Stockpiling, Learning, and the Success of New Products

Christoph Bauner*

March 3, 2013

Abstract

In a dynamic and competitive environment, factors such as the market maturity and market concentration may have significant impacts on the success of a new entrant. This paper sheds some light on how market structure influences that success in the presence of both learning and stockpiling. Using the market for sugar substitutes from 2001 to 2005, I develop and structurally estimate a novel dynamic demand model that incorporates both, aspects of learning and stockpiling. Households in the model learn their reactions to brand characteristics over time; they also engage in stockpiling in order to benefit from non-linear pricing as well as intertemporal price variations. I overcome the complexity of combining learning and stockpiling identified by previous authors by employing a new estimation technique. Parameters of the consumer utility function are then used to quantify the differential in market shares for a new entrant, comparing several counterfactual scenarios. I find that the timing of entry has a large effect on the entrant's market share in its first year in the market. Entering in a less mature market can increase an entrant's sales volume by as much as a factor of three when compared to entering the same market with more established firms four years later. Most of this effect comes from competition with the youngest incumbent brand in the market.

1 Introduction

The success of a new product is vitally important not only to the entering firm but also to consumers who may relish the added variety or benefit from prices lowered through competition. One of the most important factors influencing this success is the timing of entry. Timing impacts product success in numerous ways. A premature entry might mean reduced R&D times and thus a lower quality product; consumer preferences might change over time[1] so that market entry might be too early or too late; or a delayed entry might give competitors the opportunity to establish themselves in the market, making it hard to enter successfully. In this paper I study how this last factor, the timing of entry in a maturing market, influences the success of a new product.

This paper builds a dynamic demand model of learning and inventory. I structurally estimate this model using store and household data for sugar substitutes from 2001 and 2005. This is a particularly interesting market which during that period was going through major changes. Splenda, a brand which then was relatively new, was in the process of growing from a minor producer to the major national brand in the market. This variation in market share observed in the data allows me to study market entry in different scenarios, ranging from a market with a relatively weak leader whose market share is ever-declining to a market with Splenda as a dominant new frontrunner accounting for more than half of all sugar substitute sales.

I use pricing information on an actual new brand, Dixie Crystals Sweet Thing (henceforth DCST), that entered the market in 2004 to conduct counterfactual simulations assessing the market share DCST could have achieved at different entry dates. I find that timing of entry plays an important role in the success of the new brand. An early entry is preferable and leads to market shares which are up to three times as high as what would have been achieved with late entry. I also find that almost all of the variation in DCST's market share is mirrored by Splenda, so that the combined market share of the two is nearly constant over most of the simulated scenarios. These findings can be explained by learning about brand characteristics; while households already have enough knowledge of the old products available they might still not be well informed about the attributes of the two newer brands. Thus, the new entrants end up competing over which households they can get to sample their products and possibly become committed customers.

To realistically capture the market shares of a new entrant in counterfactual simulations,

---

[1]This is the case for instance in the fashion industry.

incorporating both learning and inventory in the demand model is important. First, learning is crucial in this context because consumers typically are not well-informed about the characteristics of newly introduced grocery products. This knowledge only grows over time with information added every time the product is sampled. The drastic changes in market shares observed in the specific market studied in this paper indicate that learning plays an important role in that market. Stockpiling is also an important feature whenever one studies storable goods. It allows households to buy large quantities at temporarily low prices and thus benefit from intertemporal substitution. Previous literature has demonstrated that accounting for inventories is important when estimating the demand for storable goods; employing a static model instead could lead to significantly misestimated elasticities.[2]

This paper extends and combines the literatures on learning and on stockpiling. While both literatures have made great strides in the past decade there is still a substantial challenge in structurally estimating a model that combines both components. As pointed out by Ching, Erdem and Keane (2011) it is computationally challenging to implement and hence estimate such a model. The difficulty lies in computing the dynamic programming problem necessary for estimating the model and arises due to the substantial size of the state space associated with both stockpiling and learning. In order to accommodate stockpiling the state space incorporates not only the inventory but also a consumption shock. For the learning component, beliefs for every good in the market are also included. In addition to these, the prices for all brands and sizes have to be tracked in the state space. Overall, with traditional methods the dynamic programming problem would have to be solved on a grid of more than 10^{15} points. This is simply infeasible even with modern day computers. The high computational demand of this estimation is exacerbated by the inclusion of random coefficients in the model, which are necessary for retaining a good description of the market as they allow households to have unobserved persistent taste differences.

I overcome this computational burden by employing two recently developed methods: First, I use the Sieve Value Function Iteration proposed by Arcidiacono et al. (2012) to solve the dynamic programming problem. Second, I estimate the parameters of the model with Ackerberg (2009)'s Importance Sampling method. Together these techniques allow me to significantly reduce the computational difficulty of the estimation.

My model and the estimation routine are flexible and can therefore be easily applied to other questions in marketing and economics that require the estimation of demand in the presence of storability and learning. Here, I use it to study the success of a market entrant.

---

[2]See Erdem, Imai and Keane (2003) or Hendel and Nevo (2006) for examples of this literature.

The results are relevant for understanding why certain products succeed and how new brands choose the launch time of their products. They also relate to studies on market structure which is intimately related to market entry.

The rest of the paper is organized in the following way: In the next section I place my work in the context of the relevant literature. After that I introduce the data and present evidence for stockpiling and learning. In section 4 I discuss my model. Section 5 explains the estimation procedure and presents the results of the demand estimation. In section 6 I describe the counterfactual simulations and discuss my results concerning the effect of timing on market entry success. Section 7 concludes.

2   Related Literature

There is an extensive literature on the optimal timing of market entry. Urban et al. (1986) and Bronnenberg, Dhar and Dubé (2009) analyze sales data in several sectors of grocery products and show that market shares depend on the order of market entry. These papers differ from the present one in that I am not concerned about the order of entry and instead concentrate on entry in markets of different maturity.

