Editorial Manager(tm) QMEC
Manuscript Draft

Manuscript Number: QMEC22R2

Title: Brand and Quantity Choice Dynamics under Price Uncertainty

Article Type: Original Research

Keywords: price expectations, pricing, scanner data, , dynamic programming, simulation, , discrete choice.

Corresponding Author: Tulin Erdem University of California, Berkeley

Other Authors: Tulin Erdem, Susumu Imai , Michael P. Keane

Brand and Quantity Choice Dynamics under Price Uncertainty

December 2001

Revised, December 2002

Note: An earlier draft of this paper was presented under the title "Consumer Price and Promotion Expectations: Capturing Consumer Brand and Quantity Choice Dynamics under Price Uncertainty." The authors thank seminar participants at AGSM, UC-Davis, UC-Irvine, UCLA, Cornell, the University of Toronto, Tel Aviv University, Wharton, NYU, MIT, Berkeley, the 1997 INFORMS conference and the 1998 Marketing Science Conference for useful comments on the earlier draft. We also thank seminar participants at the 2002 North American Winter Meetings of the Econometric Society, the University of Chicago, Stanford University, the 2002 Marketing Science Conference and the Cowles Foundation Conference on Estimation of Dynamic Demand Models for valuable comments on the later version. Comments from Jason Cummins, Steve Berry, Matt Shum, Igal Hendel, Jim Heckman, Derek Neal and Susan Athey have been particularly helpful. We especially thank to Peter Rossi for his extensive and helpful comments on both the earlier draft and the later version of this paper. This research has been funded by NSF grants SBR-9511280 and SBR-9812067.

Abstract

We develop a model of household demand for frequently purchased consumer goods that are branded, storable and subject to stochastic price fluctuations. Our framework accounts for how inventories and expectations of future prices affect current period purchase decisions. We estimate our model using scanner data for the ketchup category. Our results indicate that price expectations and the nature of the price process have important effects on demand elasticities. Long-run cross price elasticities of demand are more than twice as great as short-run cross price elasticities. Temporary price cuts (or "deals") primarily generate purchase acceleration and category expansion, rather than brand switching.

Keywords: price expectations, pricing, scanner data, dynamic programming, simulation, discrete choice.

The goal of this paper is to develop and estimate a dynamic model of consumer choice behavior in markets for goods that are: 1) frequently purchased, 2) branded, 3) storable, and 4) subject to frequent price promotions, or “deals.” In such an environment, forward-looking behavior of consumers is important. Specifically, optimal purchase decisions will depend not only on current prices and inventories, but also on expectations of future prices. There is no single “price elasticity of demand.” Rather, the effect of price changes on consumer demand will depend upon how the price change effects expectations of future prices. This depends on the extent to which consumers perceive the price change to be permanent or transitory, and the extent to which they expect competitor reaction. These, in turn, depend on the stochastic process for prices in the market (see Marshak (1952), Lucas (1976)).

In recent years a wealth of supermarket scanner data have become available that document sales of frequently purchased consumer goods. In a number of instances, panels of households have been provided with individual ID cards, so that all their purchases over long periods of time can be tracked. These data provide a valuable opportunity to study consumer choice dynamics. We will argue that such analysis is important not only for marketers wishing to predict consumer response to promotions, but also for economists interested in firm pricing behavior, antitrust policy, welfare gains from introduction of new goods, construction of price indexes, etc.

Since the pioneering work of Guadagni and Little (1983), an extensive literature has emerged that uses scanner data to study consumer choice behavior. But for the most part, this literature has relied on <u>static</u> models of consumer behavior, in the sense that consumers make decisions to maximize current period utility. Much of this literature has dealt with the issue of choice “dynamics,” where dynamics is used to refer to purchase carry over effects (or habit persistence) – i.e., does past purchase of a brand increase a consumer’s current period utility from purchase of that brand (see, e.g., Keane(1997a))? But none of the published literature examines consumer choice “dynamics” in the sense of how expectations of future prices influence the current period purchase decisions of forward looking consumers.[1]

[1] Erdem and Keane (1996) develop a model of forward looking consumers, but the focus there is on learning about brand quality in an environment where consumers have uncertainty about brand attributes. This generates a motive for trial or experimental purchases of brands to facilitate learning. Erdem and Keane model prices as iid over time, so changes in current prices do not alter expected future prices.

Understanding the role of price expectations in consumer purchase behavior is important for many reasons. For instance, evaluations of the welfare effects of mergers and welfare gains from introduction of new goods (see Hausman (1997)), rely on estimates of own and cross-price elasticities of demand for the goods in question. But the existing literature only contains static elasticity estimates. Such estimates do not account for how a price cut today affects consumer expectations of future prices, or how elasticties may differ for price cuts that are perceived to have different degrees of persistence.[2] We provide a framework for estimating dynamic price elasticities of demand for branded frequently purchased consumer goods. We will show that accounting for dynamics can have large effects on own and cross-price elasticity estimates.

More generally, our framework can be used to predict how consumers' purchase decision rules would respond to changes in the entire retail pricing process (such as, for example, a shift from high/low (H/L) pricing to "everyday low pricing" or EDLP). To our knowledge there is no prior structural work that enables one to predict consumer response to "major" pricing policy changes.[3] This problem is apparently understood by marketing practitioners. For example, in a criticism of existing models of promotion response Struse (1987), a marketing manager at General Mills, observed that: "While analysis of past events may be … useful, the real need is to better predict the future - especially under *interesting* circumstances. That is, the manager needs a forecasting method which will be robust and discriminating over a wider range of conditions than actually seen in the market since he or she needs to explore alternatives which go beyond past practice …"

Understanding consumers' dynamic responses to pricing policy changes may also be important for understanding industry dynamics. Existing dynamic oligopoly models that endogenize price (see, e.g., Berry, Levinson and Pakes (1995)) typically assume that consumer behavior is static. This may be a serious misspecification in markets where purchases are made frequently, and changes in current prices lead to important changes in expected future prices. We

---

2 In "market mapping" methods (see Elrod (1988)) cross-price elasticities of demand are critical for the evaluation of the positioning of products in unobserved (or latent) attribute space. Srinivasan and Winer (1994) and Erdem (1996) discuss how "dynamics" in the sense of habit persistence may distort such evaluations. How dynamics in the sense of price expectation formation might distort such evaluations has not been considered.

3 See Keane (1997b) for a discussion of this issue. To give an example of the problem, we would expect that price elasticities of demand would differ between an EDLP regime and a H/L regime for a variety of reasons. For instance, a price cut has different effects on expected future prices under each regime, and the expected duration until the next price cut is different under each regime. As a result, one can't use estimates obtained under the H/L regime to predict behavior under the EDLP regime, unless one uses a structural model like ours.

think work like ours will eventually prove useful for researchers seeking to elaborate the consumer side of dynamic oligopoly models.[4]

Understanding how forward looking consumers respond to temporary price cuts is important for retailers and brand managers, who want to know if price cuts merely cause consumers to accelerate purchases, or whether they also induce brand switching and/or increased category sales. Furthermore, the design of intertemporal price discrimination strategies requires an understanding of how changes in the whole price process affect consumer demand (e.g., would more frequent promotion generate sales to new consumers, or simply alter the purchase timing of existing consumers?).

As a final example, an understanding of the dynamics of consumer purchase behavior is important for the construction of price indices. To some extent, this involves the random sampling of posted supermarket prices, which will capture average offer prices.[5] But, if a large share of purchases occurs on promotion, then the average offer price of a good is not the relevant measure of its typical cost to consumers. In fact, a widespread shift from H/L pricing to EDLP, such as occurred in the US in the late 80's and early 90's (see Lal and Rao (1997) for a discussion), could cause the average posted price to fall even though the average purchase price does not, thus distorting price level estimates based on random sampling of posted prices. Our framework allows one to estimate the relationship between mean offer and accepted prices under alternative price processes.

In this paper, we estimate our model of consumer brand and quantity choice dynamics on scanner panel data provided by A.C. Nielsen. We use the data on household ketchup purchases. We choose the ketchup category for two reasons. First, it satisfies the four criteria discussed at the outset. In particular, there are frequent price promotions for ketchup. Pesendorfer (2002) finds that there is little evidence of seasonality in ketchup demand or prices, and that cost factors seem unrelated to short run price movements. He argues that a type of inter-temporal price

---

4 The computational capacity and econometric methods needed to estimate equilibrium models with forward-looking behavior on both the firm and consumer side are probably several years away. But we should note that Ching (2002) has estimated a model of the pharmaceutical industry with dynamics on both sides of the market in two stages. First, the demand side model is estimated jointly with an approximate reduced form equation for firm's pricing policy function. In a second stage the remaining supply side parameters are calibrated, treating the demand side parameters as known.

5 The BLS website (see www.bls.gov/cpi/cpifact2.htm) contains some description of the random sampling of prices at selected department stores, supermarkets, service stations, doctors' offices, rental units, etc. that underlies construction of the CPI.

discrimination strategy on the part of firms, in which the retailers play mixed strategies, most plausibly explains frequent week-to-week price fluctuations for ketchup. We agree with this analysis, which supports the view that price movements are exogenous from the point of view of consumers. We believe that similar factors are at work in most frequently purchased consumer goods markets.

Second, of the goods that satisfy our four criteria and for which scanner data have been released for public use, ketchup is the easiest category to work with. This is because the number of brand/size combinations for ketchup is lower than for the other available categories (there are 4 brands – Heinz, Hunts, Del Monte and the Store brand – that come in 3 to 5 sizes each, giving a choice set with 16 elements). We felt it was sensible to first apply our framework to this category before tackling categories with more brands and/or sizes (such as yogurt, toilet paper, cereal, etc.).

Our estimated model provides a very good fit to all the important dimensions of the data, including brand shares, size shares, purchase frequency, inter-purchase times, purchase hazard rates, brand switching matrices, and the distributions of accepted prices. In our view, this is a necessary condition in order for the model's predictions to be credible.

We use simulations of the model to evaluate the importance of price expectations. For instance, we can simulate the effect of a temporary price cut for one brand, both allowing for the effect of this price cut on expected future prices, and holding expectations fixed. Since the price process for ketchup exhibits substantial persistence, we find, as one would expect, that the current period increase of own brand sales in response to a temporary price cut is dampened by the expectations effect. However, this dampening effect is rather modest. For example, it is about 10% for the leading brand - Heinz. Interestingly, however, we find that the cross-price effects that account for expectations are roughly twice as large as cross-price effects holding expectations fixed. For example, the percentage drop in current period sales for Hunts, Del Monte and the Store brand are roughly twice as great if we account for the effect of the Heinz price cut on expected future prices of all the brands.

Two factors drive this key result: 1) if Heinz' price is lowered today it leads consumers to also expect a lower Heinz price tomorrow. This lowers the value function associated with purchase of any brand other than Heinz today. 2) Given the price dynamics in the ketchup market, a lower price of Heinz today leads consumers to expect competitor reaction, so it lowers the expected prices of the other brands tomorrow. This further lowers the value associated with purchase of those brands today.

Obviously, the quantitative significance of these two effects depends on the price process. Thus, a key point is that cross-price elasticities do not (by themselves) reveal the similarity of differentiated products in attribute space (or their degree of competition). The magnitudes of cross-price elasticities also depend on the price process - because this determines how a price cut for one brand today affects expected prices of all brands in the future. Given the importance of cross-price elasticities of demand in such areas as the analysis of mergers and the valuation of new goods, our results clearly show that accounting for consumer price expectations may be critical in these areas.

1. Background and Literature Review

Research on joint modeling of consumer brand and quantity decisions has a long tradition in both marketing and economics. Hanneman (1984) developed a unified framework for formulating econometric models of discrete (e.g., brand choice) and continuous choices (e.g., quantity decisions) in which the discrete and continuous choices both flow from the same underlying utility maximization decision.[6] Dubin and McFadden (1984) used such a model to analyze residential electric appliance holdings and consumption. In marketing, Chiang (1991) and Chintagunta (1993) also adopted the Hanneman framework and calibrated static models consistent with random utility maximization on scanner panel data.

All these models assume that consumers are myopic in that they maximize immediate utility. However, frequently purchased consumer goods typically exhibit substantial inter-temporal price variation, which suggests that for storable goods consumer expectations about future prices may play an important role in purchase timing and quantity decisions. Indeed, the evidence of forward-looking behavior in frequently purchased consumer goods markets is overwhelming. For example, in descriptive analyses, both Hendel and Nevo (2001) and Pesendorfer (2002) find that, conditional on current price, current demand is higher when past

[6] In Hanneman's framework, the commonly observed phenomenon that consumers rarely (if ever) buy multiple brands of a frequently purchased product on a single shopping occasion is shown to arise if the brands are perfect substitutes, quantity is infinitely divisible and pricing is linear. In that case, the brand and quantity decisions separate: In stage 1 it is optimal to choose the brand with the highest utility per unit, and in stage 2 the consumer chooses the number of units conditional on that brand. Keane (1997b) pointed out that this separation does not go through if available quantities are discrete, as is the case with the large majority of frequently purchased consumer goods. However, the literature typically ignores this problem, and assumes quantity is continuous, because of the computational difficulty involved in modeling choice among a multitude of discrete brand/size combinations.

prices were higher or time since last sale is longer (implying that past sales were lower, and hence that current inventories are lower). This implies that consumers "stock up" on storable goods when they see a "deal."

Shoemaker (1979) and Ward and Davis (1978) were perhaps the first (of many) studies to find evidence of "purchase acceleration," meaning that deals induce consumers to buy larger than normal quantities. Neslin, Henderson and Quelch (1985) found that advertised price cuts led to both shorter interpurchase time and larger purchase quantities for coffee. Hendel and Nevo (2001) confirm this for 3 more products, and also find that duration to <u>next</u> purchase is longer following a deal purchase. It is the combination of both increased current purchases and longer duration to next purchase that one needs forward-looking behavior to explain. While a static model with an outside good can explain a current increase in category sales in response to a temporary price cut, the increase in duration to next purchase implies that consumers time purchases to coincide with prices that are "low" relative to some inter-temporal standard.

The large literature on "reference prices," starting with Winer (1986), consistently finds that consumers base current purchase decisions not just on current prices but also on how these relate to some inter-temporal pricing standard (i.e., an average or typical price for the product). This is highly suggestive that expectations of future prices affect consumer purchase decisions.

There is also clear (recent) evidence that the Lucas Critique is quantitatively relevant. Mela, Jedidi and Bowman (1998) examine 8 years of data for a frequently purchased consumer product. During the last 6 quarters of their data there was a regime shift where deals became much more frequent. Under the new regime: 1) consumers bought less often, concentrating their purchases in deal periods, 2) consumers bought larger quantities when they did buy, and 3) overall sales were roughly constant. Mela et al conclude that, under the new regime, consumers "learned to lie in wait for deals." Furthermore, Kopalle, Mela and Marsh (1999) find (for several products) that increased frequency of promotion reduces "baseline sales" of a brand, and also increases its price elasticity of demand.

The behavior of retail prices also provides indirect evidence for the importance of forward-looking behavior by consumers. Both Pesendorfer (2002) and Hong, McAfee and Nayyar (2002) point out that it is hard to explain observed serial correlation in retail prices without consumer stockpiling behavior. In static price discrimination story, a la Varian (1980), prices should be iid over time. In contrast, suppose there exists a segment of price sensitive consumers who stockpile the good and "lie in wait for deals," creating scope for intertemporal price discrimination. As time since the last sale increases, the number of price sensitive consumers looking to buy grows, which increases potential revenue from a sale. Eventually, the retailer decides to have a sale, and then quickly returns price to the "regular" level. This positive duration dependence in the probability of a deal is in fact the price pattern observed for frequently purchased storable consumer goods.

In the marketing literature there are two influential papers that examined the purchase timing, brand choice and quantity decision of consumers for frequently purchased storable consumer goods. These are Gupta (1988) and Chintagunta (1993). Gupta models all three decisions, but the decisions are not linked, and there is no consumer taste heterogeneity. Chintagunta models all three choices in a unified utility maximization framework, and he allows for consumer taste heterogeneity. Interestingly, these two papers reach opposite conclusions regarding a key issue: Gupta concludes that most increased sales from a temporary price cut are due to brand switching, and that cross-price elasticities of demand are large. In contrast, Chintagunta finds that most increased sales from a temporary price cut are due to purchase acceleration by brand loyal consumers, and concludes that cross-price elasticities of demand are small. The Gupta results are the main evidence in the literature that is taken as unfavorable for dynamics/stockpiling behavior.

In fact, the contrast between the Gupta (1988) and Chintagunta (1993) results is exactly what one would expect if forward-looking/stockpiling behavior is important. The difference in results would then be generated by dynamic selection and endogeneity bias. To see this, consider the following example. Suppose Brand A has a deal in period t. Then, the population of people who buy the category at *t* has an over representation of people "loyal" to A. In a static logit brand choice model, such as in Gupta (1988), low price for a brand is therefore correlated with high taste for the brand. As a result, cross-price effects are overestimated. Chintagunta (1993) deals with this selection bias because he allows for taste heterogeneity. Indeed, Sun, Neslin and Srinivasan (2001) show, using simulations, that static choice models without heterogeneity drastically overstate cross-price elasticities if consumers engage in stockpiling behavior.

Recently, there have been a number of papers dealing with the issue of potential endogeneity of prices in consumer choice models (see, e.g., Nevo (2001)). In our view, much of this literature has missed the mark, because it has failed to make a crucial distinction between endogeneity stemming from aggregate (market) demand shocks and endogeneity stemming from omitted variables. Frequently purchased consumer goods typically exhibit price patterns in which prices stay flat for weeks or months at a time ("regular price"), and then exhibit short-lived drops ("deals"). We find it extremely implausible that these deals are the result of manufacturer, wholesaler or retailer responses to aggregate taste shocks, for several reasons. Why would demand for a good like ketchup or yogurt suddenly jump every several weeks and then return to normal? And how could sellers detect such a jump quickly enough to incorporate it into daily or weekly price setting? As we noted earlier, a more plausible explanation for the observed price variation is some sort of inter-temporal price discrimination, such as that considered by Pesendorfer (2002) and/or Hong, McAfee and Nayyar (2002).

On the other hand, an important reason for endogeneity of prices in demand models is the failure to account for consumer inventories, which are not observed in scanner data. If prices are persistent over time and consumers engage in stockpiling behavior, then inventories will be correlated with current prices. This causes price to be econometrically endogenous due to the omitted variables problem, even though price fluctuations are exogenous from the point of view of consumers.[7] The correct way to deal with this problem is to estimate a dynamic demand model, and to integrate out the unobserved latent inventory levels from the likelihood function. This is extremely computationally demanding, but it is exactly what we do in this paper.

In principle, an alternative to our approach would be a BLP procedure using instruments for price that are uncorrelated with inventories. But the instruments would have to be correlated with current but not lagged prices, for if they are correlated with lagged prices they would be related to inventories by construction. Given the serial correlation in prices, such instruments would be very difficult if not impossible to find.

To our knowledge there is no published research that structurally estimates a model of consumer brand and quantity choice dynamics for frequently purchased storable consumer goods under price uncertainty.[8] After our work on this project was well under way we became aware of ongoing work by Hendel and Nevo (2002), who develop a structural model that is in some ways

[7] We thank Steve Berry for pointing this out to us.

