagunta, P. (1999). Investigating household state dependence effects across categories. *Journal of Marketing Research*, *36*(4), 488.

Viard, B. (2007). Do switching costs make markets more or less competitive? The case of 800-number portability. *RAND Journal of Economics*, *38*(1), 146.