Similar research has looked at innovative industries. Bayus, Jain and Rao (1997) study a model in which an innovator planning to introduce a new product faces a trade-off between early market entry and high product performance. If the innovator expects copy-cat products to enter the market it may delay its own product launch in order to have more development time and improve product performance. This allows them to develop a stronger product and lose a smaller share of the market to their competitors. However, if the products feature network externalities and are incompatible then there is yet another outcome. Katz and Shapiro (1992) and Regibeau and Rockett (1996) show this does not hold if the products feature network externalities and are incompatible. In that case firms rush to enter their product at a time that is inefficient from a welfare perspective in order to prevent their rivals from capturing too large a market share. These papers feature direct advantages of entering late or early. Here, the incentives are less direct and instead come from households' learning over time.

Krider and Weinberg (1998) investigate the timing of movie releases. There are times of higher and lower demand over the course of a year, so production companies face a trade-off between trying to release their movie at high-demand times and trying to avoid competition. They show that stronger movies tend to be released when demand is high.

Producers of weaker films favor a time with lower demand and less competition. Thus, they may strategically delay the release. However, Krider and Weinberg (1998) do not need to feature any longer-term effects of timing that come with households' learning about brands.

On the methodological side, this paper relates to and combines two strands of literature: structural models of learning and experience goods and models of stockpiling.

A vast literature on learning has developed during the past two decades and it's beyond the scope of this paper to review it completely.[3] Here, I can only highlight a few of the more relevant examples. Erdem and Keane (1996) estimate a structural learning model in which consumers learn about attributes of several competing brands. Households are forward-looking because they may strategically sample products; to my knowledge it is the first model to incorporate the choice between a large set of brands in this kind of model. They show that increased advertising frequency has a major effect on market share in the long run. Ackerberg (2003) uses a similar model to study the differences between learning through advertising and learning through experimentation. He finds that advertising serves as an important source of information in his sample. Crawford and Shum (2005) are the first researchers to incorporate multi-attribute learning in a structural model. They study the market for pharmaceuticals and allow for learning of symptomatic and curative effects. According to their work products do well if they are strong among either one of those dimensions. Ching (2010) also studies pharmaceuticals in a model with learning. He finds evidence that prices for brand-name products are kept artificially low in order to slow down learning on generics. Osborne (2011) builds on pure learning models but adds switching costs to the model. He shows that ignoring either component might lead to misestimated elasticities. My paper differs from the ones mentioned above by incorporating stockpiling into the model. Ching, Erdem and Keane (2011) try to achieve this by proxying for a consumer's inventory using the time since his last purchase. Contrary to that, I specify and estimate a full model of inventory in conjunction with learning.

Dynamic demand models incorporating inventory have become increasingly popular since Erdem, Imai and Keane (2003) showed that the failure to account for intertemporal substitution might lead to significantly misestimated elasticities. Hendel and Nevo (2006) have a similar findings and introduce ways to significantly reduce the computational burden of estimating these models. Wang (2012) incorporates unobservable persistent household heterogeneity into an inventory model of soda sales. She finds that a model without unobservable heterogeneity would have significantly misestimated the effects of a tax on soft drinks.

---

[3]Ching, Erdem and Keane (2011) provide an extensive discussion of the relevant articles.

Hartmann and Nair (2010) study a model of tied goods. In the market for safety razor and razor blades rational consumers act forward-looking in two ways: first, they take blade prices into account when purchasing razors; second, they form expectations over prices and buy when products are cheap. The present paper differs from this strand of literature in that the dynamics in my model come from learning as well as from stockpiling.

3 Data

3.1 Overview

In this section I will introduce the data on which the estimation is based. I use weekly scanner data from 2001 to 2003 provided by IRI. The data consists of four components: A household panel, which shows weekly purchases for given consumers; a store panel, which shows prices and advertising measures for stores over time; demographic information on households in the sample; and product characteristics like brand name and package size for all products in the sample.

The household panel shows which products a household purchased in every week, how many units were purchased and at what price, and in which store the purchase was made. The store panel shows for each product the number of units bought in the week and the total revenue from those sales. In addition there is information on whether any promotional activities such as price reduction, feature or display were used and if so which ones. I also observe in which market a store is located and what type of shop (grocery store, drug store, or mass merchandise) it is.

Since for each household the stores where purchases are made are observed the entire choice set faced by the household can be uncovered by linking to the store panel. Thus, I know the prices of not only the products purchased but also the products rejected in each week.

There are 71 brands of sugar substitutes in my data and most of them are available in different package sizes. It's computationally infeasible to take prices for this number of products. Additionally, most products are not consistently available in most stores. I restrict the analysis to the four largest brands in the data set. These are Equal, Natrataste, Splenda, and Sweet'N Low and account for more than 80% of the market.

In the following I concentrate only on packages with one-serving packets. These are boxes typically filled with 50 to 200 pouches. Each pouch contains one gram of sweetener which is the equivalent of two table spoons of sugar. These pouches are typically used for

Table 1: Average Prices for Medium Sized Packs of Sugar Substitutes over Time

	2001	2002	2003	2004	2005
Equal	4.32	4.21	4.28	4.16	4.17
Natrataste	2.48	2.40	2.42	2.39	2.48
Splenda	4.33	4.26	4.21	4.11	4.30
Sweet'N Low	1.86	1.86	1.93	1.89	1.97

Note: All prices in 2001 dollars.

sweetening beverages and are by far the best-selling form of sugar substitute in my data (about 80% of purchases). I believe this is a reasonable separation of markets. Other packages, like bulk bags, are presumably mainly used for baking or similar activities. These uses require other qualities than sweetening coffee. For instance, many sugar substitutes lose their sweetness when heated too much. This would make those sweeteners not very desirable for cooking but would in no way undermine their usefulness when stirring it into hot beverages. Thus, including multiple forms of packaging into the estimation might lead to biased results. Importantly, this does not restrict households' possibility to stockpile. For most brands there are three different package sizes with one-serving pouches. Additionally, households can purchase multiple packs. By purchasing several packs of 200 servings each in a short time-frame, consumers can quickly build up a significant inventory.