[8] We note that Gönül and Srinivasan (1996) estimated a dynamic model with uncertainty about coupon availability, using data on the diaper category. But they consider only category choice and not brand choice. The category price index depends on a weighted average of coupon availability measures across brands. Prices are assumed equal across brands and over time. They also ignore quantity choice, and assume that the probability of a stockout depends only on the current purchase decision and not on the lagged inventory level.

similar to ours. In the course of presenting our model (in the next section) we will provide some discussion of how their approach differs from ours.

2. The Model

2.1. Overview

In our model, the good is storable, and households get utility from its consumption. Brands differ in the utility they provide per unit consumed. A key aspect of the model is that consumers have a per period usage requirement for the good, which is stochastic, and which is only revealed after the purchase decision is made. Thus, households run a risk of stocking out of the good if they maintain an inadequate inventory to meet the usage requirement. There is a cost of stocking out. At the same time, there are carrying costs of holding inventories, and fixed costs of making purchases. The prices of each brand evolve stochastically according to a (vector) stochastic process that is known to consumers.

The model incorporates consumer heterogeneity in two ways: First, we allow for four types of consumers in terms of their vector of utility evaluations for the brands. Second, we also allow for four types of consumers in terms of the usage rate. Thus, there are sixteen types in all. We find that this degree of heterogeneity allows us to fit the data very well. A novel aspect of our model is that a household's usage rate type evolves over time according to a Markov process. A salient feature of the data is that households will often be frequent purchasers of ketchup for several months, then stop buying ketchup for several months, etc. Allowing usage rate type to evolve stochastically over time allows us to capture this type of pattern.

A vital component of our model is the price process, which we estimate separately in a first stage, using the price data from Nielsen. We estimate a multivariate jump process that captures three key features of the data: 1) prices typically are constant for several weeks, followed by jumps, 2) the probability and direction of jumps depends on competitor prices, and 3) the direction of jumps depends on own lagged price (so the jump process is autoregressive). Consumers are assumed to know the price process for each brand, and to be aware of prices every week.

2.2. Household Utility

We assume that households have utility functions defined over consumption of each brand of a particular good and a composite other commodity. Denote the per period utility function for household *i* at time *t* by:

$$U_{it} = U(C_{i1t},...,C_{iJt},Z_{it})$$

where C_{ijt} is the quantity of brand j consumed by household *i* at time *t*, and Z_{it} is the quantity of the outside good that is consumed. Utility depends on quantities consumed rather than quantities purchased because the good in question is storable and households hold inventories. To simplify the model we assume that the composite good is not storable.

Further, we assume that utility is linear in consumption and additively separable between the storable commodity and the composite other good, so U_{it} takes the form:

$$(1) \qquad U_{it} = \sum_{j=1,J} \psi_{ij} C_{ijt} + Z_{it}$$

where ψ_{ij} represents household *i*'s evaluation of the efficiency units of consumption provided by each unit of brand *j*. The assumption of perfect substitutability among brands, and that brands generate differential utility per unit consumed, is similar to the set up in Hanneman (1984). This linear form allows us to ignore saving decisions, so that the only inter-temporal link in the model comes through inventories. We view this simplification as desirable, since the focus of our study is on inventory decisions and not saving decisions.

We model unobserved heterogeneity in consumer evaluations of the efficiency units of consumption, ψ_{ij}, by adopting a finite mixture approach (e.g., Heckman and Singer (1984), Kamakura and Russell (1989)). Thus, we assume that there are $k=1, \ldots, K$ types and we estimate type-specific parameters for the evaluation of the efficiency units of consumption, ψ_{kj}, along with the probability that a household is type *k*, which we denote by ω_k.

It is well established in the marketing literature that rich patterns of taste heterogeneity are typically needed to explain the brand switching patterns of households in frequently purchased categories. Elrod and Keane (1995) and Keane (1997a,b) discuss how brand switching patterns tend to identify distributions of consumer taste heterogeneity. As we noted earlier, we found that a model with four taste types gave a good fit to the data in general, and to brand switching patterns in particular.

We assume that households can only purchase a single brand *j* on a given purchase occasion *t*. This is consistent with the observation that for most frequently purchased consumer goods, households rarely if ever buy multiple brands on a single purchase occasion. For each brand j, the household can choose among a discrete set of available quantities (which we will enumerate in the data section).

The budget constraint for household *i* at time *t* is:

$$(2) \qquad \sum_j P_{ijt}Q_{ijt} + D_{it}(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2) + CC_{it} + SC_{it} + Z_{it} = Y_{it}$$

where P_{ijt} is the per-ounce price of brand *j* to household *i* at time *t*, Q_{ijt} is the quantity of *j* purchased by *i* at *t*, and Y_{it} is income of *i* at *t*. A crucial point is that the per-ounce price is allowed to differ by quantity (i.e., container size). We leave the dependence of per-ounce price on quantity implicit in order to conserve on notation.

The term $\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2$ in equation (2) is the fixed cost associated with a purchase, and D_{it} is an indicator variable equal to 1 if a purchase is made (and zero otherwise). In the results section we discuss why we chose to specify the fixed cost as a quadratic in container size. The term CC_{it} is the cost associated with carrying an inventory of the storable good under analysis for household *i* during time period *t*. Finally, SC_{it} is the fixed stock out cost incurred by household *i* during time period *t* if their usage requirement exceeds their inventory. We will further define CC_{it} and SC_{it} in Section 2.3.1 below.

The fixed cost can be interpreted, for instance, as the cost of going to the store, locating the product in the store, and then carrying the container home. But regardless of the story one tells to motivate this term, its role in the model is to regulate the frequency and size of purchases. A higher fixed cost will, *ceteris paribus*, lead households to purchase less frequently, and to purchase larger sizes when they do buy.

Thus, one could also view the fixed cost as simply capturing the fact that ketchup demand is part of a larger household budgeting problem. It would be highly inconvenient (and time consuming) to buy a little bit of every product one needs each week. Even if ketchup prices were constant over time, usage rates were constant, and ketchup was available in infinitely divisible quantities, households would presumably concentrate their ketchup purchases in a small percentage of weeks in order to avoid the inconvenience of making frequent small purchases.

The role of inventory carrying costs is to provide an incentive for households to smooth inventories by spreading out their purchases over time. A higher carrying cost will, ceteris paribus, induce households to avoid buying very large quantities on single purchase occasions, or buying in consecutive or nearby weeks. A crucial distinction between the fixed cost and the inventory carrying cost is that, with high fixed costs, households want to buy infrequently. But, conditional on the total number of purchases, high fixed costs do not induce a household to care if its purchases are close together or far apart. It is only the inventory carrying cost that induces the household to want to spread purchases out over time.

In the absence of inventory carrying costs, households would tend to wait for deep discounts and then buy very large stocks of ketchup. In fact, given a positive fixed cost of purchase, a price realization close enough to the lower support point of the price distribution would induce a household to buy a lifetime supply. In contrast, in simple inventory models with constant prices and usage rates, the combination of a fixed cost of purchase and an inventory carrying cost induces an optimal inter-purchase time interval, and an optimal quantity. This generates a "saw tooth" pattern in inventories and the familiar square root purchase quantity rule (see Mellen (1925), Davis (1925)).

Finally, in a model with uncertainty about usage requirements, a stock out cost generates an incentive to hold a buffer stock, and to repurchase before inventories are too close to zero. In our model, a higher stock out cost induces stronger positive duration dependence of the purchase hazard, holding price fixed. In Appendix A we provide a more detailed discussion of how the fixed cost, carrying cost and stock out cost affect key features of the data.

Next, we derive the period utility for household *i* in week *t*. Substituting for Z_{it} in (1) using (2) we obtain:

$$(3) \qquad U_{it} = \sum_{j=1,J} \psi_{ij} C_{ijt} + Y_{it} - \sum_{j=1,J} P_{ijt} Q_{ijt} - D_{it}(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2) - CC_{it} - SC_{it}$$

Because Y_{it} enters the conditional indirect utility function given purchase of each brand *j* in the same way, Y_{it} will not affect brand choice decisions and can be ignored in the model.[9] Also note that we entered the fixed cost, inventory carrying cost and stockout cost terms in the budget constraint (2), but, as is obvious from (3), it is irrelevant whether these terms enter there or in the utility function, since utility is linear in consumption

[9] An interpretation of the fact that price enters the conditional indirect utility linearly is that the marginal utility of consumption of the outside good is constant over the small range of potential expenditures on the inside good, since these expenditures will be very small relative to Y_{it}. This type of assumption is standard in marketing studies of demand for inexpensive consumer goods. It is exactly correct because we specify that utility is linear in demand for the outside good, but is still approximately correct under more general utility specifications, provided the inside good is inexpensive.

2.3. Household Inventories

2.3.1. Preliminaries

We assume that households have an exogenous stochastic usage need for the storable commodity in each period, given by R_{it}, and that they only get utility from consumption of the good up to the level determined by the usage need, and not beyond that level. Define

$$C_{it} = \sum_{j=1,J} C_{ijt}$$

Then,

$$C_{it} \leq R_{it}.$$

The inequality allows for the possibility of stock outs, in which case consumption falls short of the desired amount. We assume that R_{it} is not revealed until after the purchase decision is made at the start of period t.

The assumption of an exogenous usage need is reasonable for many of the types of goods we are interested in, such as ketchup, toilet paper, laundry detergent, etc.. For such goods, we think it is plausible – at least to a first approximation - that consumers have a satiation point beyond which they do not derive additional utility from added consumption (e.g., you don't get extra utility from using more than the recommended amount of detergent in each load of laundry, or using more ketchup beyond the ideal amount that the kids like on their hamburgers).

Another way to phrase the assumption is that, barring a stock out, the usage rate does not depend on the inventory level. Indeed, previous work in marketing (e.g., Ailawadi and Neslin 1998) suggests that this assumption holds in ketchup (the category we will study). In other words, consumers do not put less ketchup on their hamburgers when their stock is low. Rather, they use some desired amount of ketchup until they stock out - at which point they might turn to other condiments or cease eating hamburgers for awhile.

It is worth emphasizing that the assumption of an exogenous usage need does <u>not</u> mean consumption is independent of price. If price is high for an extended period of time, the households in our model will reduce consumption by suffering more frequent stock outs – as opposed to consuming any ketchup that they have in stock at a slower rate. In other words, all adjustment of consumption to price is along the extensive rather than the intensive margin.

Rather than assuming an exogenous usage requirement, we could have instead assumed that utility is concave in consumption. In that case, if price were high for an extended period of time, households would reduce consumption by slowing down their consumption rate. More generally, the optimal current consumption rate would depend on both inventories and expected future prices.

We did not adopt such a specification for two reasons. First, we don't observe actual consumption in scanner data, but only purchases. Without consumption data, we felt that identification of the extent to which households react to price changes by altering consumption along the intensive and/or extensive margin would, at best, be very tenuously identified. In particular, both the curvature of the utility function and the stock out cost regulate the duration dependence in the purchase hazard, so their separate effects would be hard to distinguish. Second, adding a weekly continuous consumption decision would vastly increase the computational burden of solving the household's optimization problem. Thus, we felt that ignoring the intensive margin was a sensible modeling choice.

We note that in some categories, such as potato chips, ice cream or cookies, consumption rates are, presumably, an increasing function of inventories. Our assumption of an exogenous usage need would be much less palatable in such categories. On the other hand, simply introducing concave utility into our model would not be a sensible strategy in such cases either. The salient feature of such categories is "temptation" as opposed to forward-looking behavior (i.e., potato chips are technologically but not practically storable – at least for most people). So we suspect that a sensible model for such categories would be one where the consumption rate depends on the stock of the good but not on expected future prices. This would require a model with myopia or a very short time horizon.

Next, we allow the distribution of the stochastic usage requirement to be heterogeneous across consumers. Thus,

$$\log R_{i_t} \sim N(\mu_l, \sigma_l)$$

where $l=1, \ldots, 4$ and l denotes the usage type, where $l=1$ has the highest usage rate, whereas $l=4$ has the lowest usage rate. We assume that usage rate type is independent of preference type. Furthermore, we assume that a household's usage rate type may vary over time following a Markov switching process. Let π_{ii} denote the probability that a household remains type i from one week to the next, and let π_{ij} denote the transition probability from type i to type j. We assume that:

$$\pi_{ij} = \frac{1.0 - \pi_{ii}}{3} \quad \forall \; i \neq j \, .$$

This says that if a household changes type, it is equally likely to change to any of the other types. Let π_i denote the initial probability of being type *i*. In order to conserve on parameters, we assume that the initial probability is related to the family size (measured at the start of the panel) in the following way:

$$\log \pi_1 = \log \pi_{10} + 2 \; f_z \; famsize$$

$$\log \pi_2 = \log \pi_{20} + \; f_z \; famsize$$

$$\log \pi_3 = \log \pi_{30}$$

$$\log \pi_4 = \log \pi_{40} - \; f_z \; famsize$$

where *famsize* is the family size.

We also allow a stock out to carry a fixed cost. Denote by I_{ijt} the inventory that household *i* holds of brand *j* at the start of period *t*. The total inventory of all brands is given by:

(4) $$I_{it} = \sum_{j=1, J} I_{ijt}$$

Thus, if household *i* purchases Q_{jt} units at the start of *t*, its maximum consumption during period *t* is I_{it}+Q_{it}. Define

(5) $$a = \frac{(I_{it} + Q_{it})}{R_{it}}$$

If *I[a < 1]=1* a stock out occurs, where *I[ ]* denotes an indicator function for the event within the brackets.

The stock out cost to household *i* in period *t* has a constant component, as well as a component proportional to the magnitude of a shortfall, and is given by:

(6) $$SC_{it} = s_0 I(R_{it} > C_{it}) + s_1 [R_{it} - C_{it}] I(R_{it} > C_{it})$$

where s_0 is the fixed cost and s_1 is the per unit cost.

We further assume that the cost of carrying inventory is given by:

(7) $$CC_{it} = c_1 \bar{I}_{it} + c_2 \bar{I}_{it}^2$$

where $\bar{I}_{it}$ is the average inventory level during period t, which is given by:

$$(8) \qquad \bar{I}_{it} = [I_{it} + Q_{it} - \frac{R_{it}}{2}] I[a \geq 1] + [a \frac{I_{it} + Q_{it}}{2}] I[a < 1]$$

and where c_1 and c_2 are linear and quadratic terms in the average inventory level. Note that the construction of $\bar{I}_{it}$ depends on whether or not a stock out occurs during the period. If there is no stock out ($a \geq 1$), it is constructed assuming that usage is spread smoothly over the period. In the event of a stock out ($a < 1$), it is constructed assuming that usage is at a constant rate prior to the stock out, and that the stock is zero afterwards.

2.3.2. Evolution of Household Inventories

At any t, a household might potentially have a number of brands in its inventory. In that case, we would need to model the order in which brands are consumed within a period. This would lead to greatly increased complexity of our model, for little payoff. In most categories of frequently purchased consumer goods, consumers almost never buy multiple brands on a single shopping occasion, and brand "loyalty" is strong, so inventory holdings will not exhibit much brand heterogeneity. So, to avoid having to model the order of consumption within a period in those rare instances where it would be relevant, we assume that in period t, after the minimum usage requirement R_{it} is realized, households use each brand in their inventory proportionately to meet their usage needs.[10]

The state of a household at time t includes its time t inventories of each brand. If there are several brands, this means that the state space for the consumer's dynamic optimization problem will grow quite large. However, under the assumption that brands are used proportionately to meet the usage requirement, a household's state can be characterized by just two variables: its total inventory, as given by (4), and its quality-weighted inventory, which we define by

$$I1_{it} = \sum_{j=1,J} \psi_{ij} I_{ijt}$$

Recall from (1) that ψ_{ij} is household i's evaluation of the efficiency units of consumption provided by each unit of brand j. This is why we call $I1_{it}$ the "quality" weighted inventory.

After purchasing Q_{ijt} units of brand j, the total stock of the storable good is $I_{it}+ Q_{ijt}$, since households are assumed not to buy multiple brands in a given time period. Because of the assumption that households use each brand proportionately to meet their usage needs, if the total

[10] Note that households would be indifferent to the order in which brands of different quality are consumed if they do not discount the future. Such indifference will hold to a good approximation if the discount factor is close to one.

amount of the storable good is greater than or equal to the minimum usage requirement R_{it}, then only a fraction *1/a* of the stock of each brand is used, where *a* is given by equation (5).

Hence, if a stock out does not occur, then, using (3), (5) and (7), the utility of household *i* in period *t*, conditional of the purchase of Q_{ijt}, can be written as:

$$(9) \qquad U_{it} = \frac{I1_{it} + \psi_{ij} Q_{ijt}}{a} + Y_{it} - P_{ijt} Q_{ijt} - c_1 [\frac{a(I_{it} + Q_{ijt})}{2}] - c_2 [I_{it} + Q_{ijt} - \frac{R_{it}}{2}]^2$$
$$- D_{it}\left(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2\right)$$

In this case, the inventory of household *i* in the following period *t+1* will be

$$(10) \qquad I_{it+1} = I_{it} + Q_{ijt} - R_{it}$$

and the quality-weighted inventory will be

$$(11) \qquad I1_{it+1} = [I1_{it} + \psi_{ijt} Q_{ijt}][1 - \frac{1}{a}]$$

However, if the total amount of the storable good, $I_{it}+ Q_{ijt}$, is less than the minimum usage requirement R_{it}, all the inventories are used and a stock out occurs. In this case the utility of household *i* in period *t* can be written, using (3), (5), (6) and (7), as:

$$(12) \qquad U_{it} = I1_{it} + \psi_{ij} Q_{ijt} + Y_{it} - P_{ijt} Q_{ijt} - c_1 [\frac{a(I_{it} + Q_{ijt})}{2}] + c_2 [\frac{a(I_{it} + Q_{ijt})}{2}]^2$$
$$- D_{it}\left(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2\right) - [s_0 + s_1 [R_{it} - (I_{it} + Q_{ijt})]]$$

Due to the stock out, both I_{it+1} and $I1_{it+1}$ are equal to zero.

2.3.3. Identification

At this point, we have laid out all the equations of our structural model of household behavior. A formal analysis of identification is not feasible for a highly complex non-linear model like ours. However, in Appendix A we present an intuitive discussion of how the key model parameters are pinned down by patterns in the data. To summarize, note that the key structural parameters are the preference weights, ψ, the means of the log usage requirements, μ, the inventory carrying cost parameters, *c*, the fixed cost of purchase parameters, τ, and the stock out cost parameters, *s*. The discussion in Appendix A includes simulations that show how changing each of these parameters leads to different types of effects on household behavior, suggesting that each parameter is separately identified. An exception is the linear term in inventory carrying costs, c_1. As we describe in Appendix A, this has almost identical effects on behavior as the linear term in the fixed cost of a purchase, τ_1. Thus, we fixed $c_1=0$.

2.4. The Price Process

A key component of our model is the vector stochastic process for the prices of each brand/size combination. In order to have confidence in our model's predictions of how price expectations affect brand and quantity choice dynamics, it is important that our assumed price process be realistic. Thus, our price process must capture three important features that are typical of observed price data for most frequently purchased consumer goods: 1) prices typically are constant for several weeks, followed by jumps, 2) the probability and direction of jumps depends on competitor prices, and 3) the direction of jumps depends on own lagged price. To capture these features of the data we specify the multivariate jump process described below.