One of the most salient features of the data is how the market shares of these brands change over time (see figure 1). Splenda is rising fast from a market share of about 10% to close to 50% while other brands are on a decreasing trend. This is particularly remarkable as it is apparently not caused by a major change in prices. Table 1 shows that mean prices in real terms did not fluctuate very much.[4]

Some households are not observed very often in the data, for instance because they did not report regularly. Since I'm interested in the dynamic effects of learning and inventory it is important to follow households' behavior over a relatively long time frame. Therefore households observed for less than three years continuously are dropped. In the reduced data set I observe 439 households. Table 2 shows some summary statistics on their demographics and their shopping behavior.

Most households in the sample are middle income households; the median household earns between $35,000 and $44,999, which is close to the national average. The average

---

[4]Table 1 shows average prices across stores for medium sized (3.5 oz, 4.5 oz for Sweet'N Low) packs. Prices for other sizes show a similar lack of movement.
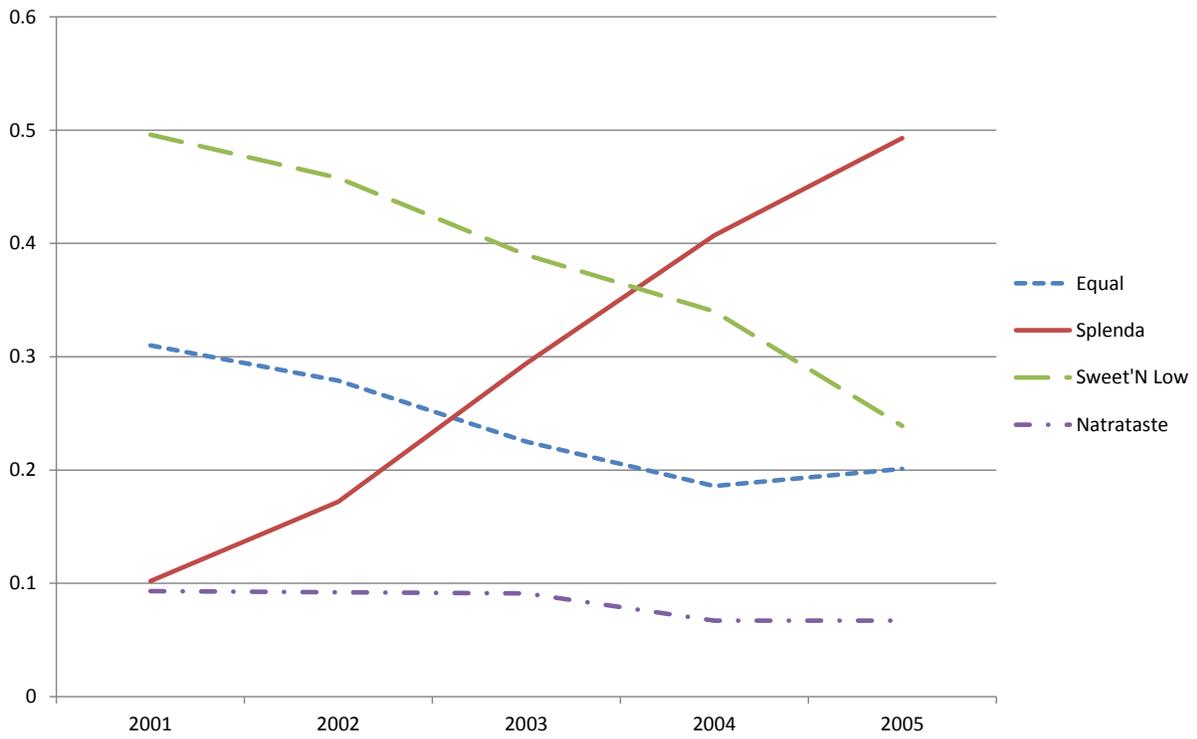
Figure 1: Market Shares of Four Leading Brands over Time.

family size is is relatively small at 2.33; one factor contribution to this is that IRI does not differentiate household sizes greater than six. Households buy sugar substitutes roughly three times per year; they purchase on average about 12 ounces of sugar substitute and pay about $10 for it. These numbers might seem relatively low; however, taking into account that sugar substitutes tend to be very light weight they are reasonable. In fact, 12 ounces correspond to roughly 350 servings for the brands considered for this study; thus, the average household purchases roughly one serving per day.

Table 2: Household Panel Summary Statistics

	Avg.	Std. Dev.
Demographics		
Family Size [a]	2.33	1.14
Income ($)	35,000-44,999 [b]	
Purchase Behavior		
Number of Purchases	2.95	2.86
Avg. Annual Volume Purchased (oz)	11.80	12.16
Avg. Annual Spending ($)	10.16	11.32

[a] Family size is capped at 6 in the IRI data.
[b] Displayed is the income bracket of the median household.

Some variables from the store panel are summarized in table 3. The first six rows of the table show average prices and standard deviations of prices across brands and package sizes. Typically smaller sizes are more expensive on a per ounce basis. A notable exception is Sweet'N Low for which the large package comes at a price of $0.47 per ounce compared to $0.42 for the medium sized package. It's not immediately clear what drives consumers to buy the larger package under these conditions. However, it is important to note that there is considerable variation in prices. Particularly, temporary price reductions do not necessarily apply to all package sizes of a brand at the same time so that price orderings are sometimes reversed. This means that the incentives for stockpiling fluctuate over time, a feature that is important since it helps with identification of the model. Table 3 also shows the advertising frequency for each brand. As can be seen, advertising is quite rare, particularly for Splenda which nonetheless increases its market share dramatically over time. In addition, most of the advertising measures in the data are relatively small in scale (e.g. small store displays). Thus, it is unlikely that these measures have a major effect on market shares. The last three lines of the table show the relative purchase frequencies for each size and brand. Natrataste

offers only medium sized packages in my sample, so these naturally account for 100% of sales. For the other brands, medium sized packages also seem to be most popular. However, since large packs are roughly twice the size a higher volume is sold through them.