A key problem that we face is that the number of brand/size combinations is very large for the typical frequently purchased consumer good (e.g., in the case of ketchup it is 16). And per ounce prices for the same brand typically differ across sizes. This creates two problems. First, it is not feasible to estimate a vector price process including each of the 16 brand/size combinations, because of the substantial proliferation of parameters that would be entailed (i.e., consider the size of the variance/covariance matrix of the vector of price innovations). Second, if the price process exhibits persistence, so that current prices alter expected future prices, the expected value of the household's next period state will depend on the current price of each brand/size combination. Thus, we must keep track of an infeasibly large number of state variables when solving the household's dynamic optimization problem.

To arrive at a practical solution of this problem, we exploit a common feature of most frequently purchased consumer goods categories. In most categories, there is one clearly dominant (or most popular) container size. That is, the large majority of sales are for a particular size. Thus, our solution is as follows: First, we estimate a vector process for the prices of the most common size (e.g., 32 ounces in the case of ketchup) of the alternative brands. This process captures the patterns of persistence and competitor reaction observed in the data. Second, we specify (for each brand) a process for the differentials of the per ounce prices of the "atypical" sizes relative to the most common size. We assume that the price differentials between the atypical sizes and the most common size are iid over time (except for constant mean differentials that capture the fact that per ounce prices differ systematically across sizes).

The assumption that price differentials between the atypical sizes and the common size of each brand are iid over time greatly simplifies the solution of the dynamic optimization problem. It means that the only state variables we need to keep track of are the prices of the common size of each brand. Without this assumption, the estimation of our model would be completely infeasible. In our view, the assumption is probably fairly innocuous. Since most purchases are of the most common size, value functions should not be too sensitive to prices of atypical sizes.

To proceed, we first specify the price process for the most common size of each brand, and then specify how price for atypical sizes move relative to the common size prices. The price of the most common size of brand *j*, denoted by *c*, is assumed to stay constant from one week to the next with probability π_{1jt}. That is:

$$P_{jt}(c) = P_{j,t-1}(c) \quad \textit{with probability } \pi_{1jt}, \text{ for } j=1,J,$$

where:

$$(13) \qquad \pi_{1jt} = \frac{\exp[\delta_{0j} + \delta_{1j}(P_{jt} - \overline{P}_{t-1}) + \delta_2 (P_{jt} - \overline{P}_{t-1})^2]}{1 + \exp[\delta_{0j} + \delta_{1j}(P_{jt} - \overline{P}_{t-1}) + \delta_2 (P_{jt} - \overline{P}_{t-1})^2]}; \quad \overline{P}_{t-1} = (1/4)\sum_{j=1}^{4} P_{j,t-1}$$

Thus, the probability of a price change is $\pi_{2jt} = 1 - \pi_{1jt}$. In this case, the process is posited to be

$$(14) \qquad \ln[P_{jt}(c)] = \beta_{0j} + \beta_{1j}\ln[P_{j,t-1}(c)] + \beta_2\{(1/4)\sum_{l=1}^{4}\ln[P_{l,t-1}(c)]\} + \varepsilon_{jt}$$

where the vector of price shocks has a multivariate normal distribution

$$\varepsilon_t \sim N(0, \Sigma).$$

Note that equation (13) specifies the probability of a price change as a logistic function. To capture competitive reaction, the probability that a brand changes its price is allowed to depend on the difference between the brand's current price and the mean price of the other brands. Equation (14) specifies that if prices do change they follow an autoregressive process (in logs). Competitor reaction is captured in (14) by the parameter β_2 that multiplies the mean (log) price of the competitor brands.

Finally, the price process for the atypical sizes is specified as:

$$\ln P_{jt}(z) = b_{1j}(z) + b_{2j}(z)\ln P_{jt}(c) + v_{jt}(z),$$

where *c* again indicates the common size and *z* indexes the atypical sizes. We also assume

$$v_{jt}(z) \sim N(0, \sigma_v^2)$$

The price process parameters are estimated in a first stage using the price data, prior to estimation of the choice model. They are treated as known in the second stage, at which point we plug them into the consumer's dynamic optimization problem. The vector autoregressive jump (or switching) process for prices of the common size is estimated by maximum likelihood, while the price processes for the atypical sizes are estimated by OLS regression.

In the first stage we estimate the price process faced by a typical household, which is subtly different from the price process that exists in particular stores. To estimate the price process for a particular brand/size, we first construct the price history for that brand/size that was faced by each individual household over the weeks of our sample period. We then pool these household specific price histories together in the estimation. Thus, variation in price due to uncertainty about which store will be visited in the next period is subsumed in the household level price process that we estimate.

To justify this approach, we assume that the sequence of stores visited by a household over successive weeks is determined by a process that is exogenous to the brand and quantity choice process. This exogenous random variation in the store visited from week-to-week leads to mixing of the store level price processes, thus generating an additional source of variation in the prices a household faces. This assumption of exogeneity of the store visit process would probably not be a good assumption for big ticket items (say diapers) where price advertising might influence the store one visits. But we doubt that this is an important factor for inexpensive items like ketchup.

Our model makes the strong assumption that consumers observe the price process realizations each week. We considered two types of alternatives to this basic model. One is a model in which consumers only see prices and can only make purchases in the weeks in which they visit a store. Then the dynamic optimization problem can be simply modified by specifying a weekly probability of a store visit. An agent at time t who is in a store and observing a set of prices must take into account probability he/she might not visit a store next week (and therefore won't be able to make a purchase or see prices next week) when deciding whether to purchase at time t. But we found that this model produced essentially identical results to our model, because the large majority of households visit a store in the large majority of weeks.

A second more extreme alternative is to assume that consumers only see prices in the weeks they actually purchase the good. This could be rationalized by a model in which consumers first decide whether to buy the good in a given period, and only then go to the store and observe prices. But we reject this option out of hand, because such a model could not possibly explain the purchase acceleration effects that are clearly present in the data.

Having completely described our model, we can provide some discussion of how it differs from that of Hendel and Nevo (2002). Their model is in many ways similar to ours, but a key difference is that they specify utility as a concave function of consumption and do not have a stock out cost parameter. In this framework, a high marginal utility of consumption near zero would induce consumers to try to avoid stock outs. They also assume that the utility from a brand is derived entirely at the moment of purchase. Hence, a household's state depends only on its total inventory (and not how it is allocated among different brands). This assumption allows Hendel and Nevo to achieve a separation of the brand choice and quantity choice problems - households solve a dynamic optimization problem to choose optimal quantity each period, and then choose brands (conditional on quantity) in a static framework.[11]

While the Hendel-Nevo approach leads to an important computational simplification, this of course comes at some cost. The complete separation of the brand and quantity choice problems breaks down if there is unobserved taste heterogeneity. In that case, the distribution of brand preferences in the selected sample of consumers who chose to buy a positive quantity in any given period will, in general, differ from population distribution of brand preferences (in a way that depends on prices). As we discussed in section 1, this is a source of bias in any estimation of price elasticities of demand based on static choice models. The Hendel and Nevo approach is likely to be most efficacious for categories in which the relation between usage rates and inventory is a first order problem while flexible modeling of unobserved consumer heterogeneity is of second order importance. In contrast, estimation of our model is more computationally demanding. But the main advantage of our approach is that we can easily accommodate unobserved heterogeneity.

It is worth noting that unobserved heterogeneity in brand preferences can have important implications for how consumers optimize in the presence of inter-temporal price variation. To give just one example, consider a consumer who is very "loyal" to a particular name brand.

[11] Taken literally, this assumption implies that brands are identical in attribute space (so they all generate the same utility when consumed), but that households' perceptions of brands alter which brands they like to purchase. Such perceptions might be generated by "persuasive" or "image" advertising. However, if the discount factor is close to one, then to a good approximation it is irrelevant whether brands deliver different flow utilities when consumed, or if the expected present value of the brand specific flow utility is received at the time of purchase.

Suppose he/she is low on inventory, and faces a situation where current prices are high for his/her preferred name brand. This consumer has an incentive to buy a small quantity of the inexpensive store brand in order to tide him/herself over until a future time when the price of his/her favorite name brand is lower, anticipating that he/she can "stock up" on the favorite brand at that time.[12] Such "stop gap" purchase behavior depends crucially on unobserved heterogeneity that generates a strong preference for a particular name brand.

For instance, in the above example, a different consumer who was not "loyal" to a single name brand, but who preferred all name brands about equally, would not buy the store brand as a stop gap measure unless all name brand prices were high. Such a consumer would be much more likely to switch among the store brands as their prices fluctuate over time.

2.5. The Household's Dynamic Programming Problem

The household's optimal purchase timing, brand choice and quantity decisions can be described by the solution to a dynamic programming problem (see, e.g., Rust (1987), Pakes (1987), Wolpin (1987), Eckstein and Wolpin (1989), Erdem and Keane (1996)) with inventory I_{it+1}, quality weighted inventory $I1_{it+1}$ and prices of the common size, P_{tj} for $j=1,J$, as the state variables.[13] We assume that households solve a stationary problem.[14]

Households are assumed to make their purchase decisions after they observe the prices at period t but before they observe their period t usage requirement (R_{it}).[15] Now let us define the value function associated with the purchase of brand j and quantity Q before the realization of the usage requirement to be

$$V_{jQt}(I_{it}, I1_{it}, P_t) = E_{R_t} V_{jQt}(I_{it}, I1_{it}, P_t, R_{it}) + \tfrac{1}{\gamma} e_{it}(j,Q)$$

where $e_{it}(j,Q)$ is a stochastic term known to the household at the time of purchase but not observed by the analyst. To obtain multinomial choice probabilities (see McFadden (1974), Rust

---

[12] In the data we examine, the store brand is indeed bought in small quantities much more commonly than the name brands. This is precisely the mechanism our model uses to explain this phenomenon.

[13] In describing the households' problem, we suppress the dependence of the value functions on household type, which depends on preference type and usage rate type. We also suppress the dependence of price on the household i that arises because different households shop in different stores.

[14] As described in Appendix B, we obtain a stationary solution for the value functions by artificially assuming a terminal period where all value functions equal zero, and then backsolving from that period until the state specific value functions converge to a fixed point.

[15] As described in section 2.2.1, the usage rate is stochastic for two distinct reasons. Conditional on the household's usage rate type, there is an iid stochastic shock to the usage rate each period. But also, the household's usage rate type varies over time according to a Markov process. We subsume both types of uncertainty when we take the expectation over the usage rate realization.

(1987)), this error term will be assumed extreme value and i.i.d. distributed. γ denotes the inverse of the scale factor, which is proportional to the standard deviation of the extreme value distributed error term.

The value function associated with the above problem for household *i* at period *t* is

$$V(I_{it}, I1_{it}, P_t) = E_{e_{it}} Max_{j,Q} \{E_{R_t} V_{jQt}(I_{it}, I1_{it}, P_t, R_{it}) + \tfrac{1}{\gamma} e_{it}(j,Q)\} \tag{15}$$

In writing the alternative specific value functions $V_{jQt}(I_{it}, I1_{it}, P_t, R_{it})$ there are two cases to consider. First, if $I_{it} + Q_{ijt} > R_{it}$ there is no stock out. In that case, using (5), (8) and (9) and applying Bellman's principle, the value function associated with brand *j* for household *i* at time period *t* is

$$\begin{aligned} V_{jQt}(I_{it}, I1_{it}, P_t, R_{it}) = {} & (I1_{it} + \psi_{ij} Q_{ijt}) \frac{R_{it}}{I_{it} + Q_{ijt}} + (Y_{it} - P_{jt} Q_{ijt}) - c_1 \bar{I}_{it} - c_2 \bar{I}_{it}^2 \\ & - D_{it}\left(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2\right) + \beta E_{P_{t+1}} V(I_{it+1}, I1_{it+1}, P_{t+1}) \end{aligned} \tag{16}$$

with $I_{i,t+1}$ given by (10) and $I1_{i,t+1}$ given by (11).

If $I_{it} + Q_{ijt} < R_{it}$ there is a stock out. Then, using (8) and (12), the value function is:

$$\begin{aligned} V_{jQt}(I_{it}, I1_{it}, P_t, R_{it}) = {} & I1_{it} + \psi_{ij} Q_{ijt} + Y_{it} - P_{jt} Q_{ijt} - c_1 \bar{I}_{it} - c_2 \bar{I}_{it}^2 \\ & - D_{it}(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2) - [s_0 + s_1[R_{it} - (I_{it} + Q_{ijt})]] \\ & + \beta E_{P_{t+1}} V(I_{it+1}, I1_{it+1}, P_{t+1}) \end{aligned} \tag{17}$$

and next period inventory levels will be such that I_{it+1}=0 and $I1_{it+1}$=0. Note that V_{0t}, the value of the no purchase option, is obtained just by substituting Q_{ijt} = 0 in either equation (16) or (17).

Equations (16) and (17) capture the notion that households may not make the choice that maximizes the expected time *t* payoff, but rather will also consider the consequences of their time t decisions for expected future payoffs. For example, if a household expects that a substantial price cut for their favorite brand is likely at *t+1*, it may be optimal to make no purchase at *t*, even if this means running a high risk of a stock out, because it is optimal to try to arrive at *t+1* with inventories as low as possible. On the other hand, if a substantial price cut for a favorite brand occurs at *t*, it may be optimal to buy heavily - thus incurring substantial carrying costs at *t* and in the near future - due to the expected utility flow from consuming the brand over the next several periods.

We are now in a position to write out the probability that a household chooses to buy a particular brand/size combination conditional on its state (which includes inventories and the current price vector). Denote by d_{ijt} an indicator equal to 1 if household *i* buys brand *j* at time *t*, and equal to 0 otherwise, and let $d_{i0t}=1$ denote the no purchase option. Since we have assumed that the alternative specific taste shocks $e_{it}(j,Q)$ in equation (15) follow an i.i.d. extreme value distribution, the probability that household *i* purchases brand *j* in quantity *Q* at time *t* is given by a multinomial logit type expression:

$$(18) \qquad \Pr ob[(d_{ijt}=1,Q_{ijt})\,|\,I_{it},I1_{it},P_t]=\frac{\exp\left\{E_{R_{it}}\left[V_{jQt}\left(I_{it},I1_{it},P_t,R_{it}\right)\right]\right\}}{\sum\limits_{l=0,..,J;Q}\exp\left\{E_{R_{it}}\left[V_{lQt}\left(I_{it},I1_{it},P_t,R_{it}\right)\right]\right\}}$$

With regard to the summation in the denominator, *Q* must belong to a discrete set of available sizes, which may in general be different for every brand *j*. Also note that $l=0$ corresponds to the no purchase option, and there is slight ambiguity in notation because in that case *V* does not have a *Q* subscript. Finally, note that the probability of no purchase is obtained by substituting V_{0t} for V_{jQt} in the numerator of (18).

2.6. The Solution of the Dynamic Programming Problem

Given the very large number of points in the state space, we do not solve for the value function at each point. Instead, following Keane and Wolpin (1994), we evaluate the value function only at a finite grid of points, assigned randomly over (*I*, *I*1, *P*) space. We then fit polynomials in (*I*, *I*1, *P*) to the values on these grid points, and use them to interpolate the value function at points outside the grid points.

Using a polynomial in state variables to approximate the value function has an additional advantage: the integrations of the value function with respect to price shocks that appears in (16) and (17) can be done separately for each polynomial term in price that appears in the approximation. These integrations can be done analytically, since the price shocks in (14) are normal. Also, the integration with respect to the usage requirement shocks that appears in (15) can be done simply using quadrature integration. We describe the details of the solution of the dynamic programming problem and of our approximation methods in Appendix B.

2.7. The Likelihood Function and the Initial Conditions Problem

In our model, household choices are stochastic from the perspective of the econometrician for four reasons: The econometrician does not observe a household's preference type, its usage rate type, or its inventory levels. And furthermore, the econometrician does not observe the idiosyncratic extreme value distributed taste shocks for brand-size combinations. The choice probabilities given in equation (18) assume that only the taste shocks are not observed by the econometrician. However, we need to form choice probabilities by integrating over all the state variables that are unknown to the econometrician. Thus, we also need to integrate over the latent taste types, usage rate types and inventory levels.

Note that we face an initial conditions problem since we do not know the inventory levels of households at the start of the data set (see Heckman (1981)). We integrate out the initial conditions in the following way: We assume that the process had a true start that occurred t_0 periods prior to the start of our data, so that households had zero inventories at that point. Call this t=1. Our model specifies probabilities that each household is each usage rate type at t=1 (these were denoted π_1 through π_4 in section 2.3.1). Conditional on an initial usage rate type and a preference type, we simulate the household's purchase and consumption process for t_0 weeks (this requires us to draw prices, usage rates and usage rate types), bringing us up to the start of the observed data.[16] Call the first period of observed data $t=t_0+1$. Doing this M times, we obtain M simulated initial inventory levels and initial usage rate types. This process is repeated for each of the L possible initial usage rate and preference types, and for each household in the data. Thus, for each household we get $L{\cdot}M$ draws of initial inventories. In our application, we set M=10 and t_0 = 246 (which is equal to twice the number of weeks of observed data). Also note that $L=4{\cdot}4=16$.

Suppose that we observed consumption of households during the sample period, which runs from $t=t_0+1$ to $t=T$. Then we could form the simulated likelihood[17] of household i's observed choice history as follows:

$$\mathrm{L}_i = \log\left\{\sum_{k=1}^{K}\omega_k \sum_{l=1}^{L}\pi_l\left[\frac{1}{M}\sum_{m=1}^{M}\prod_{t=t_0+1}^{t_0+T}\Pr ob\left(d_{ijt}^0, Q_{ijt}^0 \,\middle|\, I_{it}(I_{lkit_0}^m), I1_{it}(I_{lkit_0}^m, I1_{lkit_0}^m), l_{lit_0}^m, P_{it}^0, \Psi_k\right)\right]\right\}$$

where d_{ijt}^0 denotes the observed choice for household i at time t, Q_{ijt}^0 denotes the observed quantity for household i at time t, and P_{it}^0 denotes the price vector faced by household i at tim e t.

---

[16] To draw prices we use a block bootstrap in which we sample 10 week long sequences of prices from the actual price data. This was done in an attempt to retain the serial correlation properties of prices present in the data.
[17] See Keane (1993, 1994) for a discussion of simulated maximum likelihood methods for discrete panel data.

ω_k is the population proportion of taste type *k*, π_l is the probability the initial usage rate type is *l*, and Ψ_k is the vector of taste parameters for taste type *k*. Here, $I_{it}(I^m_{lkit_0})$ denotes the inventory level of the household *i* at time t, conditional on simulation *m* of the initial inventory level and type, as well as on the households choice and consumption history up to time t. The object $I1_{it}(I^m_{lkit_0}, I1^m_{lkit_0})$ is similarly defined. $l^m_{lit_0}$ is the usage rate type for the household at t_0, according to simulation *m* and conditional on draw *l* for the initial usage rate type.[18]

Unfortunately, we also face a problem of unobserved endogenous state variables, because we do not observe households' usage rate realizations (or equivalently, their consumption levels) even during the sample period. Thus, even if we knew the initial inventory level at the start of the observed data, we could not construct in-sample inventory levels. We deal with this problem by simulating each of the *M* inventory histories constructed above forward from $t=t_0+1$ to $t=T$. Then we form the simulated likelihood contribution for household *i*'s observed choice history as follows:

$$\mathsf{L}_i = \log\left\{\sum_{k=1}^{K}\omega_k \sum_{l=1}^{L}\pi_l\left[\frac{1}{M}\sum_{m=1}^{M}\prod_{t=t_0+1}^{t_0+T}\Pr ob\left(d^0_{ijt}, Q^0_{ijt}\,\middle|\, I^m_{lkit}, I1^m_{lkit}, l^m_{lit}, P^0_{it}, \Psi_k\right)\right]\right\}$$

where I^m_{lkit} is the inventory level at time *t* for household *i* of taste type *k* and initial usage rate type *l*, according to the *m*th draw sequence. The quality weighted inventory $I1^m_{lkit}$ is defined similarly. And l^m_{lit} denotes household *i*'s usage rate type at time t according to draw *m* and conditional on initial type *l*.

3. Data Description

We estimate the model introduced in Section 2 on A.C. Nielsen scanner panel data from Sioux Falls, SD. The data set contains 2797 households and covers a 123 week period from mid-1986 to mid-1988. Every market of any significant size in the city of Sioux Falls was included in the study, so that we should have fairly complete data on the purchases of the participating households.[19]

Three national brands (Heinz, Hunt's and Del Monte), together with Store brands, capture more than 96% of total sales in this market. We therefore restricted the analysis to these

[18] In writing the likelihood, we have left the integration over the latent usage rate types from time t_0+1 through t_0+T implicit.

[19] We will of course miss purchases that were made out of town.

four brands, and eliminated households that bought other, minor, brands. Among these four, Heinz is clearly the dominant brand, with roughly a 66% share of all purchases, followed by Hunts at 16%, Del Monte at 12% and Store brands at 5%. The sizes available are 14, 28, 32, 40 and 64 ounces. But Hunt's is not available in 14 and 28 ounce sizes, and the Store Brands are not available in the 40 and 64 ounce sizes. Of all the 16 available brand/size combinations, Heinz 32-ounce is the market share leader with 36% share.

We wanted to limit the sample to households who are regular ketchup users because it seems unlikely that our model would be relevant for households who are not regularly in the market. A careful inspection of the data revealed that some households would be heavy ketchup users for several months, and then seem to never purchase again. We are uncertain if this is because these households actually stopped buying ketchup, or perhaps because of some problem with the data.[20] In order to obtain a sample of households who appeared to be regular ketchup buyers throughout the 123 week period, we subdivided the period into three 41 week sub-periods. Then, we took only households who bought at least once during each sub-period. This reduced the sample size from 2797 households to 996 households.

Figure 1 reports the distribution of households by total number of ketchup purchases during the 123 week period. We discovered that with only 4 usage rate types our model had difficulty simultaneously fitting the fat right tail of very heavy ketchup users, along with the large number of light users. This problem is compounded by the fact that, as we noted earlier, households usage intensity often seems to vary greatly over the 123 week period. Thus, our usage rate heterogeneity distribution has to play the dual role of explaining the dispersion in purchase frequency <u>across</u> households (Figure 1), and the heterogeneity <u>within</u> households in purchase intensity over time. Adding more usage rate types would solve the problem, but computational barriers precluded us from pursuing that course.

Hence, we decided to further screen the sample down to households who bought at least 4 times and bought no more than 16 times over the 123 week period. This further reduced the sample size from 996 to 838.

We also had to decide on which purchase quantities would be included in the choice set. As we noted, there are 5 sizes of ketchup container (14, 28, 32, 40 and 64 oz.), but households

---

[20] We speculate that some households may have moved out of Sioux Falls, but that this wasn't recorded, or perhaps that the ID cards malfunctioned for some households.

could purchase other quantities by buying multiple containers. However, we found that 7 options accounted for more than 99% of all ketchup purchases: 1) buy a single container of one of the 5 sizes, 2) buy two bottles of the 14-ounce size, or 3) buy two bottles of the 32-ounce size. Since option (2) generates a 28 oz purchase, and option (3) generates a 64 oz purchase, we decided to limit the discrete set of quantities that any household can buy to just {14, 28, 32, 40, 64}.

A feature of the data is that not every brand size/combination is available in every store in every week. Table 1 reports the sample frequencies with which each brand/size was present in the choice sets of the households in the data (conditional on the stores they visited each week). This variability in the choice sets was accounted for in both the solution of the DP problem and the construction of the likelihood for our model. We ignored this in the presentation of the model, because it would be notationally cumbersome. Essentially, we assume the households in our model know the probabilities in Table 1, and that they take these into account when constructing their expected value functions.

Tables 2 and 3 contain some descriptive statistics about prices. Table 2 reports the mean (offer) price of each of the 32 oz sizes in cents.[21] Note that Heinz, the most popular brand, is also the most expensive. Table 3 reports the mean price per oz differentials between the various sizes and the 32oz size. Notice that, in most cases, the 32 oz size is actually cheaper, on a price per oz basis, than the larger sizes.

4. Empirical Results

4.1. Parameter Estimates for the Price Process

Table 4 reports the maximum likelihood estimates of the parameters of the price process for the per ounce price of the 32 ounce sizes. The top panel of the table reports the parameters in the logit for the probability that price remains constant from one week to the next. The most interesting coefficient here is δ_2, the coefficient on the squared difference between own price and mean competitor price. This term is negative, indicating that when the price differential is large

---

[21] The price variable used in the estimation is the price paid before coupons (i.e., the shelf price). Including the redeemed coupon value in the price of the purchased brand would create a serious endogeneity problem. This is because we do not observe what coupons the households could have used for the brands they chose not to buy. Including the coupon value only in the price of the brand actually bought is like including a dummy for the brand purchased (interacted with coupon value) as an explanatory variable in the choice model. That is, one is including a transformed version of the dependent variable as an independent variable!! While this has often been done in scanner data research, it is clearly a serious misspecification. Erdem, Keane and Sun (1999) show that it leads to serious exaggeration of price elasticities of demand.

the contribution of this term to the logit becomes a large negative. Thus, a brand's price is less likely to stay constant if it departs greatly from competitors' prices.

The bottom panel of Table 4 reports the parameters of the autoregressive process for log per ounce prices in the event that there is a price change. Note the autoregressive coefficient on own lagged price is .4473, while the coefficient on the average price of competitors is .1482. This is again consistent with competitor reaction, since it implies that P_{t+1} tends to be higher relative to P_t if competitors' prices are higher.

Also interesting are the covariances between the price shocks. Note that the covariances among the price shocks for the three national brands (Σ_{23}, Σ_{23}, and Σ_{23}), are very small, and in two out of three cases negative. This suggests that when brands change prices simultaneously in a given week, there is no clear tendency for the prices to move in the same direction. This suggests that common demand shocks are not driving the price changes, which is consistent with our argument that price movements are largely exogenous from the point of view of consumers.

Table 5 reports the OLS estimates of the processes for how the per ounce prices of the "atypical" sizes differ from that of the common 32 ounce size. Interestingly, the fact that the slope coefficients are in many instances small suggests that these prices do not move very closely together.[22] This again suggests that brand specific demand shocks are not what drives price fluctuations.

4.2. Parameter Estimates of the Choice Model

Table 6 presents the simulated maximum likelihood estimates of our dynamic model of consumer choice behavior. Consider first the taste parameter estimates for the four taste types. These are interpretable as cents per ounce. Thus, type 1 households receive a monetary equivalent utility of $4.09 \cdot 32 = \$1.31$ from consuming a 32 oz container of Heinz. Type 1's have a clear preference for Heinz over the other three brands. And type 1's account for 51% of the population, which is consistent with Heinz' dominant position. This "loyal" type will buy Heinz almost exclusively.

Type 2's and 3's also prefer Heinz to the other brands, which illustrates just how dominate Heinz is in the this market. However, type 2's like Hunts almost as much as they like Heinz, and type 3's like Del Monte almost as much as they like Heinz. Type 2's will tend to

[22] This is a weakness of our approach, since it means our model of the price process fits the behavior of the "atypical" sizes less well than we would like. But, as noted in Section 2.4, we expect that value functions will not be too sensitive to the price process for the atypical sizes since they are bought much less frequently than the 32oz size.

switch over time between Heinz and Hunts, while type 3's will tend to switch between Heinz and Del Monte. The type 4's, who make up only about 4% of the population, like the store brand much more than do the other types. They will tend to switch between Heinz and the Store brand over time.

The next section of Table 6 contains the estimates of the parameters that characterize stock out costs, inventory carrying costs and the fixed cost of a purchase. Note that the linear inventory carrying cost term was set to zero for identification reasons, as discussed in Appendix A. The quadratic inventory term (.001157) implies that the cost of carrying a stock of 32 ounces is only about 1 cent per week. Based on this, inventory carrying costs may seem to be trivial. It should be noted, however, that the quadratic term would become important if households tried to hold very large inventories (e.g., at 100 ounces it becomes 11.6 cents per week, and at 200 ounces it becomes 46 cents per week). Thus, this parameter plays a key role in the model (i.e., simulations of our model imply that households rarely hold inventories in excess of 64 ounces, and practically never hold more than 80 ounces).

The stock out cost is about 12 cents. In contrast, the fixed cost of making a purchase of a 32 ounce size is $228 – 4.73\ (32) + 0.06\ (32)^2 = \$1.38$. This slightly exceeds the typical price of the 32oz size. This estimate seems quite reasonable if one interprets the fixed cost as consisting primarily of the utility cost (i.e., time cost) of going to the store to make the purchase. However, since we know households go to the store in the large majority of weeks, this interpretation is not "realistic." More plausibly, what the high fixed cost really captures is that it would be highly inconvenient to make frequent small purchases of ketchup and other consumer goods, rather than concentrating ones purchases for each good into a small number of weeks. Given the low stock out cost and the large fixed cost of making a purchase, it is not surprising that simulations of the model imply stock outs are very common (see Section 4.3 below).

Our estimates of the quadratic in container size implies that the fixed cost of purchasing the 32 oz size is lower than the fixed cost of purchasing any other size. This seems plausible, given that the 32 oz size is typically prominently displayed in the store (sometimes including end of aisle displays, in aisle displays, etc.), while other sizes may take more effort to locate. Obviously, our fixed cost parameters are capturing time and search costs, not just the physical effort involved in carry containers. It is worth noting that ketchup purchases are quite heavily concentrated at the most popular (32oz) size (see Section 4.3). Our model can generate that the

32oz is clearly the most popular size even without making the fixed cost a quadratic in size, but not to the same degree seen in the data.[23]

We turn next to a discussion of the usage rate parameters for each of the four usage rate types. Type 1's have a very high usage rate – about 23 ounces per week on average.[24] But the probability a household remains a type 1 from one week to the next is only .35. The mo del uses the type 1's to capture instances in the data where households are observed to buy large amounts of ketchup in consecutive (or nearby) weeks. We speculate that these unusual episodes are probably due to events like container breakage or instances where families throw large parties or cook outs.

Type 2's and 3's exhibit much more moderate usage rates, and also much greater persistence over time. For instance, type 2's use about 8 ½ ounces per week on average. They have a week-to-week probability of staying type 2's of .9958, which implies there is about a 20% chance they change type within a year. Type 3's use about 2 ounces per week on average.

Note that usage rate parameters for type 4's are not reported in the table. In the estimation process, the model wanted to generate one type with a very low usage rate. This enables the model to explain instances where households go several months without buying any ketchup. Thus, at some point in the estimation process we simply fixed the usage rate for type 4's at zero.

The last set of estimates, reported at the bottom of Table 6, are the probabilities that a household is each of the four usage rate types in the initial week of the data. The most common initial type is actually the zero usage rate type (e.g., 34.2% if family size is set to zero). As we would expect, the family size coefficient suggests that larger families are more likely to be the higher usage rate types (initially). The estimate implies that the probability of being a zero usage rate type drops by about 3.5% with each additional family member (e.g., the probability of being the zero usage rate type drops to 29.8% if family size is 4).

4.3. Goodness of Fit

Table 7 compares the sample choice frequencies and simulated choice frequencies for all brand/size combinations. Overall, the fit of the model is very good on this dimension. The probability that a household makes a ketchup purchase in any given week in the data is 6.768%.

---

23 This same problem – the extent of preference for the most popular size is hard to explain - has been noted in many past marketing studies. These typically invoke size specific preferences (or "size loyalty") to explain the phenomenon.

24 To obtain this figure, use the μ_1 and σ_1 from Table 6, and plug them into the formula $exp(\mu_1 + \sigma_1^2/2)$.

Simulation of our model generates a probability of 6.771%. The model also fits the brand shares extremely well. For example, the Heinz share is 66.4% and the model predicts 64.6%.

The only dimension in which the model (slightly) fails is generating the size distribution of purchases. The 32 ounce share is slightly underestimated (61% in the simulation vs. 64% in the data) and the 28 ounce share is also slightly underestimated (13.7% in the simulation vs. 17.0% in the data). Both these errors get pushed into the 40 ounce share, which is seriously overestimated (11.4% in the simulation vs. 5.4% in the data).

Obviously, inventories are unobserved, so we cannot compare the model's inventory predictions to the data. The simulation implies that households carry a mean inventory of 7.5 ounces, and that they are stocked out (have no ketchup at all) in 2/3 of the weeks. This rate may seem high, but we have no data to compare it against. Also, recall that our estimates imply that roughly 1/3 of households have zero desired usage. These households are not really "stocked out" in the standard sense of the term, but simply do not want ketchup (i.e., they bear no stock out cost). It is also useful to recall that, in our model, households adjust consumption rates along the extensive margin (i.e., percentage of weeks they have ketchup available to consume) rather than along the intensive margin (i.e., rate of ketchup consumption in weeks when it is available). Thus, price changes effect consumption via their effect on stock out frequency, and the stock out frequency is therefore closely related to the price elasticity of demand.

An interesting aspect of the data is that the share of the 14 ounce size in total brand sales is much greater for the Store brand (29%) than for the name brands (7% for Heinz and 4% for Del Monte). The model captures this pattern quite well, despite the fact that there is no specific parameter that could pick it up (i.e., we do not have brand/size specific taste parameters). Thus, the pattern is generated by the basic structure of our behavioral model itself. We can show in a greatly simplified version of our model that if a household has low inventory (i.e., it is at risk of stocking out) and the prices of preferred name brands are high, it is optimal to buy a small amount of the cheapest brand as a stop gap measure while waiting for prices to fall. This basic mechanism presumably carries over to the more complex model estimated here.

Figures 2, 3 and 4 provide evidence on how the model fits choice dynamics.[25] Figure 2 compares the simulated and actual distributions of inter-purchase times (in weeks). The main

---

[25] Of course, in the data, some spells are left or right censored. To make the simulations of the distribution of interpurchase times, the survivor function and the hazard comparable to those in the data, we imposed the same

failure of the model is that it somewhat underestimates the frequency of very short inter-purchase spells. For instance, the percent of the time that people buy again in just one week is 3.8% in the data vs. 2.7% in the simulation. By four weeks the model predicts 5.3% vs. 5.9% in the data. But other than that, the agreement of the simulated and actual distributions is quite impressive. The modal inter-purchase time is 6 weeks, and this is correctly predicted by the model. The model is also accurate regarding the amount of mass in the vicinity of the mode. In data, 18.8% of spells are in the 5 to 7 week range, compared to 17.7% in the simulation. The model (very) slightly overestimates the percent of inter-purchase spells in the 8-22 week range, and is quite accurate for spells of over 22 weeks.

Another way to look at the data is to look at the survivor function for no-purchase spells, which is reported in Figure 3. Here, the agreement between the model and the data is quite good. Consistent with the observations made above, the simulated survivor function from the model is slightly above that in the data in weeks 1-16, because the model predicts too few short spells. And the simulated survivor function drops a bit below the data in the 21-37 week range – because too many spells are predicted to end in that range. But the divergence between the data and simulated survivor functions is never more than a few percent.[26] In the data the survivor function first drops below 50% at 10 weeks (i.e., 47.8% of no-purchase spells survive more than 10 weeks). The model survivor function implies that 50.2% of spells survive past 10 weeks, and dips below 50% at 11 weeks. The model and data survivor functions both drop below 20% at 18 weeks.

Figure 4 reports hazard rates for the hazard of making a purchase. Again the empirical and simulated hazard rates line up quite well. The hazard rate for the data is rather jagged due to noise, especially after about 30 weeks, since less than 10% of all no purchase spells survive that long (see the survivor function). The model predicts that purchase hazard is quite low immediately after a purchase, and then rises to the vicinity of 8% after about 7 weeks. It then

---

censoring on the simulated data. However, we found that this led to only trivial changes in the simulated distributions. This contrasts with the usual experience with unemployment duration data, where truncation typically has large effects. The reason for the difference lies in the different nature of these two types of data. In unemployment spell data, the sample usually consists of people who became unemployed in a particular week. Thus, the finite length of the sampling period leads to right censoring of longer spells. In our data, in contrast, the sample begins at random points during no-purchase spells of households. And the sampling frame of nearly three years is long relative to even the longest no-purchase spells. This means that short spells are just as likely to be right censored as long spells when the data set ends.

[26] The maximum divergence is at week 7. In the data 61.9% of no-purchase spells survive past week 7, and the model predicts 65.8%.

stays fairly flat at that level regardless of spell length. Note that empirical hazard is very similar to the simulated hazard up through week 16, and by that point over 70% of spells are ended (see survivor function). The empirical and simulated hazards diverge a bit after week 16. The difference is that, while the simulated hazard stays near 8%, in the actual data the hazard sags to the 6-7% range in weeks 16-32, and after week 35 it averages around 10%.

Table 8 reports on how the model fits the distribution of accepted (per ounce) prices. The top two panels of the table contain mean offer prices from the data vs. simulation of the model. These are virtually identical. The bottom two panels contain mean accepted prices from the data vs. the simulation. The mean price for each brand/size combination is reported as a price per ounce. For example, for Heinz, the mean offer price in the data is 3.596 cents per ounce for the 32 ounce size, or $1.15. The mean accepted price is $1.12. In the simulation, these figures are $1.15 and $1.11, respectively. Note that mean accepted price is only a few cents below mean offer price, which is consistent with the fact that a large fraction of consumers have a strong preference for Heinz, thus isolating it from strong price competition.

As we would expect, differentials between offer and accepted prices are generally much larger for Hunts, Del Monte and the Store Brand. For example, for Del Monte, the mean offer price in the data is $1.05 and the mean accepted price is 96 cents. The simulation also generates predictions of $1.05 and 96 cents, respectively. For the Hunts 32oz, the offer/accepted differential is about 5 cents in both the data in the simulation. Overall, the fit of the model to the accepted price distribution is remarkably good.[27]

Finally, Table 9 compares the brand transition matrix in the data vs. that generated by simulation of our model. Some features of the transition matrix for the data are quite striking. First, note that a household that buys Heinz on a given purchase occasion has a 79% probability of buying Heinz again on the next purchase occasion. But the pattern is strikingly different for the other three brands. For example, a household that buys Del Monte on a given purchase occasion has only a 34% probability of buying it again on the next purchase occasion. It actually has a higher probability of buying Heinz (41%). Indeed, Heinz is so dominant in this market that this basic pattern holds for all three alternative brands. In general our model fits the transition matrix quite well, except that we understate the own transition rate for Del Monte by a third.

---

[27] The only exceptions are the Heinz 64 oz and the Del Monte 40 oz. For each of these, the model substantially under-predicts mean accepted price.

4.4. Policy Experiments

Our model could potentially be used to study the impact of a multitude of possible policy experiments. In this section we report the results of two types of experiment that are of particular interest. First, we discuss a transitory price cut experiment. This experiment is aimed at evaluating the importance of price expectations in the determination of own and cross price effects on demand. Second, we evaluate the effects of three types of permanent changes in pricing policy: a permanent reduction in mean price, a permanent reduction in price variability, and a simultaneous reduction in both the mean and variance of prices.