Table 3: Store Panel Summary Statistics

	Equal	Natrataste	Splenda	Sweet'N Low
avg. price small[a]	2.60	-	2.73	1.26
std. dev. price small[a]	0.29	-	0.16	0.16
avg. price medium[b]	4.24	2.43	4.24	1.89
std. dev. price medium[b]	0.32	0.32	0.36	0.17
avg. price large[c]	7.35	-	7.32	3.91
std. dev. price large[c]	0.39	-	0.28	0.16
advertising frequency (%)	2.0	4.5	0.8	3.4
% of purchases small[a]	34.7	-	33.5	34.6
% of purchases medium[b]	37.4	100	45.1	38.7
% of purchases large[c]	27.9	-	21.4	26.7

[a] 1.75 oz
[b] 3.5 oz for Equal, Natrataste and Splenda; 4.5 oz for Sweet'N Low
[c] 7 oz for Equal, Natrataste and Splenda; 8.75 oz for Sweet'N Low

3.2 Evidence of Stockpiling and Learning

I now show some evidence that stockpiling and learning are common characteristics in the market and that therefore they need to be included in an accurate estimation.

In table 4 I show the number of brands sampled by households. There are two contrary effects going on when looking at this: On the one hand, households have preferences over brands and to some degree are loyal to brands they like. On the other hand, households experiment in order to find the best available product. Of course households with weak preferences might also be swayed if prices for their less preferred brands are reduced. The table is restricted to households buying sugar substitutes at least twice over a three-year span since a single purchase does not leave room for a meaningful statement about switching behavior. Roughly one third of the remaining households buy only a single brand, while two thirds purchase two or more different brands. 10% of households actually sample all four brands; this provides some initial evidence that learning might be prevalent in this market. The alternative explanation would be that these households do not have strong preferences

10

over any of the four brands. This, however, seems unlikely since the brands in the market differ significantly among several dimensions including price, taste and sweetening agent.

Table 4: Brand Switching Behavior

Number of Brands Tried	Frequency (in%)
1	36.58
2	37.84
3	15.35
4	10.23

Figure 2 provides some evidence for stockpiling in this market. It shows the share of a households' annual sweetener purchases that is purchased in a single trip. For the purpose of this figure, I exclude households who buy only one unit of the smallest package size since these consumers naturally purchase their whole supply in one trip. I also exclude households whose total annual purchase is greater than largest package available. These consumers obviously have to buy several packages which makes the results harder to interpret. The figure shows that a majority of these households purchase their whole demand in a single shopping trip. This is clear evidence that consumers either purchase large packages in order to benefit from non-linear pricing or buy several units at once, presumably when the items are on sale.

An example for the purchase behavior of a household from the data is provided in table 5. The consumer, identified in my data as household number 1816157, starts out purchasing Sweet'N Low. However, at some point he seems to have decided to strategically sample other brands, buying small or medium sized packages of three different products in three consecutive purchases. Apparently these tries were not very satisfactory; at the end of the observation period the household goes back to purchasing Sweet'N Low and eventually even switches to the large package size.

Table 5: Purchase Behavior over Time

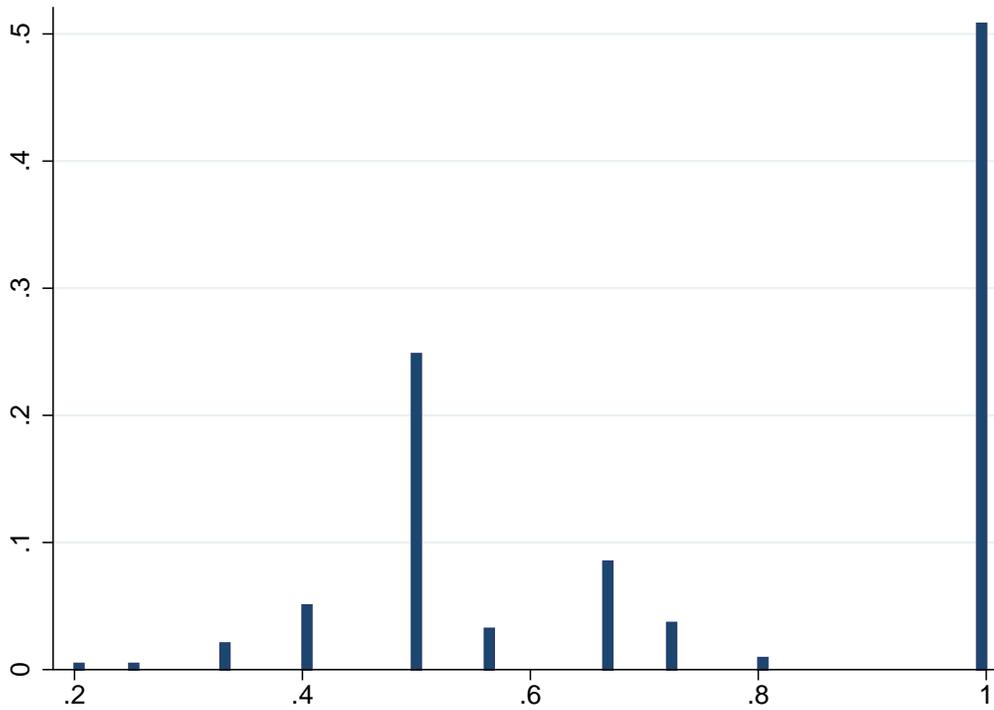| Week	Brand	Volume
1/14/2001	Sweet'N Low	4.5
8/26/2001	Sweet'N Low	4.5
5/5/2002	Sweet'N Low	4.5
6/23/2002	Nutrasweet	3.5
11/17/2002	Equal	1.75
1/19/2003	Private Label	3.5
2/23/2003	Sweet'N Low	4.5
3/30/2003	Sweet'N Low	8.75

Figure 2: Largest Purchase as Fraction of Total Annual Purchase by Household

4 Model

4.1 Overview

In this model of learning and inventory there are H risk-neutral households, indexed by $h = 1, \ldots, H$, who might purchase sugar substitutes. They have the choice between K products and an outside option denoted as product 0. Product k ($k = 1, \ldots, K$) is available in different package sizes, given by the set J_k. A package of product k and size j contains a quantity of q_{kj} servings.

In each week households observe the available products and prices and then decide how much to consume, whether to purchase anything, and if so which brand and package size to purchase.

Households differ by their preferences over goods; for instance a household might be indifferent between all sweeteners or it might strongly prefer Splenda to all other brands. However, households don't know perfectly the utility associated with a good. Instead, they have a belief over this utility which is unbiased but not necessarily accurate. Every time a household buys a product it receives another signal about the quality of this product and over time its belief becomes more accurate. Households are aware of this process and may strategically sample products for the purpose of learning about their preferences.

Households do not have to consume their purchase immediately. Instead, they can choose to store some or all of the sugar substitutes bought today for future use. This allows them to make use of temporary price reductions as well as nonlinear pricing in order to reduce their purchase costs.

4.2 Utilities

Household h chooses current consumption (c_{ht}) as well as a brand (k_{ht}) and a package size (j_{ht}) to purchase in order to maximize the sum of discounted future utilities. I denote purchase choices by $\wp_{ht} = \{k_{ht}, j_{ht}\}$ where it is convenient and define $\mathfrak{P}_t$ as the set of all purchases $\wp_{ht}$ possible at time t. Formally, household h's objective at time t is to solve the

following infinite horizon problem:

$$V_h(\theta_{ht}) = \max_{\{c_{h\tau}, \wp_{h\tau}\}} \sum_{\tau=t}^{\infty} \delta^{\tau-t} E[u_{h\tau}(c_{h\tau}, k_{h\tau}, j_{h\tau})|\theta_{ht}] \tag{1}$$

$$\text{s. t. } 0 \leqslant c_{h\tau}, \quad 0 \leqslant i_{h\tau}, \quad j_{h\tau} \in J_k$$

$$i_{h\tau} = i_{h\tau-1} + j_{h\tau} - c_{h\tau}$$

$$\wp_{h\tau} \in \mathfrak{P}_\tau$$

where θ_{ht} represents the state in period t, δ is the discount factor and u_{ht} and i_{ht} are household h's flow utility and inventory, respectively, at time t. That is, the household finds his optimal consumption and purchase subject to feasibility constraints.

Current flow utility when buying size j of product k is given by the following functional form:

$$u_{ht}(c_{ht}, j_{ht}, k_{ht}) = \alpha^h \log(c_{ht} + \eta_{ht}) - \beta_1^h i_{ht} - \beta_2^h i_{ht}^2 - \zeta^h p_{htkj} + q_{kj} \Gamma_{tk}^i + \epsilon_{htkj} \tag{2}$$

This can be split into several parts: First, the utility of consumption is

$$\alpha^h \log(c_{ht} + \eta_{ht}).$$

Here, α^h measures the marginal utility of consumption and η_{ht} is a random shock, allowing households' preferred consumption levels to change over time. Second, there is a cost associated with keeping an inventory i_{ht}. This cost is given by

$$\beta_1^h i_{ht} + \beta_2^h i_{ht}^2$$

and accounts for the inconvenience of having to store large quantities of goods; in other words, it captures the opportunity cost of domestic space. The last part captures the utility of the shopping experience. This is:

$$-\zeta^h p_{htkj} + q_{kj} \Gamma_{tk}^h + \epsilon_{htkj}$$

ζ^h measures the marginal utility of income. It is lost since the household has to pay p_{htkj}, the price it observes at time t for its chosen product k and size j. Γ_{tk}^h is the taste household h has for good k. This allows different households to prefer different brands. Finally, ϵ_{htkj} is an idiosyncratic random shock following a type-I extreme value distribution.

14

Households' preference for the outside good is normalized to $\Gamma_0 = 0$. Also, in that case the household obviously does not have to make a payment, i.e. $p_{ht0} = 0$. Thus, when a household does not purchase anything the flow utility becomes:

$$u_{ht}(c_{ht}, 0) = \alpha^h \log(c_{ht} + \eta_{ht}) - \beta_1^h i_{ht} - \beta_2^h i_{ht}^2 + \epsilon_{ht0} \tag{3}$$

Note from equation (2) that households realize the utility associated with a specific brand immediately after purchasing the good. This is a simplification introduced by Hendel and Nevo (2006). It allows for keeping only a single inventory (as opposed to keeping track of inventories for all brands) and thus reduces the number of state variables. Additionally, this assumption removes the need of specifically modeling how households choose between consuming the various products in their storage.

4.3   Realized Utility

Note the household taste parameter Γ_{tk}^h has an index for time t. This is because households' tastes can change over time due to a shock influencing product experience according to the following assumption.

Assumption 1. *The realized utility derived from buying any good k is:*

$$\Gamma_{tk}^h = \gamma_k^h + \varepsilon_{kt}^h \tag{4}$$

where γ_k^h is household h's average utility for product k and ε_{kt}^h is an iid shock following a normal distribution with mean zero and standard deviation σ^h and is unknown to consumers at the time of purchase.

Assumption 1 implies that consumers' usage experience randomly changes from one unit to the next. This is a good description of reality due to two factors: First, consumers actual tastes might slightly vary. For instance, the reaction to certain foods might vary depending on which nutrients the body lacks. Second, there might be actual variation in product quality, like torn packages or spoiled ingredients.