4.4.1. Effects of Transitory Price Changes: Evaluating the Importance of Expectations

In our first experiment we simulate the effect of a 10% temporary (i.e., one week in duration) price cut for all sizes of the leading brand, Heinz. This change in price at time t will alter the expected future prices of Heinz, and all the other brands. Using our model, we can simulate the "total" effect of the temporary price cut on demand, which includes this change in expectations. We can also calculate an "expectations fixed" effect, in which households do not update their forecast of future prices when Heinz changes its time t price. To implement this, we simply use the original (rather than the reduced) Heinz price when constructing the future components of the alternative specific value functions given by (16) and (17).

To conduct the experiment, we first generate 10,000 simulated price histories that last 246 weeks (twice the sample period in our data). Four each of the four taste types, we then simulate the behavior of 10,000 households, each facing one of these price histories. The 40,000 simulated households are then weighted according to our estimates of the population type proportions. Details of the simulation procedure are presented in Appendix C.

We start each household with zero inventories at $t=1$, just as when we integrate out the initial conditions in simulating our likelihood function. We simulate the effect of a price cut at week 80, under the assumption that the distribution of inventories would have converged to the stationary distribution by that point. This leaves 167 weeks over which to trace out the impulse response to the price cut. By using 10,000 simulated price histories, we essentially integrate over the distribution of initial prices and inventories that exist at the time of the price cut, as well as over the distribution of price changes (for Heinz and other brands) that occur after the price cut.

Table 10 reports the effects of the temporary price cut on purchase probabilities for Heinz and all other brands in the week of the price cut, week 1, and in subsequent weeks through week

15. The table reports percentage changes in the number of purchases. We found that after week 15 effects on demand were trivial, so we do not report them. Under each brand heading in the table, the first column reports the "total" effect, and the second column reports the "expectations fixed" effect. The first column indicates that purchases of Heinz increase 41.3% percent in the week of the price cut (corresponding to an elasticity of demand of roughly –4). This very large own effect is consistent with a large body of work in marketing (using scanner data) showing large effects of temporary price cuts on demand for many frequently purchased consumer goods.

Consider now the cross-price effects. The 10% price decrease for Heinz results in decreases in demand in week 1 of about 4% for Hunts, 3.6% for Del Monte and 3.1% for the Store Brand, implying cross-price elasticties of demand in the range of .30 to 40.

Note that total demand in the category rises 25.3%. This indicates that the price cut for Heinz is not just stealing customers away from the other brands. Rather most of the increase in Heinz sales results from either "purchase acceleration" or category expansion.

Next, consider the effect of the price cut, holding expectations of future prices fixed. As we would expect, the positive effect on Heinz sales is now greater; 45.3 % compared to 41.3 % when expectations adjust. The reason is that, with expectations held fixed, given the high degree of persistence in the price process, consumers expect that at $t+1$ the price of Heinz will very likely be near its original (pre-promotion) level. Thus, there is an added incentive to purchase at time t (i.e., "make hay while the sun shines"). Still, the most striking thing about this effect is that it is rather small. The own price elasticity of demand holding expectations fixed is only 10 % greater than the elasticity when expectations adjust.

The striking result in the table is the impact of expectations on cross-price elasticities of demand. When expectations are held fixed, these are reduced by roughly 50 %. The expectations fixed cross-price elasticities of demand for Hunts, Del Monte and the Store Brand are only about .20. Two factors drive this result: 1) if Heinz' price is lowered today it leads consumers to also expect a lower Heinz price tomorrow. This lowers the value function associated with purchase of <u>any</u> brand other than Heinz today. 2) Given the price dynamics in the ketchup market, a lower price of Heinz today leads consumers to expect competitor reaction, so it lowers the expected prices of the other brands tomorrow.

Table 11 reports the results of the exact same experiment, except that there we report the effects of the price cut on quantities demanded, rather than on purchase probabilities. The basic story is exactly the same. The only additional point worth noting is that the price cut for Heinz causes both the Heinz quantity sold and the overall category quantity sold to increase by about 10% more than did the purchase incidence. This implies that, in response to the price cut, consumers are also switching to somewhat larger quantities (conditional on purchase). This is as we would expect, and is again consistent with purchase acceleration.

Finally, we report the effect of the temporary price cut on total quantity sold over weeks 1 through the end of the simulation (a period of 167 weeks). As a percentage of average weekly sales, the sales of Heinz increase 35.5%, while the sales of Hunts, Del Monte and the Store Brand decline –7.9 %, -6.9 % and –6.6 %, respectively. Overall category sales increase 20.5%.[28] Thus, it is clear that the short run increase in the level of sales due to the temporary price cut is not wiped out, even in the long run, by sales reductions in later periods. The temporary price cut not only produces purchase acceleration, but also generates some additional Heinz and category sales that otherwise would not have occurred.

4.4.2. Effects of Permanent Changes in Pricing Policy

The real strength of a structural approach to demand estimation is that we can forecast how consumer behavior will respond to fundamental changes in pricing policy. In this section we analyze three such policy changes: 1) a permanent 10% cut in the mean price of Heinz, 2) a permanent 50% reduction in the standard deviation of Heinz prices around their mean, and 3) a combined experiment where we lower both the mean and variance in Heinz prices. The results of these three experiments are reported in Table 12. For both Heinz and the competitor brands we report the percentage changes in purchase incidence, total quantity sold (i.e., sales weighted by the container size), sales revenue, and mean accepted price.

The top panel of Table 12 reports results from a permanent 10% reduction in the mean price of Heinz. This price reduction was applied to all sizes.[29] Note that the purchase frequency for Heinz increases 33.1%, while the total quantity of Heinz sales increases 35.6%. Thus the price cut generates some shift towards purchase of larger sizes. As we would expect, the long run

[28] Dividing these figures by 167, we see that percentage increase in Hunts sales over the whole 167 week period is less than 2 tenths of 1 percent. Since essentially all the change happens in the first several weeks, we see that asymptotically, as *t* grows large, the effect of the temporary price cut on the percentage increase in total sales is (of course) approaching zero.

[29] To determine the new stochastic process for prices, we reduced all Heinz prices in the data by 10%, and then re-estimated the price process of Section 4.1. Since the price process includes terms that capture competitor reaction, these parameters must adjust so that the new price process generates the same distribution of competitor prices as did the original price process (despite the lower prices of Heinz).

elasticity of quantity demanded with respect to the permanent price cut (-3.5) is less than the short long elasticity with respect to a transitory price cut (-4.5).

It is also interesting to compare short-run vs. long-run cross-price elasticities of demand. Recall from Table 11 that the short-run cross-price elasticities with respect to transitory price changes were in the range of .30 to .40. Here, we see that the long-run elasticities with respect to permanent price changes are in the range of .75 to 1.0. Thus, the long run cross-price elasticities are much greater than the short-run elasticities. Finally, note that the 10% price cut for Heinz leads to a 19.7% increase in overall category demand in the long-run. As we would expect, this is substantially less than the 31.3% short-run category expansion effect we found for a transitory price cut in Table 11.

The finding that long-run cross price elasticities of demand greatly exceed short-run cross-price elasticities is a key result of our analysis. This result implies that the degree of competition between brands is substantially greater than short-run elasticity estimates would indicate. But, in interpreting this result, it is important to bear in mind that we are not modeling competitor reaction to the permanent change in Heinz pricing policy.[30] Thus, our experiment involves permanently lowering the price of Heinz relative to other brands. Obviously, this will induce a certain degree of brand switching (from other brands to Heinz). In contrast, when we simulated a transitory price cut for Heinz, this induced consumers to expect competitor reaction, in the form of lower prices for the other brands in the future. Thus, to some extent, the transitory price cut generates delay as opposed to switching, thus dampening the cross-effect.

There is a second mechanism that also dampens the short-run cross-price effect relative to the long-run effect. Given a transitory price cut for Heinz, a "switch" to Heinz will, for households with relatively large inventories, require a "purchase acceleration" ahead of the time when they would have otherwise bought again. Such households may be deterred from switching because it will entail extra inventory costs in the short run. With a permanent price cut this mechanism is not operative – a household with large current inventories can simply delay the Heinz purchase until some future point when inventories are sufficiently run down.

In other words, given a permanent Heinz price cut, a household can raise its steady state share of Heinz purchases without being subject to any short run spike in inventories. But, with a

[30] To do so we would need to develop and estimate a market equilibrium model, which is beyond the scope of this paper. As we noted in the introduction, the technology to estimate market equilibrium models with forward-looking behavior on both the firm and consumer sides is probably several years away.

purely transitory price cut, a household whose current inventory is relatively high can only take advantage of the sale by bearing the cost of a short run inventory spike. Intuitively, households that already have a 64 oz bottle of ketchup at home will not want to buy Heinz even if it is on sale this week, because they don't want to waste more room in their kitchen cabinets on ketchup. Thus, households with high inventories are insensitive to transitory price cuts.

It is important to consider the implications these findings for conventional estimations of demand elasticities using static models. The fact that elasticities with respect to permanent and transitory price cuts differ substantially means that conventional estimates will be quite sensitive to whether the price changes present in the data under analysis are primarily persistent or transitory changes. Of course this is not a new point. For instance, Keane and Wolpin (2002) recently made a similar point regarding estimates of elasticities of various behaviors such as labor supply, fertility, marriage and welfare participation with respect to permanent vs. transitory changes in welfare benefit rules, and, much earlier, Lucas and Rapping (1969) considered elasticities of labor supply with respect to permanent vs. transitory wage changes. But it appears that previous work on scanner data has not paid serious attention to this issue.

The second panel of Table 12 reports the effects of a 50% reduction in the variance of Heinz prices. To implement this experiment, we first calculated the mean offer price for each Heinz container size. We then compressed these offer prices around the size specific means, and re-estimated the price process of Section 4.1. Note that the Heinz purchase probability falls 5.3%, while total quantity of Heinz sold drops by 6.6%. This implies there is some shift towards the purchase of smaller sizes. Heinz total sales revenue drops by 4.2%. Revenue drops less than quantity because the mean accepted price increases by 2.6%. This is as we would expect in a search model, given a reduction in the dispersion of the offer price distribution.

Our final experiment was designed to determine whether it would be possible for Heinz to increase profits by simultaneously reducing the mean and variance of prices. This corresponds to a policy of having fewer sales, while also maintaining price at a consistently lower mean level. This experiment is of some interest, because there was a widespread shift from a policy of frequent sales to a policy of "every day low pricing" (EDLP) for many frequently purchased consumer goods in the late 1980s and early 1990s (i.e., after our sample period ended).

Of course, not knowing the cost function, we can't in general determine how changes in pricing policy would affect profits. However, if we assume that production cost is only a function of total quantity sold, then we can compare profits under policies that generate equal sales quantities simply by comparing revenues. Given a 50% reduction in the variance of its offer prices, we determined that Heinz would need to reduce its mean offer price by 2.2% in order to hold the quantity of sales constant. We report this experiment in the bottom panel of Table 12.

Note that the change in pricing policy has a positive effect on Heinz revenue, which increases by ½ of one percent. Thus, our model of consumer demand does imply that "Heinz" had an incentive to try this type of change in strategy. Of course, we are abstracting from the fact that "Heinz" is not a unitary actor. Actual retail price setting for a brand involves a complex interaction between retailers, wholesalers and manufacturers. Manufacturers use systems of incentives to attempt to induce particular pricing strategies on the part of retailers. For discussions of this topic, see, e.g., Lal (1990), Neslin, Powell and Schneider-Stone (1995) and Neslin (2002). It is beyond the scope of our analysis to explain how any increase in Heinz revenues resulting from the policy change would be distributed among the various actors in the supply chain, or to consider how such a policy change might be instigated.

The Heinz pricing policy change actually reduces demand for all competing brands and for the ketchup category as a whole. So whether a retailer would have an incentive to try such a change in strategy is ambiguous. Of course, not knowing wholesale costs for the various brands, we cannot determine how the policy change affects total category profits for the retailer. Within a range of plausible estimates for markups, our estimates in Table 12 imply reduced profits from other brands but ambiguous effects for the category as a whole.[31]

The reason demand for competing brands is reduced is the very dominant position of Heinz in the market. Even the type 2 and 3 households, who account for the bulk of Hunts and Del Monte sales, respectively, actually slightly prefer Heinz. Thus, a large fraction of their sales derive from situations in which the Heinz price is relatively high. The variance reduction reduces the extent of such events.[32]

---

[31] For example, assuming all brands have a 20% markup, and that marginal cost is constant, the estimates imply an increase in Heinz profits of 3.2%, reduced profits from Hunts, Del Monte and the Store brand of –9%,-11% and –7% respectively, and a small decline of –0.4% in category profits. Assuming smaller markups for the smaller brands would easily swing the sign of the net category effect, as would assuming smaller markups in general.

[32] Of course, the variance reduction also reduces the magnitude of Heinz sales. But there is an asymmetric effect of the variance reduction because, loosely speaking, as long as Heinz price is at or below its mean, households are much more likely to buy Heinz than the competing brands anyway.

Even if we adopt the abstraction of each brand as a unitary actor, our policy experiment is also limited because it holds competitors' pricing policy rules fixed. Thus, the policy change could still be undesirable if Heinz expected it to induce competitor reactions that would adversely affect Heinz profits. At best, an analysis of demand side response alone can only tell us whether a policy change would have some potential for increasing profits given the predicted nature of consumer reaction. It cannot reveal whether a policy change would still increase profits once competitor reaction is factored in.

Nevertheless, our results in the bottom panel of Table 12 do suggest that a strategy based on reducing price variance would have offered some promise. For such a strategy to be successful, a necessary condition is that it induce a substantial increase in mean accepted price. Notice that under the experiment, consumers are predicted to buy the same quantity of Heinz but at a mean accepted price that is ½ of one percent higher. This may represent a significant percentage increase in margins.

Finally, our experiment also illustrates why a price index constructed by randomly sampling offer prices would be misleading during a period in which retailers switched to an EDLP strategy. In our experiment the mean offer price for Heinz falls 2.2%, yet the mean accepted price for Heinz rises 0.5%, and that for the category as a whole rises 0.85%. Thus, such a price index would falsely imply that the price of ketchup had fallen, when in reality the effective price of ketchup to consumers had increased.

5. Conclusion

We have shown that our dynamic model of consumer brand and quantity choice dynamics under price uncertainty does an excellent job of fitting data on consumer purchase behavior in the market for a particular frequently purchased consumer good, namely ketchup.

Our results indicate that increased brand sales resulting from a temporary price cut are mostly due to a combination of purchase acceleration and category expansion, rather than brand switching. Given the stochastic process for prices present in our data, cross-price elasticities of demand with respect to temporary price cuts are modest compared to the own price elasticities.

More generally, our work suggests that estimates of own and particularly cross-price elasticities of demand may be very sensitive to the stochastic process for price and how households form expectations of future prices. In particular, the magnitude of cross-price elasticities of demand depends not just on the similarity of goods in attribute space, but also on the extent to which changes in current prices affect expected future prices for the own brand and other brands. This in turn will depend on the price process itself, which is just another way of saying hat price elasticities of demand are reduced form, and not “structural,” parameters. These findings suggest that researchers working on merger analysis and evaluation of welfare gains from the introduction of new goods should be careful about interpreting cross-price elasticities as measures of the degree of competition between brands.

A second main finding of our work is that estimates of cross-price elasticities with respect to permanent or long-run price changes are substantially greater (i.e., by a factor of two) than estimates of cross-price elasticities with respect to transitory or short-run price changes. The short-run estimates are dampened by the presence of both inventory carrying costs and expected competitor reaction. For this reason, the long-run estimates provide a better measure of the intensity of competition between brands.

Finally, we showed how a pricing policy change that involves a simultaneous change in mean offer prices and price variability can create a substantial wedge between the change in mean offer vs. accepted prices. Thus, price indices based on sampling of offer prices can potentially be highly misleading as measures of changes in effective costs to consumers during periods when price variability is changing.

It is beyond the scope of this paper to determine if retailers confronted with the consumers in our model would choose pricing patterns with positive duration dependence in the probability of sales. Models of sales that capture this pattern, like Pesendorfer (2002) and Hong, McAfee and Nayyar (2002) are quite stylized. But the essential dynamic that these models capture is that demand is increasing in duration since the last sale, because consumer’s inventories are dwindling. This implies that the potential revenue from holding a sale is increasing over time, which, combined with appropriate assumptions about the supply side, creates at least the potential for positive duration dependence in the probability of sales. Our demand side model does imply that demand is increasing in duration since the last sale, so it may be consistent with duration dependence in pricing. On the other hand, in a static demand model, or a model without inventory carrying costs, demand is not a function of inventory, so demand cannot depend on duration since the last sale. Thus, consumer forward-looking behavior, along with storability and inventory carrying costs, appear to be essential ingredients for any realistic equilibrium model that seeks to generate positive duration dependence in the probability of sales.

References

Ailawadi, Kusum L. and Scott A. Neslin. (1998). “The Effect of Promotion on Consumption: Buying More or Consuming Faster,” Journal of Marketing Research 35 (3), 390-398.

Berry, S., J. Levinsohn and A. Pakes. (1995). “Automobile Prices in Market Equilibrium,” Econometrica 63, 841-890.

Chiang, J. (1991). "A Simultaneous Approach to Whether, What and How Much to Buy Questions", Marketing Science 10, 4 (Fall), 297-315.

Ching, Andrew (2002). "Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs After Patent Expiration,” Working Paper, Ohio State University.

Chintagunta, Pradeep. (1993). "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households", Marketing Science 12, 2 (Spring), 184-209.

Davis, Ralph C. (1925). “Methods of Finding Minimum-Cost Quantity in Manufacturing,” Manufacturing Industries, (April) 353-56.

Dubin, Jeffrey.A. and Daniel McFadden. (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption", Econometrica 52, 2 March), 345-362.

Eckstein, Zvi and Kenneth Wolpin. (1989). "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models", The Journal of Human Resources 24, 4, 562-597.

Elrod, Terry. (1988). “Choice Map: Inferring A Product Map from Observed Choice Behavior,” Marketing Science, 7 (Winter), 21-40.

Elrod, T. and M. Keane (1995). A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data. Journal of Marketing Research, 32, 1-16.

Erdem, Tülin. (1996). “A Dynamic Analysis of Market Structure based on Panel Data,” Marketing Science 15 (4), 359-378.

Erdem, Tülin and Michael P. Keane. (1994). "Decision-Making under Uncertainty: Capturing Choice Dynamics in Turbulent Consumer Goods Markets", Marketing Science 16 (1), 1-21.

Erdem, Tülin, Michael P. Keane and Baohong Sun. (1999). “Missing Price and Coupon Availability Data in Scanner Panels: Correcting for the Self-Selection Bias in the Choice Model Parameters,” Journal of Econometrics, 89, 177-196.

Gönül, Füsun and Kannan Srinivasan. (1996). “Estimating the Impact of Consumer Expectations of Coupons on Purchase Behavior: A Dynamic Structural Model.” Marketing Science 15 (3), 262-279.

Guadagni, Peter M. and John D. C. Little. (1983). "A Logit Model of Brand Choice Calibrated on Scanner Data", *Marketing Science*, 2 (Summer), 203-238.

Gupta, Sunil. (1988). "Impact of Sales Promotions on When, What, and How Much to Buy", Journal of Marketing Research 25 (November), 342-355.