Using (4) the expression for the *realized* flow utility can be characterized in the following way:

$$u_{ht}(c_{ht}, j_{ht}, k_{ht}) = \alpha^h \log(c_{ht} + \eta_{ht}) - \beta_1^h i_{ht} - \beta_2^h i_{ht}^2 - \zeta_1^h p_{htkj} + q_{kj}\left(\gamma_k^h + \varepsilon_{kt}^h\right) + \epsilon_{htkj} \tag{5}$$

4.4 Household Decisions

Households do not know ahead of time what the realized value of Γ_{tk}^h will be. They also don't know their true product preference γ_k^h. Instead, they have a belief over Γ_{tk}^h. Φ_{kt}^h and φ_{kt}^h specify the mean and the standard deviation, respectively, of the cdf associated with this belief. Before sampling a product for the first time a household has an initial belief $\left(\Phi_{k0}^h, \varphi_{k0}^h\right)$. After every purchase the household updates its belief according to the Bayesian updating procedure:

$$
\begin{aligned}
\Phi_{kN+1}^h &= \frac{\Phi_{kN}^h \sigma^{h2} + \Gamma_{tk}^h \varphi_{kN}^{h2}}{\sigma^{h2} + \varphi_{kN}^{h2}} \\
\varphi_{kN+1}^{h2} &= \frac{1}{\left(\frac{1}{\varphi_{kN}^h}\right)^2 + \left(\frac{1}{\sigma^h}\right)^2}
\end{aligned}
\tag{6}
$$

The following assumption describes the timing of learning:

Assumption 2. *Households' beliefs are updated immediately after the purchase.*

Assumption 2 states that households learn post purchase, not post each use. This is a simplification that makes learning tractable. Without it, learning would depend on consumption which is not be observed. It is valid since households rarely make several purchases of sugar substitutes in close succession. Thus, the households belief at the time of the next purchase does not differ significantly from an alternative model in which learning gets triggered by consumption.

Households make purchase decisions based on their beliefs over all products. That is, they behave as if flow utility were:

$$
u_{ht}^{\Phi}(c_{ht}, j_{ht}, k_{ht}) = \alpha^h \log(c_{ht} + \eta_{ht}) - \beta_1^h i_{ht} - \beta_2^h i_{ht}^2 - \zeta^h p_{htkj} + q_{kj}\Phi_{tk}^h + \epsilon_{htkj}
\tag{7}
$$

Substituting (7) into the infinite horizon problem, the purchase probabilities can be calculated by the standard multinomial logit formula.

4.5 States

So far, utility has been expressed as a function of the current state θ_{ht}. In order to solve the utility maximization problem we need to calculate the transition of states. State variables include prices for all goods and sizes, inventory, the consumption shock η_{ht} and households' beliefs over product quality. In the previous sections I have already discussed how inventory

and households' beliefs evolve. The following assumptions describe the dynamics of the remaining state variables:

Assumption 3. *The consumptions shock η_{ht} is iid over time and households.*

The consumption shock allows households' consumption to vary over time, even if states were to be identical otherwise. Thus, it accounts for changes in unobservable variables that influence households' consumption. Consumption shocks are assumed to be independent over time; thus, its distribution remains static.

Assumption 4. *Prices follow an exogenous first-order Markov process.*

Assumption 4 is a common way of reducing the state space in inventory models.[5] It implies that households' choices depend on prices for only one period. While it is not clear that there is an equilibrium in which prices follow a first-order Markov process, I follow Hendel and Nevo's (2006) argument that this reasonably approximates shoppers' memory.

5 Estimation

In the previous section I develop the model which incorporates unobserved heterogeneity on all structural parameters. The goal of the estimation is to uncover the distributions of those parameters. I call the set of structural parameters

$$\Psi^h = \{\alpha^h, \beta_1^h, \beta_2^h, \zeta^h, \sigma^h, \gamma^h\}$$

where $\gamma^h = \{\gamma_1^h, \gamma_2^h, \gamma_3^h, \gamma_4^h\}$ is the set of household preferences for all products. Ψ^h is random from the econometrician's perspective; it is distributed according to the joint cdf $G(\Psi|\Upsilon, \Sigma)$, where Υ and Σ are parameters governing the shape of G. In order to simplify the problem I assume that the off-diagonal elements of Σ are zero. In practice, I choose the marginal cdf $g(\upsilon, \sigma)$ to be log-normal for α^h, ζ^h, and σ^h since these parameters are naturally bounded at zero. The βs and γs follow normal distributions so that the results are restricted as little as possible.

For any given state θ_{ht} the probability of household h's choice at time t for a set of preferences is:

$$cp_{ht}(\theta_{ht}) = \frac{\exp(V_h(\wp_{ht}|\theta_{ht}))}{\sum_{\wp_j \in \mathfrak{P}_t} \exp(V_h(\wp_j|\theta_{ht}))} \tag{8}$$

---

[5]See for instance Hendel and Nevo (2006).

where $V_h(\wp_{ht}|\theta_{ht})$ is given by solving (1) with respect to c_{ht} for the different purchase option. In this specific model not all state variables are observed by the econometrician. Particularly, inventory and all beliefs are unobserved. Therefore, the econometrician has to integrate out over them. Thus, the likelihood of observing household h's behavior at time t is:

$$L_{ht} = \int_{\Theta} cp_{ht}(\theta_{ht})dF_{ht}(\theta_{ht}) \tag{9}$$

where the states θ_{ht} are from the econometrician's perspective randomly distributed over the set Θ according to the cdf $F_{ht}(\theta)$. The cdf of states depends on households and time; for instance, a household in an expensive neighborhood might rarely observe certain low prices.

Taking into account now that for each household many choices are observed I can write down the likelihood for all choices household i makes:

$$L_h = \prod_{t=1}^{T} \int_{\Theta} cp_{ht}(\theta_{ht})dF_{ht}(\theta_{ht}) \tag{10}$$

Here, T is the last period observed in the data.

While (10) looks innocuous it actually incorporates the dynamic programming problem. Note that calculating the likelihood of all household choices as the product of L_{ht} over time requires intertemporal independence. This only holds here because the current state θ_{ht} is used as conditioning argument. Remember, however, that states are not independent over time and in fact depend on previous states as well as consumer choices. Thus, the dynamics of the model are hidden in the probability distribution over states, $F_{ht}(\theta)$.

Now remember that even given the distributional parameters Υ and Σ we do not know V_h. That is because in the random coefficient setup we never uncover parameters for individual households Ψ^h. When calculating the likelihood this complicates matters in that I have to integrate over Ψ^h. Accounting for this fact, (10) has to be rewritten as:

$$L_h = \int_{\Psi} \prod_{t=1}^{T} \int_{\Theta} cp_{ht}(\theta_{ht}|\Psi)dF_{ht}(\theta_{ht}|\Psi)dG(\Psi|\Upsilon,\Sigma) \tag{11}$$

In the above equation I integrate out over household parameters Ψ^h when calculating choice probabilities. Now remember that the estimation takes the decisions of all households in the

data into account. Thus, the likelihood to be maximized is:

$$L = \prod_{h=1}^{H} \int_{\Psi} \prod_{t=1}^{T} \int_{\Theta} cp_{ht}(\theta_{ht}|\Psi)dF_{ht}(\theta_{ht}|\Psi)dG(\Psi|\Upsilon, \Sigma) \tag{12}$$

Typically, problems of this kind are solved using the nested fixed point (NFXP) algorithm introduced by Rust (1987). For each parameter try the expected value function (EV) has to be calculated for each state and action. It can then be used to calculate the new choice probabilities. The problem with this approach is the extremely large state space I face. With inventory, consumption shock, prices for ten brand-size combinations and beliefs for four brands there is a total of 20 state variables and three choice variables. Even if I were to perform those calculations on very coarse girds of five values per state variable, ten products and five consumption values, I would have to solve the EV at about 5 quadrillion points. Even if this were possible it would take a prohibitive amount of time since this calculation has to be repeated for every parameter try in the NFXP algorithm. To reduct the computational burden I employ several estimation techniques that significantly reduce computation time.

5.1 EV calculation

To calculate the EV I employ the Sieve Value Function Iteration (SVFI) method proposed by Arcidiacono et al. (2012). When faced with solving the dynamic programming problem in models with large state spaces many researchers have successfully relied on approximation techniques (e.g. Hendel and Nevo (2006) among others). SVFI tries to provide a better approximation while avoiding making the problem infeasible. The basic idea is to start out with an arbitrary guess for for the EV and then iteratively improve it over time.

To be more precise, SVFI approximates the EV using a simpler function (the "sieve function") of the state variables. In this case I choose the sum of state variables and pairwise interactions of state variables as my sieve function. Following Arcidiacono et al. (2012) I denote the j-th term in the sum by $w_j(\theta)$ and the coefficient associated with this term by $\rho(j)$. Then I can write the sieve function as:

$$\mathfrak{V}(\theta) = \sum_{j} \rho(j)w_j(\theta) \equiv \rho W(\theta) \tag{13}$$

where the second equality simplifies the expression by moving to matrix notation. Starting out with an arbitrary guess for ρ I can easily evaluate (1) for a given state θ and determine

the approximation error made. To update the sieve function I calculate the value function for a set of states; then I regress the resulting values on the state variables and use the resulting coefficients as the new ρ. That is, the new value for ρ, $\hat{\rho}$, can be found by solving:

$$\hat{\rho} = \underset{\rho}{\operatorname{argmin}} \sum_{\theta} \left(\rho W(\theta) - E[u_{ht}(c_{ht}, k_{ht}, j_{ht}) + \delta \rho W(\theta)|\theta] \right)^2 \tag{14}$$

The newly found $\hat{\rho}$ can then be use to calculate the new sieve approximation; this process is repeated until convergence.

This method is relatively fast; thus, it's possible to use a large number of interactions for the approximation (I use about 200 terms). Arcidiacono et al. (2012) show that by using a sufficiently large sieve space one can achieve a high degree of accuracy.

While the SVFI method allows me to find the EV function relatively quickly, incorporating this into the NFXP algorithm still requires resolving the dynamic programming problem a large number of times. Thus, this procedures would still be too slow for practical purposes. In order to avoid this problem I employ the importance campling method proposed by Ackerberg (2009).

5.2 Importance Sampling

Ackerberg (2009) shows how importance sampling can be used to dramatically reduce computation times of the maximum likelihood procedure. The idea is to perform the actual likelihood calculation a limited number of times for pre-drawn sets of structural parameters. When searching over the distributional parameters governing G a weighted sum of those likelihoods is calculated and changes in the estimates for Υ and Σ, $\hat{\Upsilon}$ and $\hat{\Sigma}$ respectively, only lead to reweighting which is relatively fast.

More precisely, in the first step I solve the dynamic programming problem following the method described above for 3,000 randomly drawn parameter sets $\tilde{\Psi}_1, \ldots, \tilde{\Psi}_{3000}$. For each household I then randomly draw 30 of those parameter sets on which to calculate the likelihood as given by (10). Even this limited number of likelihood calculations is time-sensitive due to the dynamic nature of the problem and the large state space. The reason for this is that an exact calculation would require calculating utilities over a relatively fine grid of state variables and actions; as noted above, the sheer number of calculations is too large to be feasible. In order to get around this problem I use a simulated likelihood approach.