Hanneman, W. Michael. (1984). "Discrete/Continuous Models of Consumer Demand", Econometrica 52, 541-61.

Hausman, Jerry A. (1997). "Valuation of New Goods under Perfect and Imperfect Competition," In T.F. Bresnahan and R.J. Gordon (eds.), The Economics of New Goods, The University of Chicago Press, Chicago IL.

Heckman, James J. (1981). "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time, Discrete Data Stochastic Process." In C. Manski and D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications. Cambridge: MIT Press.

Heckman, Jim and B. Singer. (1984). " A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," Econometrica 52 (2), 271-320.

Hendel, Igal and Aviv Nevo. (2001). "Sales and Consumer Inventory," working paper, University of California, Berkeley.

Hendel, Igal and Aviv Nevo (2002), "Measuring the Implications of Sales and Consumer Stockpiling," working paper, University of California, Berkeley.

Hong, Pilky, R. Preston McAfee and Ashish Nayyar. (2002) "Equilibrium Price Dispersion with Consumer Inventories," Journal of Economic Theory, forthcoming.

Kamakura, Wagner and Gary Russell. (1989). "Probabilistic Choice Model for Market Segmentation and Elasticity Structure", Journal of Marketing Research 26 (Winter), 379-90.

Keane, Michael P. (1993). "Simulation Estimation for Panel Data Models with Limited Dependent Variables." In G.S.Maddala, C.R.Rao and H.D.Vinod (eds.), Handbook of Statistics, Elsevier Science Publishers.

Keane, Michael P. (1994). "A Computationally Practical Simulation Estimator for Panel Data", Econometrica 62 (1), 95-116.

Keane, Michael P. (1997a). "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," Journal of Business and Economic Statistics, 15(3), 310-327.

Keane, M. (1997b). "Current Issues in Discrete Choice Modeling," Marketing Letters 8(3), 307-

322.
Keane, Michael P. and Kenneth I. Wolpin. (1994). "Solution and Estimation of Dynamic Programming Models by Simulation", Review of Economics and Statistics 76 (4), 648-672.

Keane, Michael P. and Kenneth I. Wolpin (2002). “Estimating Welfare Effects Consistent with Forward Looking Behavior, Part I: Lessons from a Simulation Exercise,” Journal of Human Resources 37 (3), 570-599.

Kopalle, Praveen, Carl F. Mela and Lawrence Marsh. (1999). “The Dynamic Effect of Discounting on Sales: Empirical Analysis and Normative Implications,” Marketing Science 18:3, 317-332.

Lal, Rajiv. (1990). “Price Promotions: Limiting Competitive Encroachment,” Marketing Science 9 (3), 267-42.

Lal, Rajiv and Ram Rao. (1997). “Supermarket Competition: The case of Every Day Low Pricing,” Marketing Science 16 (1), 60-80.

Lucas, R.E. (1976). "Econometric Policy Evaluation", in The Phillips Curve and Labor Markets. Supplement to the Journal of Monetary Economics.

Lucas, R.E. and Rapping. (1969). “Real Wages, Employment, and Inflation,” Journal of Political Economy 77 (5), 721-754.

Marschak, J. (1952). “Economic measurements for Policy and Prediction.” In W.C. Hood and T.J. Koopmans (eds.), Studies in Econometric Metho. New York: Wiley. 1-26.

Mela, Carl F., Sunil Gupta and Donald R. Lehmann. (1997). “The long-term impact of promotions and advertising on consumer brand choice,” Journal of Marketing Research 34 (2), 248-261.

Mela, Carl F., Kemal Jedidi and Douglas Bowman. (1998). “The long-term impact of promotions on consumer stockpiling,” Journal of Marketing Research 35 (2), 250-262.

Mellen, George F. (1925). “Practical Lot Quantity Formula,” Management and Administration, 10:3, 155.

McFadden, Daniel. (1974). "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zarembka (ed.), Frontiers of Econometrics. New York: Academic Press, 105-42.

Neslin, Scott, Caroline Henderson and John Quelch. (1985). "Consumer Promotions and The Acceleration of Product Purchases", Marketing Science 4, 2 (Spring), 147-165.

Neslin, Scott A., Stephen G. Powell and Linda Schneider-Stone. (1995). "The Effects of Retailer and Consumer response on Optimal manufacturer Advertising and Trade Promotion Strategies," Management Science 41 (May), 749-66.

Neslin, Scott. (2002). Sales Promotion. Cambridge: Marketing Science Institute Relevant Knowledge Series.

Nevo, Aviv. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry," Economoetrica, 69 (2), 307-342.

Pakes, Ariel. (1987). "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks", Econometric*a* 54, 755-784.

Pesendorfer, Martin. (2002). "Retail Sales: A Study of Pricing Behavior in Supermarkets," Journal of Business 75:1, 33-66.

Rust, J. (1987). "Optimal replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," Econometrica, 55(5), 999-1033.

Shoemaker, Robert W. (1979). "An Analysis of Consumer reactions to Product Promotions." In Neil Beckwith, Michael Houston, Robert Mittelstaedt, Kent Monroe and Scott Ward (eds.), Educators' Conference Proceedings, 244-8. Chicago: American Marketing Association.

Srinivasan, T.C. and Russell S. Winer. (1994). "Using Neoclassical Consumer-Choice Theory to Produce a Market Map From Purchase Data," Journal of Business and Economic Statistics 12 (January), 1-9.

Struse, Rudolph W. (1987). "Commentary: Approaches to Promotion Evaluation: A Practitioner's View Point," Marketing Science 6:2 (Spring), 150-151.

Sun, Baohong, Scott Neslin and Kannan Srinivasan. (2001). "Measuring the Impact of Promotions on Brand Switching under Rational Consumer Behavior," working paper, Carnegie Mellon University.

Varian, Hal. (1980). "A Model of Sales," American Economic Review, 70 (4), 651-659.

Ward, R.W. and J.E. Davis. (1978). "Coupon Redemption," Journal of Advertising Research 18 (4), 51-58.

Winer, Russell S. (1986). "A Reference Price Model of Brand Choice for Frequently Purchased Products", Journal of Consumer Research 13 (September), 250-256.

Wolpin, Kenneth I. (1987). "Estimating a Structural Search Model: The Transition from School to Work", Econometrica 55, 801-818.

Appendix A: Identification

Our model is too complex to for us to provide analytic results on identification, so we instead provide an intuitive discussion of how the key model parameters are pinned down by patterns in the data. We also present, in Table A1, simulations of how increasing each of the key model parameters affects key features of simulated data. These simulations are useful for understanding how various parameters have different effects, and are therefore identified.

First, we discuss the parameters that determine the inventory carrying cost (*CC*), fixed cost of purchase (*FC*) and stock out cost (*SC*). Recall that the equations for these are:

Carrying Cost: $CC = c_1 I + c_2 I^2$

Fixed Cost: $FC = \tau_0 + \tau_1 Q + \tau_2 Q^2$

Stock Out Cost: $SC = s_0 + s_1 (R - C)$

First, consider the linear component of fixed cost (τ_1) and the linear term in inventory carrying costs (c_1). If the quantity Q that a consumer buys is used at a constant rate over time (i.e., R is fixed), and/or there is no discounting, it is irrelevant whether carrying costs are spread out over the period the good is consumed (reflected in c_1), or whether the present value of carrying costs is born up front (reflected in τ_1). Thus, c_1 and τ_1 would not be separately identified.

In our model, there is discounting, and usage rates R do fluctuate over time, so the parameters c_1 and τ_1 would have subtly different effects on behaviour, but it would not be surprising if these are hard to detect. Indeed, when we tried to iterate on both parameters, we found that the likelihood was extremely flat along a locus in (c_1, τ_1) space, and that the two parameters would run off in opposite directions. Thus, we discovered that we cannot separately identify these two parameters, and so we constrained $c_1 = 0$. It is comforting that this identification problem was made obvious by our search algorithm, and that such problems did not emerge for any other model parameters.

Next consider τ_0, the constant in the fixed cost of purchase function. A large τ_0 would induce one to minimize the frequency of purchases, and to buy large quantities when one does buy. Thus, it is pinned down by data on the frequency and size of purchases.[33] It is important to note that while a high τ_0 discourages frequent purchases, it does not affect or induce duration dependence in the purchase hazard. For example, with a high τ_0 one wants to avoid having many purchases during a year, but, conditional on the total number, one doesn't care if purchases are spread out or close together. Indeed, simulations of our model, which we report in Table A1 indicate that an increase in τ_0 shifts down the purchase hazard, but has little effect on its shape.

Next, consider the quadratic terms c_2 and τ_2. These might at first appear to be subject to the same sort of identification problem that affects c_1 and τ_1. If a consumer had $I=0$, usage rate was fixed and/or there were no discounting, and furthermore, if the consumer knew that he/she

[33] Suppose we had not constrained $c_1 = 0$. Higher inventory carrying costs would also cause one to buy small quantities frequently, so as to smooth inventories over time. Thus, a reduction in τ_0 and in increase in c_1 would both lead to more frequent small purchases and hence smoother inventories. However, the former would increase overall demand, while the later would reduce it. So the likelihood would not be flat in these two parameters.

would not buy again until Q was used up, then there would be a locus of c_2 and τ_2 values that would generate equal present values of fixed plus carrying costs for a purchase Q.

However, c_2 and τ_2 are separately identified by variation in inventories and the probability of subsequent purchases. A large c_2 says one should avoid buying a large quantity if current inventory is already large and/or one thinks it is likely one would want to buy again in the near future (say, because a deal is likely). In contrast, a large τ_2 says one should avoid buying large sizes regardless of one's state. This will tend to make time between purchases shorter, leading to less positive duration dependence in the purchase hazard.

In contrast to fixed costs, higher inventory carrying costs should induce more positive duration dependence in the purchase hazard. A value of $c_2 > 0$ induces one to smooth inventories, to the extent that one wishes to avoid very high inventory spikes, but it leaves one rather indifferent to fluctuations of inventories around low levels. In other words, since $c_2 > 0$ induces a convex carrying cost function, the marginal cost of carrying inventories will be small until inventories grow quite large. Thus, purchase probability is increasing in duration since last purchase.

The simulations reported in Table A1 are consistent with these assertions. They indicate that while increases in c_2 and τ_2 both shift down the purchase hazard, the increase in c_2 makes the hazard steeper (i.e., greater positive duration dependence), while the increase in τ_2 makes the hazard flatter.

Next consider the role of stock out costs. The critical role of these parameters is to induce positive duration dependence in purchase probabilities. With a stock out cost, the probability of a purchase is increasing in duration since last purchase, holding price fixed. As we noted earlier, fixed costs of purchase cannot induce positive duration dependence, so there is no danger of confounding these parameters with the stock out cost parameters.[34] However, an inventory carrying cost also makes purchase more likely as duration since last purchase increases. But a higher inventory carrying cost reduces demand, while a higher stock out cost increases demand, so these parameters have different effects. All these statements are verified by the simulations in Table A1.

Now consider the tastes for consumption (or utility weights) ψ and the usage rate R. An increase in ψ or R each increases demand. However, they have different effects on the duration dependence of purchase probabilities. A higher usage rate R causes important changes in the duration dependence in the purchase hazard, while a higher ψ does not. As the simulations in Table A1 show, the effects of increasing usage rates on the duration dependence of the purchase hazard are rather complex. For high and medium usage rate types, the increase in R leads to less positive duration dependence (i.e., the relative frequency of short inter-purchase spells increases). But for low usage rate types the hazard increases for intermediate length spells relative to both short and long spells.

Table A1 also describes how key parameters affect accepted prices. It is interesting that most parameters seem to have negligible effects on accepted prices. It is also interesting that the utility weights have ambiguous effects. Of course, in a static model, an increase in the utility weights would unambiguously raise accepted prices. This is no longer true in a dynamic model, where consumers can search for good prices over time. One clear cut effect is that if we raise the

[34] In addition, our fixed cost parameters (constant, linear and quadratic) are also pinned down by the relative purchase frequencies for the different sizes.

Heinz utility weight for type 1 consumers (i.e., the ones who strongly prefer Heinz) it raises accepted prices for Heinz. Other effects are more complex.

Appendix B: The Solution of the Dynamic Programming Problem

In this appendix we describe the details of how we solve the dynamic programming (DP) problem faced by households in our model. The solution of the DP problem proceeds as follows. In order to construct the value function $V(I_{it}, I1_{it}, P_t)$ in equation (15) we need to construct the objects $E_{R_{it}}\left[V_{jQt}\left(I_{it}, I1_{it}, P_t, R_{it}\right)\right]$. Given those objects, the expected maximum taken in equation (15) has simple closed form. The $E_{R_{it}}\left[V_{jQt}\left(I_{it}, I1_{it}, P_t, R_{it}\right)\right]$ are expectations (over usage rate realizations) of the alternative specific value functions, which are given by equations (16) and (17). Using (16) and (17), along with (8), and letting $F(R)$ denote the cumulative distribution function of the usage rate, we obtain:

(A1) $E_{R_{it}}\left[V_{jQt}\left(I_{it}, I1_{it}, P_t, R_{it}\right)\right]$

$$
\begin{aligned}
&= \Bigg\{ \frac{\left[I1_{it} + \psi_j Q_{ijt}\right]}{I_{it} + Q_{ijt}} \frac{\int_0^{I_{it}+Q_{ijt}} R_{it}\, dF(R)}{\Pr ob(R_{it} \le I_{it} + Q_{ijt})} + Y_{it} - P_{jt} Q_{ijt} - D_{it}\left(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2\right) \\
&\quad - \frac{c_1 \int_0^{I_{it}+Q_{ijt}} \left[I_{it} + Q_{ijt} - \frac{1}{2} R_{it}\right] dF(R) + c_2 \int_0^{I_{it}+Q_{ijt}} \left[I_{it} + Q_{ijt} - \frac{1}{2} R_{it}\right]^2 dF(R)}{\Pr ob(R_{it} \le I_{it} + Q_{ijt})} \Bigg\} \Pr ob(R_{it} \le I_{it} + Q_{ijt}) \\
&\quad + \Bigg\{ I1_{it} + \psi_j Q_{ijt} + Y_{it} - P_{jt} Q_{ijt} - D_{it}\left(\tau_1 + \tau_2 Q_{ijt} + \tau_3 Q_{ijt}^2\right) \\
&\quad - \frac{c_1 \left(\frac{\left(I_{it} + Q_{ijt}\right)^2}{2}\right) \int_{I_{it}+Q_{ijt}}^{\infty} \left(\frac{1}{R_{it}}\right) dF(R) + c_2 \left(\frac{\left(I_{it} + Q_{ijt}\right)^2}{2}\right)^2 \int_{I_{it}+Q_{ijt}}^{\infty} \left(\frac{1}{R_{it}}\right)^2 dF(R)}{\Pr ob(R_{it} > I_{it} + Q_{ijt})} \\
&\quad - s_0 - \frac{s_1 \int_{I_{it}+Q_{ijt}}^{\infty} \left(R_{it} - I_{it} - Q_{ijt}\right) dF(R)}{\Pr ob\left(R_{it} > I_{it} + Q_{ijt}\right)} \Bigg\} \Pr ob(R_{it} > I_{it} + Q_{ijt}) \\
&\quad + \beta E_{R_{it}}\left[E_{P_{t+1}} V(I_{it+1}, I1_{it+1}, P_{t+1}) \middle| I_{it}, I1_{it}, P_t, Q_{ijt}\right]
\end{aligned}
$$

Note that the first term in brackets corresponds to usage rate realizations that generate no stock-out, while the second term in brackets corresponds to cases in which a stock-out does occur. The univariate integrals over R_{it} in (A1) can be done analytically.

The last term in (A1) is the future component, $E_{R_{it}}\left[E_{P_{t+1}} V(I_{it+1}, I1_{it+1}, P_{t+1}) \middle| I_{it}, I1_{it}, P_t, Q_{ijt}\right]$, which we must somehow compute. In our model, we assume that households solve an infinite horizon stationary problem. However, as computational device for solving the DP problem, we assume there is terminal period T at which the future component is exactly zero at all state points. Then, at $t=T-1$, equation (A1) takes a simple form, since the last term drops out, and we can calculate the $E_{R_{iT-1}}\left[V_{jQ,T-1}(I_{iT-1}, I1_{iT-1}, P_{T-1}, R_{iT-1})\right]$ values analytically. These can then be substituted into equation (15) to obtain values for the $V(I_{i,T-1}, I1_{i,T-1}, P_{t,T-1})$. Given these, it is straightforward to construct the future component terms in (A1) that are relevant for $t=T-2$. Given these, we can calculate the $E_{R_{iT-2}}\left[V_{jQ,T-2}(I_{iT-2}, I1_{iT-2}, P_{T-2}, R_{iT-2})\right]$ values analytically, and so on. This process of solving a finite horizon DP problem by working backward from a terminal period in which the value functions are known is called "backsolving."

Our computational procedure for solving the infinite horizon DP problem is to backsolve the finite horizon DP problem for a sufficiently large number of time periods so that the value functions at each state point become stable, meaning that they cease to change significantly as we move further back. This approach to solving infinite horizon problems is quite common. We will make the criterion for stability more precise below.

It is important to note that the state variables in our DP problem, I_{it}, $I1_{it}$ and P_t, are continuous. Therefore, in contrast to problems with a finite number of state points, it is not possible to solve exactly for $V(I_{it+1}, I1_{it+1}, P_{t+1})$ at every state point. Thus, exact solution of the DP problem is impossible, and an approximation method must be used. We therefore introduce a polynomial approximation for $V(I_{it+1}, I1_{it+1}, P_{t+1})$.

The polynomial for $V(I_{it+1}, I1_{it+1}, P_{t+1})$ is a function of total inventories, I_{it}, the ratio of quality adjusted inventories to total inventories $I1_{it}/I_{it}$, and the vector of prices of the common size of each brand, $P_{jt}(c)$ for $j=1,\ldots,4$. To be precise, we specify:

$$(A2)\quad V(I_{it}, I1_{it}, P_t) = \sum_{\substack{k=0,\ldots,K\\ l=0,\ldots,L}} \sum_{\substack{m(1),\ldots,m(4)\\ \in M(k,l)}} C_{k,l,m(1),\ldots,m(4)} I_{it}^k \left[\frac{I1_{it}}{I_{it}}\right]^l \prod_{j=1}^{4} [\ln P_{jt}(c)]^{m(j)}$$

where the $C_{k,l,m(1),..,m(4)}$ are parameters to be estimated, and $M(k,l)$ is a set whose elements satisfy the following conditions:

$$\sum_{j=1}^{4} m(j) = 0,1 \text{ or } 2 \text{ if } k=0,\ l=0;$$

$$\sum_{j=1}^{4} m(j) = 0 \text{ or } 1 \text{ if } l=0 \text{ and } k=1,2,3, \text{ or } l=1 \text{ and } k=0,1;$$

$$\sum_{j=1}^{4} m(j) = 0 \text{ if } l=1 \text{ and } k=2,3,4$$, or $l=0$ and $k=4$, or $k=0$ and $l=2,3,4$, or $k=1$ and $l=2$.

The structure of the set $M(k,l)$ is set up so as to ignore various high order interaction terms, thus achieving a more parsimonious specification. For instance, the cases where the $m(j)$ sum to 2 and $k=0$, $l=0$ correspond to squared terms in prices and interactions between each pair of prices. The requirement that $k=0$ and $l=0$ in this case means that these second order price terms are not interacted with inventories. The total number of coefficients in the approximating polynomial is 48, and the R^2 is .99.

The $C_{k,l,m(1),..,m(4)}$ parameters are estimated by OLS regression. To obtain the sample of data points on which the regression is run, we calculate $V(I^g, I1^g, P^g)$ on $G = 1000$ inventory/price grid points $(I^g, I1^g, P^g)$. To set up the grid, we first set the inventory grid points I^g to be the Chebychev quadrature points on the interval from 0 to 80. The values of quality adjusted inventories and prices at each grid point are then set as follows:

First, we generate a fraction of inventories that is allotted to each brand.[35] Given these fractions and the I^g, we can construct brand specific inventories I_j^g, $j=1,\ldots,4$. We then multiply these by the utility weights ψ_j to obtain the quality weighted inventory $I1^g$.

Second, the four (brand specific) prices are drawn i.i.d. from a uniform distribution on the interval from 35 cents to 200 cents. This exceeds the range of prices for the 32 oz size observed in the data. These prices are then divided by the standard size 32 to obtain prices per ounce.

Having defined the grid over which we calculate the value functions, we can now return to the issue of convergence of the backsolving process. We backsolve until we reach a point where, in going back one additional period, the maximum percentage change in the value functions across all grid points is less than 0.1%.

We now discuss why we use $I1_{it+1}/I_{it+1}$ as an argument in the polynomial approximation for $V(I_{it+1}, I1_{it+1}, P_{t+1})$, rather than $I1_{it+1}$ itself. One reason is that $I1_{it+1}$ is highly collinear with I_{it+1}, while $I1_{it+1}/I_{it+1}$ is not. Thus, the OLS regression we use to estimate the $C_{k,l,m(1),..,m(4)}$ is better behaved if we use the ratio.

A second reason is more subtle. The ratio specification has a computational advantage that arises because $I1_{it+1}/I_{it+1}$ does not depend on R_{it}. To see this, note that if $R_{it} \le I_{it} + Q_{ijt}$ then $I_{it+1} = I_{it} + Q_{ijt} - R_{it}$. Thus, using (5) and (11), we have:

$$I1_{it+1} = \left(I1_{it} + \psi_j Q_{ijt}\right)\left(\frac{I_{it} + Q_{ijt} - R_{it}}{I_{it} + Q_{ijt}}\right)$$

and hence:

$$\frac{I1_{t+1}}{I_{it+1}} = \left[\frac{I1_{it} + \psi_j\, Q_{ijt}}{I_{it} + Q_{ijt}}\right]$$

[35] To generate the inventory shares for each brand, we draw three uniform random numbers on the interval (0,1). Denote these by u_1, u_2 and u_3. Then numerator of the share for brand 1 is set to $u_1 u_2 u_3$, that for brand 2 is set to $(1-u_1)u_2 u_3$, that for brand 3 is set to $(1-u_1)(1-u_2)u_3$, and that for brand 4 is set to $(1-u_1)(1-u_2)(1-u_3)$. The denominator of the shares are set to the sum of the four numerators. This construction guarantees that the inventory shares of the four brands sum to one.

Thus, $I1_{t+1}/I_{t+1}$ does not depend on R_{it}.[36] We now describe why this is advantageous.

Updating (A2) by one period and substituting for I_{it+1} and $I1_{it+1}/I_{it+1}$, we obtain:

$$(A3)\quad V(I_{it+1}, I1_{it+1}, P_{t+1})$$
$$= \sum_{\substack{k=0,\dots,K \\ l=0,\dots,L}} \sum_{\substack{m(1),\dots,m(4) \\ \in M(k,l)}} C_{k,l,m(1),\dots,m(4)} \left(I_{it} + Q_{ijt} - R_{it}\right)^{k} \left(\frac{I1_{it} + \psi_j Q_{ijt}}{I_{it} + Q_{ijt}}\right)^{l} \prod_{j=1}^{4} [\ln P_{jt+1}(c)]^{m(j)}$$

Since R_{it} does not appear in the $I1_{it+1}/I_{it+1}$ term, the expectation over R_{it} is simply:

$$(A4)\quad E_{R_{it}} V(I_{it+1}, I1_{it+1}, P_{t+1})$$
$$= \sum_{\substack{k=0,\dots,K \\ l=0,\dots,L}} \sum_{\substack{m(1),\dots,m(4) \\ \in M(k,l)}} C_{k,l,m(1),\dots,m(4)} \left[\frac{I1_{it+1}}{I_{it+1}}\right]^{l} \prod_{j=1}^{4} [\ln P_{jt+1}(c)]^{m(j)} E_{R_{it}} (I_{it} + Q_{ijt} - R_{it})^{k}$$

Now, for each k, the terms $E_{R_{it}} (I_{it} + Q_{ijt} - R_{it})^{k}$ can be calculated analytically using a simple quadrature procedure.

Finally, to obtain the future component, $E_{R_{it}} \left[E_{P_{t+1}} V(I_{it+1}, I1_{it+1}, P_{t+1}) \middle| I_{it}, I1_{it}, P_t, Q_{ijt}\right]$, which is the last term in (A1), we must also take an expectation with respect to price realizations at $t+1$. We have:

$$(A5)\quad E_{P_{t+1}} E_{R_{it}} V(I_{it+1}, I1_{it+1}, P_{t+1})$$
$$= \sum_{\substack{k=0,\dots,K \\ l=0,\dots,L}} \sum_{\substack{m(1),\dots,m(4) \\ \in M(k,l)}} C_{k,l,m(1),\dots,m(4)} \left[\frac{I1_{it+1}}{I_{it+1}}\right]^{l} E_{R_{it}} (I_{it} + Q_{ijt} - R_{it})^{k} E_{P_{t+1}} \prod_{j=1}^{4} [\ln P_{jt+1}(c)]^{m(j)}$$

where, using equations (13) and (14), the last term in (A5) can be written:

$$E_{P_{t+1}}\left[\prod_{j=1}^{4} \ln P_{jt+1}(c)^{m(j)}\right] = E_{\varepsilon_t} \prod_{j=1}^{4} \Bigg([\ln P_{jt}(c)]^{m(j)} \pi_{1jt}$$
$$+ \left[\beta_{1j} + \beta_{2j}[\ln P_{jt}(c)] + \beta_{3j}(1/4)\sum_{l=1}^{4}[\ln P_{lt}(c)] + \varepsilon_{jt}\right]^{m(j)} \pi_{2jt} \Bigg)$$

The integration over realizations of the error term for prices ε_{jt} in equation (13) can be done analytically, since we assume these errors are normally distributed.

[36] Note that if $R_{it} > I_{it} + Q_{ijt}$, then $I_{it+1} = 0$ and $I1_{it+1} = 0$, so the ratio is undefined. However, the Chebychev quadrature points that we use always have $I_{i,t+1} > 0$, so this problem does not arise.

Appendix C: Simulation of the Model

This appendix describes how we simulate data from the model, both to evaluate model fit and to conduct policy experiments. The first step is to generate 10,000 simulated price histories that last 246 weeks (twice the sample period in our data). Recall that in our model there are four taste types and four initial usage rate types, giving a total of 16 types. We simulate the behavior of 10,000 households of each taste type. Each of these faces one of the 10,000 simulated price histories. Within each taste type, 2500 households are assigned to each initial usage rate type. Thus, we simulate a total of 40,000 households, 2500 for each of the 16 types. In order to form sample statistics, the simulated households are weighted according to our estimates of the population type proportions. We discard the data from the first 79 weeks.

Given a simulated price history, we simulate the choice history for a household of a particular taste and initial usage rate type as follows:

1) Assume initial inventory is zero.

2) Use equation (18) to determine the probability of each of the 17 choice options at $t=1$. These are conditional on the $t=1$ price vector P_1, the initial inventories $I_1=0$ and quality weighted inventories $I1_1=0$, and the household's initial usage rate type. Denote the set of choice probabilities by $\{p_1, \ldots, p_{17}\}$. Define $q_0 = 0$ and $q_k = \sum_{l=1,k} p_l$. Draw a uniform random variable u_1 on the interval [0,1]. Option j is chosen iff $q_{j-1} < u_j < q_j$.

3) Draw the $t=1$ usage requirement from a log normal distribution. Update inventory using equation (10) to obtain I_2, and update quality weighted inventory using (11) to obtain $I1_2$.

4) Use the usage rate type transition probabilities π_{ij} for $j=1, \ldots,4$ to draw the household's $t=2$ usage rate type. This is done using a uniform draw, following the same type of algorithm used in step (2).

5) Use equation (18) to determine the probability of each of the 17 choice options at $t=2$, conditional on the $t=2$ price vector P_2, the $t=2$ inventory levels I_2 and $I1_2$, and the $t=2$ usage rate type. A particular choice option is drawn using a uniform random draw, just as in step 2.

Steps analogous to these are repeated until a complete history is obtained.

It is worth noting that simulation of data from the model is trivial once we have solved the dynamic optimization problem and can form the conditional choice probabilities (18), because the inventories and the latent usage rate types that enter the conditioning set are fully observed along the simulated choice path. This contrasts with construction of the likelihood function, which is very difficult because inventories and usage rate type realizations are unobserved in the actual data, and therefore must be integrated out of the choice probability expressions.

It is also worth noting that, whenever we implement a policy experiment, we hold fixed the uniform and log normal random draws that determine the choice history. In that way, all changes in behavior are due to changes in the prices facing the household or changes in the choice probabilities determined by equation (18), rather than due to simulation induced noise.

Table 1: Probability of availability of various brands and sizes.

Sizes (oz)	Brands			
	Store Brand	Del Monte	Heinz	Hunts
14	0.8840	0.4268	1.0000	0.0
28	0.7060	0.7817	0.9975	0.0
32	0.8840	1.0000	0.9968	0.9968
40	0.0	0.5630	0.9968	0.6071
64	0.0	0.9264	0.9968	0.9968

Table 2: Mean price of the 32 oz. size (in cents)

Store Brand	Del Monte	Heinz	Hunts	Total
90.70	105.07	115.09	104.93	104.33

Table 3: Average % Difference in per oz. prices from 32 oz. size

Oz. Size	Store Brand	Del Monte	Heinz	Hunts	Total
14	33.43	67.88	54.66		48.97
28	43.47	51.53	37.26		43.51
40		32.44	33.55	39.86	35.03
64		-9.93	16.00	-6.76	0.00

Table 4: Estimates of the Price Process Coefficients

Parameters in Logit for Probability of Price Staying Constant		
Store Brand intercept	δ_{01}	1.829 (0.01890)
Del Monte intercept	δ_{02}	0.6170 (0.00901)
Heinz intercept	δ_{03}	0.3079 (0.00980)
Hunts intercept	δ_{04}	0.7655 (0.00814)
Store Brand slope coefficient	δ_{11}	1.139 (0.0460)
Del Monte slope coefficient	δ_{12}	2.004 (0.0327)
Heinz slope coefficient	δ_{13}	1.908 (0.0145)
Hunts slope coefficient	δ_{14}	1.577 (0.0371)
Square term coefficient	δ_{2}	-0.1453 (0.0239)
Parameters of the Autoregressive Process for Log Price Change		
Store Brand Intercept	β_{01}	0.3851 (0.00428)
Del Monte intercept	β_{02}	0.4375 (0.00415)
Heinz intercept	β_{03}	0.5068 (0.00470)
Hunts intercept	β_{04}	0.4534 (0.00455)
Slope coefficient	β_{1}	0.4473 (0.00330)
Square term coefficient	β_{2}	0.1482 (0.00516)
Variance Covariance Matrix Parameters		
Σ_{11}		0.00402 (3.22E-5)
Σ_{12}		0.00121 (5.10E-5)
Σ_{13}		0.00148 (6.37E-5)
Σ_{14}		0.00014 (4.53E-5)
Σ_{22}		0.01189 (5.71E-5)
Σ_{23}		-0.00042 (9.04E-5)
Σ_{24}		0.00218 (6.14E-5)
Σ_{33}		0.00891 (9.63E-5)
Σ_{34}		-0.00050 (6.42E-5)
Σ_{44}		0.00820 (4.93E-5)

Note: The brand subscripts are defined as follows: Store Brand=1, Del Monte=2, Heinz=3, Hunts=4.

Table 5: OLS Results for Log Prices of Atypical Sizes Relative to 32oz

Constant Terms

Size (oz)	Store Brand	Del Monte	Heinz	Hunts
14	1.409	1.642	1.814	
28	0.971	1.338	1.191	
40		1.566	1.639	1.265
64		0.445	1.143	0.614

Slope Coefficients

Size (oz)	Store Brand	Del Monte	Heinz	Hunts
14	-0.090	-0.003	-0.085	
28	0.419	0.197	0.324	
40		-0.211	-0.059	0.228
64		0.561	0.211	0.411

Standard Errors

Size (oz)	Store Brand	Del Monte	Heinz	Hunts
14	0.046	0.033	0.045	
28	0.128	0.107	0.087	
40		0.145	0.060	0.096
64		0.127	0.150	0.119

Table 6: Parameter Estimates for the Structural Model

Parameter	**Symbol**	**Estimate**	Standard Error
Utility Type 1			
Store Brand Utility Weight	Ψ_{11}	0.0373	0.108
Del Monte Utility Weight	Ψ_{12}	0.4520	0.176
Heinz Utility Weight	Ψ_{13}	4.0860	0.135
Hunts Utility Weight	Ψ_{14}	1.3924	0.174
Utility Type 2			
Store Brand Utility Weight	Ψ_{21}	0.0099	0.174
Del Monte Utility Weight	Ψ_{22}	2.2218	0.150
Heinz Utility Weight	Ψ_{23}	3.3812	0.144
Hunts Utility Weight	Ψ_{24}	3.2828	0.145
Utility Type 3			
Store Brand Utility Weight	Ψ_{31}	0.1205	0.295
Del Monte Utility Weight	Ψ_{32}	2.7410	0.188
Heinz Utility Weight	Ψ_{33}	3.0087	0.212
Hunts Utility Weight	Ψ_{34}	1.3046	0.397
Utility Type 4			
Store Brand Utility Weight	Ψ_{41}	3.2608	0.214
Del Monte Utility Weight	Ψ_{42}	0.3618	0.339
Heinz Utility Weight	Ψ_{43}	2.4825	0.199
Hunts Utility Weight	Ψ_{44}	2.8661	0.198
Type Probabilities			
Utility Type 1	Π_1	0.5148	0.025
Utility Type 2	Π_2	0.3426	0.026
Utility Type 3	Π_3	0.0996	0.020
Other Utility Function Parameters			
Precision of Utility Shocks	γ	0.03125	3.51E-4
Discount factor	β	0.99	—

Table 6: Continued

Parameter	Symbol	Estimate	Standard Error
Parameters of Stockout Costs, Inventory Carrying Costs and Fixed Costs of Purchase			
Stockout cost: Constant	s_0	11.528	1.446
Stockout Cost: Linear Term	s_1	0.001728	0.370
Inventory Carrying Cost: Square Term	c_2	0.006236	5.16E-5
Cost of Purchase: Constant	τ_1	228.46	3.527
Cost of Purchase: Size	τ_2	-4.7263	0.271
Cost of Purchase: Size^2	τ_3	0.06119	0.0016
Usage Rate Process: Type 1			
Mean	μ_1	3.0186	0.063
Standard Deviation	σ_1	0.5111	0.011
Usage Rate Process: Type 2			
Mean	μ_2	1.5222	0.0033
Standard Deviation	σ_2	1.1224	0.0054
Usage Rate Process: Type 3			
Mean	μ_3	0.5267	0.034
Standard Deviation	σ_3	0.5253	0.042
Usage Rate Type Persistence			
Type 1	Π_{11}	0.3511	0.0060
Type 2	Π_{22}	0.9958	9.53E-4
Type 3	Π_{33}	0.9101	3.95E-4
Type 4	Π_{44}	0.9049	3.41E-4
Usage Rate Types, Initial Probability			
Type 1	Π_{01}	0.1245	0.031
Type 2	Π_{02}	0.2770	0.101
Type 3	Π_{03}	0.2565	0.045
Family Size Effect on Usage Rate	f_z	0.03484	0.097

Table 7: Choice Frequencies in Data vs. Model Predictions

Sample Choice Frequencies

Size (oz)	Brand				Size Total
	Store	Del Monte	Heinz	Hunts	
14	0.0159	0.0049	0.0489		0.0698
28	0.0050	0.0156	0.1498		0.1752
32	0.0326	0.0904	0.3643	0.1540	0.6413
40		0.0032	0.0444	0.0060	0.0535
64		0.0029	0.0571	0.0049	0.0649
Brand Total	0.0535	0.1170	0.6646	0.1649	1.0000

Purchase probability: 0.06768

Simulated Choice Frequencies

Size (oz)	Brand				Size Total
	Store	Del Monte	Heinz	Hunts	
14	0.0183	0.0079	0.0426		0.0688
28	0.0158	0.0145	0.1064		0.1367
32	0.0310	0.0723	0.3636	0.1431	0.6100
40		0.0123	0.0848	0.0167	0.1139
64		0.0050	0.0486	0.0169	0.0705
Brand Total	0.0651	0.1155	0.6461	0.1768	1.0000

Purchase probability: 0.06775
Stockout probability: 0.6665
Average Inventory Level: 7.5226

Table 8: Average Offer and Accepted Prices in Data vs. Model Predictions

Mean Offer Prices - Data

Oz. Size	Store Brand	Del Monte	Heinz	Hunts
14	3.752	5.154	5.492	
28	4.024	4.830	4.901	
32	2.836	3.288	3.596	3.280
40		4.007	4.742	4.502
64		2.845	4.137	3.024

Mean Offer Prices – Simulation of the Model

Oz. Size	Store Brand	Del Monte	Heinz	Hunts
14	3.752	5.154	5.491	
28	4.022	4.827	4.897	
32	2.833	3.284	3.594	3.273
40		4.013	4.743	4.500
64		2.835	4.133	3.023

Mean Accepted Prices - Data

Oz. Size	Store Brand	Del Monte	Heinz	Hunts
14	3.747	5.078	5.535	
28	4.045	4.706	4.749	
32	2.760	2.996	3.509	3.114
40		4.145	4.619	4.470
64		2.580	3.909	2.993

Mean Accepted Prices – Simulation of the Model

Oz. Size	Store Brand	Del Monte	Heinz	Hunts
14	3.737	5.136	5.464	
28	3.666	4.649	4.674	
32	2.785	3.006	3.463	3.099
40		3.657	4.663	4.392
64		2.638	3.302	2.789

Note: The figures in the table are cents per ounce. For accepted prices, brand totals are obtained by dividing aggregate brand sales revenue by the aggregate quantity sold of the brand (i.e., purchases of larger sizes receive more weight).

Table 9: Brand Switching Matrix in Data vs. Model Predictions

Data

	Store Brand	Del Monte	Heinz	Hunts
Store Brand	0.2719	0.1338	0.4233	0.1711
DelMonte	0.0583	0.3407	0.4111	0.1898
Heinz	0.0340	0.0698	0.7895	0.1067
Hunts	0.0678	0.1576	0.4516	0.3230

Simulation of the Model

	Store Brand	Del Monte	Heinz	Hunts
Store Brand	0.2363	0.0930	0.4661	0.2047
DelMonte	0.0520	0.2270	0.4850	0.2360
Heinz	0.0468	0.0834	0.7422	0.1276
Hunts	0.0780	0.1474	0.4643	0.3103

Note: The left column reports the brand bought on the previous purchase occasion. The top row indicates the brand bought on the current purchase occasion.

Table 10: Effects of Temporary 10% Heinz Price Decrease On Purchase Frequencies for Heinz and Other Brands When Expectations are Adjustable ("Full") or Fixed

Week	Heinz		Hunts		Del Monte		Store Brand		Total	
	Full	Fixed	Full	Fixed	Full	Fixed	Full	Fixed	Full	Fixed
1	41.30	45.28	-3.99	-1.93	-3.58	-1.81	-3.11	-1.88	25.32	28.53
2	-2.07	-2.31	-1.23	-1.38	-1.23	-1.29	-1.07	-1.21	-1.75	-1.96
3	-1.56	-1.72	-0.94	-1.02	-0.80	-0.89	-0.79	-0.88	-1.31	-1.45
4	-1.40	-1.54	-0.68	-0.72	-0.57	-0.63	-0.58	-0.65	-1.13	-1.24
5	-0.80	-0.92	-0.53	-0.55	-0.37	-0.41	-0.39	-0.43	-0.68	-0.77
6	-0.55	-0.58	-0.26	-0.27	-0.21	-0.21	-0.18	-0.19	-0.44	-0.46
7	-0.42	-0.42	-0.15	-0.19	-0.12	-0.14	-0.24	-0.26	-0.33	-0.34
8	-0.25	-0.24	-0.10	-0.12	-0.06	-0.07	-0.09	-0.10	-0.19	-0.19
9	-0.17	-0.19	-0.02	-0.04	-0.05	-0.06	-0.05	0.03	-0.12	-0.14
10	-0.30	-0.15	-0.04	-0.04	-0.08	-0.06	0.04	-0.05	-0.21	-0.11
11	-0.13	-0.14	-0.04	-0.04	-0.08	-0.07	-0.03	0.04	-0.09	-0.10
12	-0.21	-0.21	-0.03	-0.04	-0.02	-0.03	-0.12	-0.04	-0.15	-0.15
13	-0.05	-0.04	0.01	-0.01	-0.01	-0.01	-0.01	-0.11	-0.04	-0.03
14	-0.15	-0.01	0.02	0.01	0.00	0.00	0.01	-0.00	-0.10	-0.01
15	0.02	0.04	0.03	0.04	0.01	0.03	0.01	0.02	0.02	0.04

Note: The table reports the effects of a temporary 10% price cut for Heinz on simulated weekly sales frequencies for Heinz and the other brands over a 15 week period. Changes are reported in percent terms. The first column (for each brand) shows the effect when expectations of future prices are allowed to adjust (which we denote here as "full" expectations). The second column (for each brand) shows the effect holding expectations of future prices fixed.

Table 11: Effects of Temporary 10% Heinz Price Decrease On Purchase Quantities for Heinz and Other Brands When Expectations are Adjustable ("Full") or Fixed

Week	Heinz		Hunts		Del Monte		Store Brand		Total	
	Full	Fixed	Full	Fixed	Full	Fixed	Full	Fixed	Full	Fixed
1	45.07	49.19	-4.02	-1.90	-3.64	-1.78	-3.15	-1.82	27.93	31.28
2	-2.14	-2.38	-1.22	-1.38	-1.13	-1.30	-1.08	-1.23	-1.81	-2.01
3	-1.51	-1.67	-0.91	-0.99	-0.80	-0.89	-0.81	-0.90	-1.29	-1.42
4	-1.61	-1.75	-0.65	-0.69	-0.56	-0.62	-0.60	-0.67	-1.26	-1.37
5	-0.87	-0.97	-0.54	-0.55	-0.36	-0.41	-0.41	-0.45	-0.73	-0.80
6	-0.57	-0.60	-0.26	-0.27	-0.20	-0.20	-0.19	-0.21	-0.46	-0.48
7	-0.43	-0.44	-0.14	-0.18	-0.11	-0.13	-0.24	-0.25	-0.33	-0.35
8	-0.26	-0.24	-0.13	-0.15	-0.06	-0.06	-0.09	-0.10	-0.20	-0.20
9	-0.18	-0.20	-0.02	-0.04	-0.04	-0.05	0.03	0.01	-0.13	-0.14
10	-0.44	-0.15	-0.04	-0.03	-0.08	-0.07	-0.05	-0.05	-0.31	-0.11
11	-0.12	-0.13	-0.04	-0.04	-0.08	-0.03	0.03	-0.03	-0.09	-0.10
12	-0.33	-0.32	-0.03	-0.04	-0.02	-0.01	-0.03	-0.04	-0.22	-0.22
13	-0.06	-0.04	-0.01	-0.01	-0.01	0.00	-0.10	-0.10	-0.04	-0.04
14	-0.19	-0.02	0.02	0.01	0.01	0.03	-0.01	-0.00	-0.12	-0.01
15	0.02	0.05	0.03	0.04	0.02	0.04	0.01	0.02	0.02	0.04

Note: The table reports the effects of a temporary 10% price cut for Heinz on simulated weekly sales quantities for Heinz and the other brands over a 15 week period. Changes are reported in percent terms. The first column (for each brand) shows the effect when expectations of future prices are allowed to adjust (which we denote here as "full" expectations). The second column (for each brand) shows the effect holding expectations of future prices fixed.

Table 12: Predicted Effects of Permanent Changes in Heinz Pricing Policy

Permanent 10 % drop in Mean Offer Price of Heinz

	Store Brand	Del Monte	Heinz	Hunts	Total
Purchase Probability	-8.078	-9.071	33.071	-10.100	18.038
Purchase Quantity	-8.075	-8.821	35.581	-10.186	19.844
Revenue	-8.052	-9.008	23.363	-10.242	13.151
Accepted Price (Mean)	-0.048	-0.205	-9.012	-0.063	-5.585

Permanent 50 % drop in the Standard Deviation of Heinz Offer Prices

	Store Brand	Del Monte	Heinz	Hunts	Total
Purchase Probability	-6.351	-10.172	-5.278	-6.994	-6.200
Purchase Quantity	-7.522	-11.145	-6.642	-7.256	-7.295
Revenue	-7.385	-10.683	-4.221	-7.110	-5.485
Accepted Price (Mean)	0.141	0.520	2.594	0.157	1.952

Combined 2.2% Permanent drop in Mean and 50 % drop in Standard Deviation for Heinz Price

	Store Brand	Del Monte	Heinz	Hunts	Total
Purchase Probability	-6.746	-12.157	1.052	-9.347	-2.774
Purchase Quantity	-7.340	-13.021	≈0	-9.660	-3.622
Revenue	-7.238	-12.669	0.537	-9.545	-2.802
Accepted Price (Mean)	0.110	0.404	0.530	0.127	0.851

Note: The table reports the percentage changes in each of the indicated quantities for the period after the policy change, compared to a baseline simulation under the present pricing policy. The mean accepted prices are obtained by dividing aggregate sales (over all sizes) by aggregate quantity.

Table A1: Effect of Increasing Fundamental Parameters on Key Features of the Data

Parameter	Purchase Hazard: **General Level**	Purchase Hazard: **Duration Dependence**	Purchase Frequency: **Overall**	Purchase Frequency: Large Size vs. Small Size	**Mean Accepted Price**
Carrying Cost					
Quadratic - c_2	-	+	-	-	≈ 0
Fixed Cost of Purchase					
Constant - τ_0	-	0	-	+	≈ 0
Linear - τ_1	-	0	-	-	≈ 0
Quadratic - τ_2	-	-	-	--	≈ 0
Stock Out Cost					
Constant – s_0	+	+	+	+	≈ 0
Linear – s_1	Small +	0	Small +	Small +	≈ 0
Utility Weights - ψ	+	0	+	+	Ambiguous (small + own effect for preferred brand)
Usage Rate					
High Type	+ at early weeks only	--	+	+	+ for 64 oz. Only
Medium Type	- at early weeks (≤ 5) + at weeks 6+	--	≈ 0	+	≈ 0
Low Type	+ at intermediate weeks only	+ at early weeks - at later weeks	+	+	≈ 0

Note: "+" denotes increase and "-" denotes decrease. A "- -" denotes that increasing a parameter results in an exceptionally sharp decrease for some data feature, relative to the magnitude of the parameter's effect on other data features

Figure 1: Observed Frequency of Total Purchases

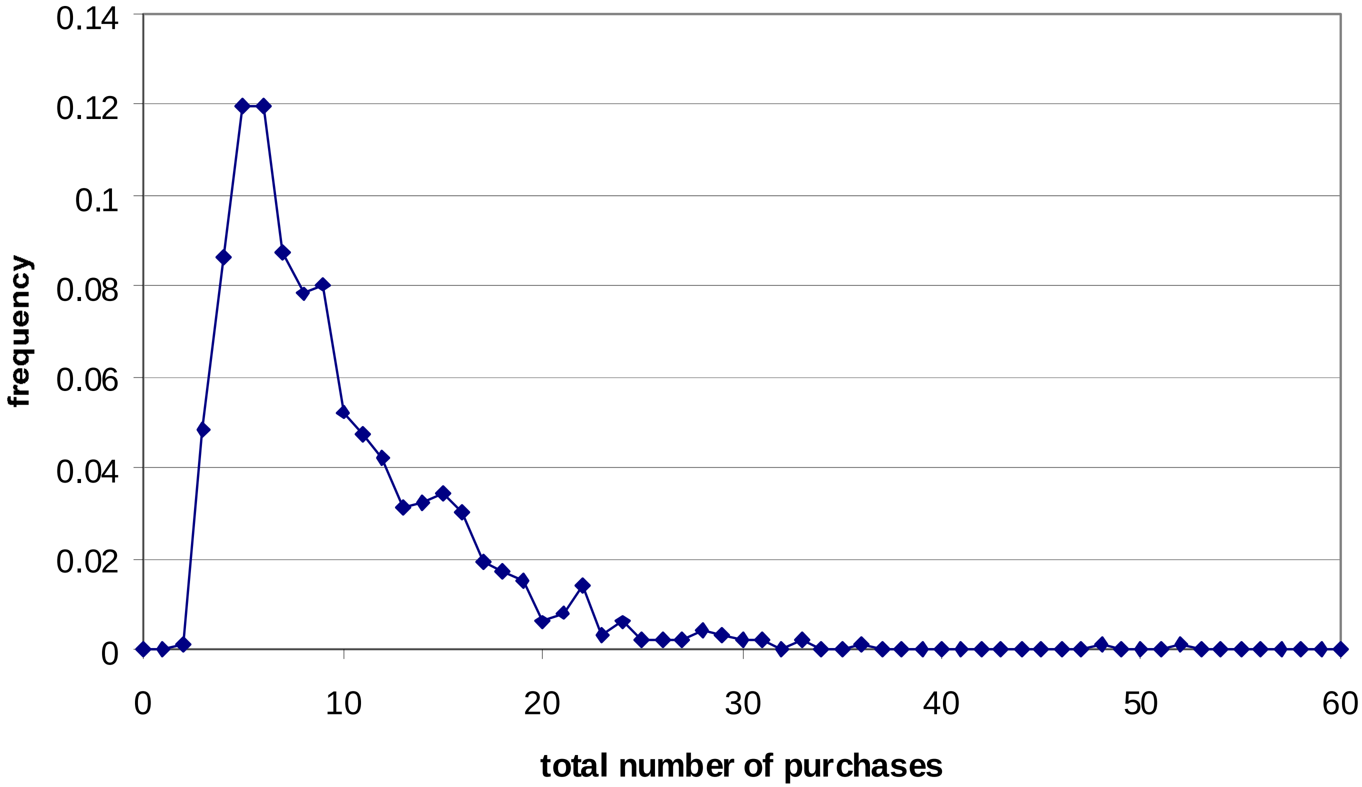

Figure 2: Interpurchase Time Distribution

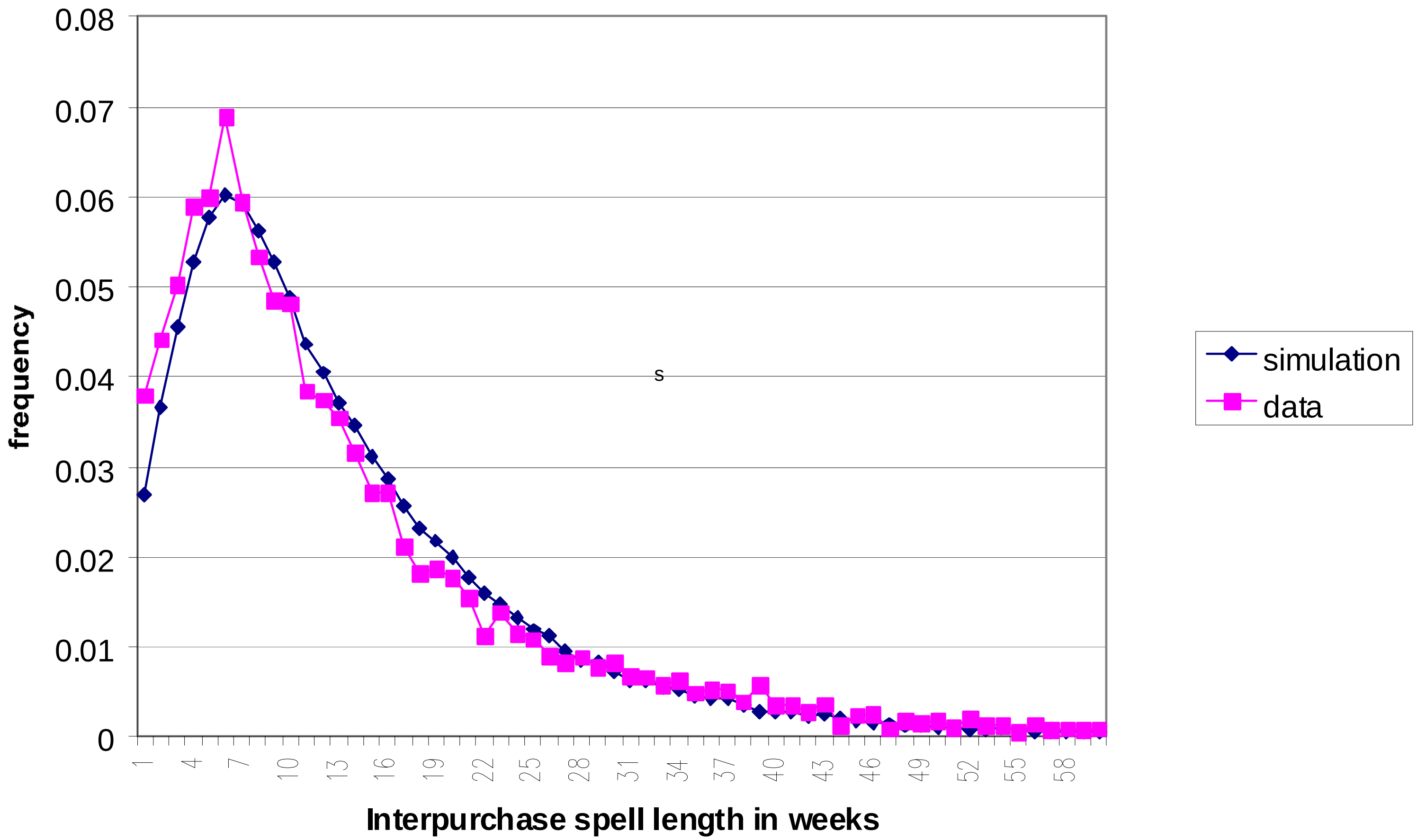

Figure 3: Survivor Function

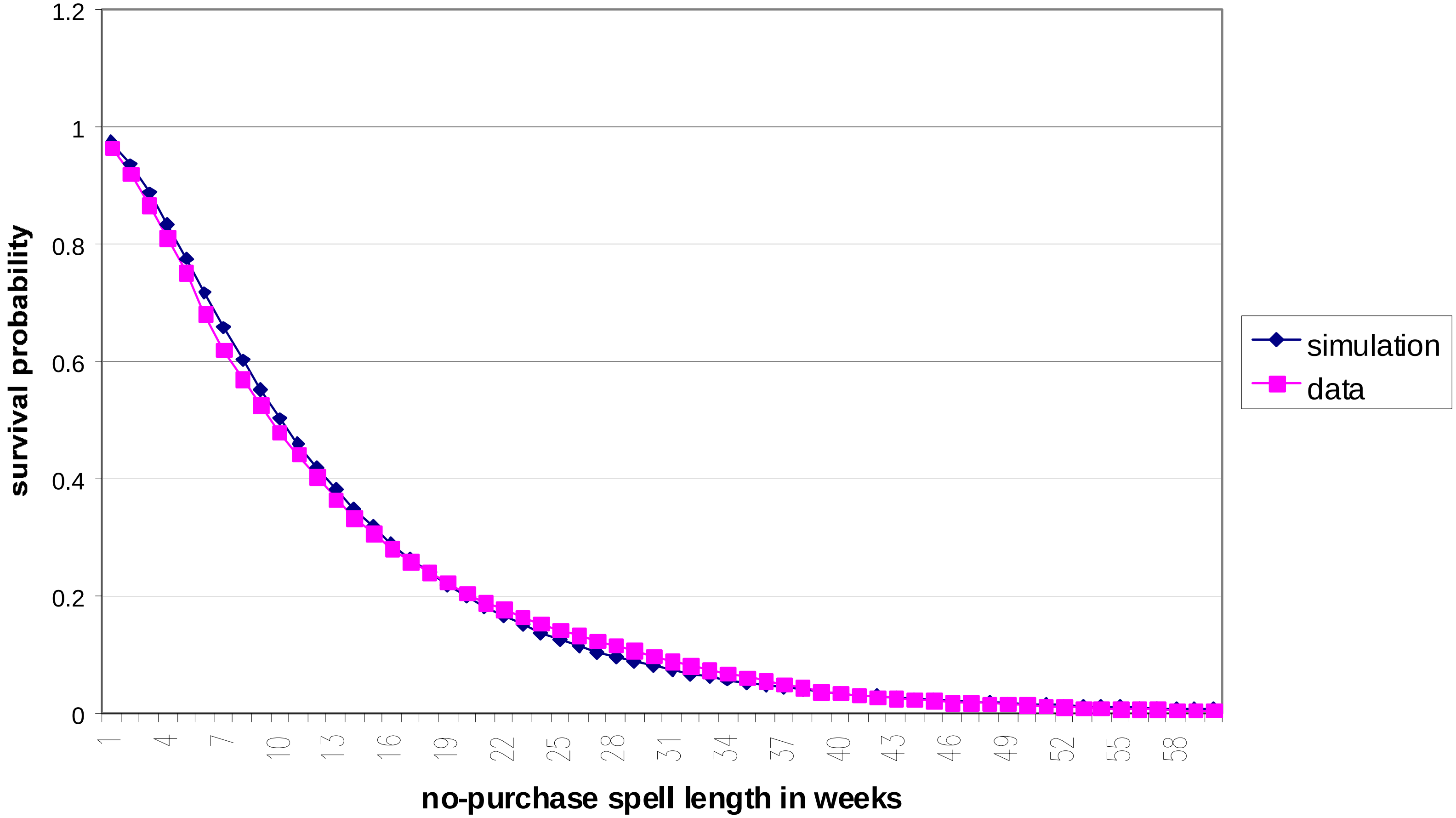

Figure 4: Purchase Hazard