A particular problem are initial inventories and beliefs. These can be difficult to estimate since they are based on unobserved behavior from before data was collected. To deal with

20

inventory I follow Hendel and Nevo's (2006) approach. That is, I first randomly draw start values for the variables in question. Then I let the model run through the first few observations in order to arrive at reasonable values for inventories. This method works well in my case since I observe households over a very long timespan. Remember from above that for the estimation I only use households for which I have at least three years of data. Dealing with initial beliefs in this way would be problematic since they systematically change over time. Therefore I estimate the initial beliefs in the main estimation routine.

At the end of this process I have a set of likelihoods $\{L_h^1, \ldots, L_h^{30}\}$ for each household. I now form the weighted sum of those likelihoods. Remember that under any set of distributional parameters $(\hat{\Upsilon}, \hat{\Sigma})$ I can determine the density of parameter draw $\tilde{\psi}_n$. This density, $g(\tilde{\psi}_n | \hat{\Upsilon}, \hat{\Sigma})$, is a measure for how likely it is to draw these parameters under the specified distribution. The idea behind importance sampling is to weight the predetermined likelihoods with a ratio of densities, $\frac{g(\tilde{\psi}_n | \hat{\Upsilon}, \hat{\Sigma})}{g(\tilde{\psi}_n | \hat{\Upsilon}_0, \hat{\Sigma}_0)}$. Here, denominator is the distribution from which the $\tilde{\psi}_n$ were drawn while the numerator is the distribution following from the current try of Υ and Σ. The aim is now to find $\hat{\Upsilon}$ and $\hat{\Sigma}$ such that large likelihoods receive a large weight. In other words, the final estimates $\hat{\Upsilon}$ and $\hat{\Sigma}$ are chosen in order to make parameter tries more likely to explain household behavior well. Formally, the objective is to maximize the following with respect to $\hat{\Upsilon}$ and $\hat{\Sigma}$:

$$\tilde{L} = \prod_{h=1}^{H} \left(\sum_{d=1}^{30} L_h^d \frac{g(\tilde{\psi}_h^d | \hat{\Upsilon}, \hat{\Sigma})}{g(\tilde{\psi}_h^d | \hat{\Upsilon}_0, \hat{\Sigma}_0)} \right) \tag{15}$$

This re-weighting only requires relatively simple calculations and can therefore be performed very fast; the much slower likelihood calculation does not have to be repeated which explains the time savings through this method.

5.3 Identification

Here I informally discuss the identification of the model. Identification is actually not very problematic, having mostly been established by previous literature. Standard learning models are unproblematic in terms of identification.[6] Wang (2012) establishes identification of an inventory model with random coefficients. Thus, the main concern is to separately identify the effects of learning and stockpiling. Both influence how much a consumer purchases at a given time. Learning leads households to make small purchases in order to be able to test

---

[6]See for instance Ackerberg (2003) for a discussion of identification in this context.

products before committing. In contrast to that, stockpiling is a tactic households use in order to make use of non-linear pricing and temporary price reductions and thus is an incentive to make few purchases of large quantities. The important thing to notice is that these incentives change over time. First, learning becomes less important the more experience a household has. Second, relative prices between differently sized packages change in my data so that the incentive to purchase for future use varies. Additionally, experience and prices also vary across households, allowing me to identify the random coefficients in the model.

5.4   Parameter Estimates

Tables 6 and 7 show the results of the structural estimation. Table 6 displays the parameter distributions and standard errors which I estimate using bootstrapping with 2000 subsamples. Table 7 converts the parameters reported in table 6 to means and standard deviations. Overall the parameter estimates seem sensible. All parameters have the expected signs and the magnitudes are reasonable.

The inventory cost parameters β_1 and β_2 are very low which is in line with common sense. Sugar substitutes are small and light-weight which makes them easy to transport and store. Thus, we would not expect the disutility from stockpiling to be high. Similarly, it is reasonable that the randomness of brand preferences (σ) is relatively low. For an industrially manufactured product it seems natural that differences in objective quality are small. This means that it is likely that households can learn about their preferences very quickly, which is supported by the estimated results for σ. The differences in product preferences are not very surprising either. The ordering of preferences conforms roughly to the ordering of market shares. Splenda, which over the course of my sample grows from a small producer to the dominant brand in the market, gives households by far the most utility. Natrataste is estimated to be a low-quality product; this explains the small market share it holds throughout the sample. Between these extremes we find the estimated preferences for Equal and Sweet'N Low. On average, households in my sample prefer Equal to Sweet'N Low (even though this order is reversed for almost 1 in 5 households). This is despite the fact that Sweet'N Low's market share is above that of Equal in every year in my sample. However, Sweet'N Low's market share falls much faster than Equal's as Splenda gains consumers.

To get a sense of how well the model matches the data I simulate household behavior based on the estimated parameters and compare the resulting purchase numbers to market shares observed in the data. Figure 3 shows the results of this exercise.

The figure shows that the model can explain market shares very well. While the simulated

Table 6: Parameter Estimates

		υ	σ
Marg. Utility of Consumption	α	-0.2352 (0.0219)	0.4796 (0.0151)
Cost of Inventory (linear)	β_1	5.1e-4 (1.3e-5)	2.9e-4 (9.1e-6)
Cost of Inventory (quadratic)	β_2	6.9e-4 (1.1e-5)	2.3e-4 (7.2e-6)
Marg. Utility of Income	ζ	-0.0141 (0.0269)	0.1543 (0.0178)
Std. Dev. of Shocks	σ	-1.4687 (0.0407)	0.4362 (0.0254)
Preference for Equal	γ_1	2.2037 (0.0860)	1.0655 (0.0628)
Preference for Splenda	γ_2	4.9183 (0.0902)	0.8005 (0.0633)
Preference for Sweet'N Low	γ_3	1.1041 (0.0898)	0.6388 (0.0600)
Preference for Natrataste	γ_4	0.8224 (0.0909)	0.9650 (0.0625)

Note: Standard deviations in parentheses.

Table 7: Parameter Distributions

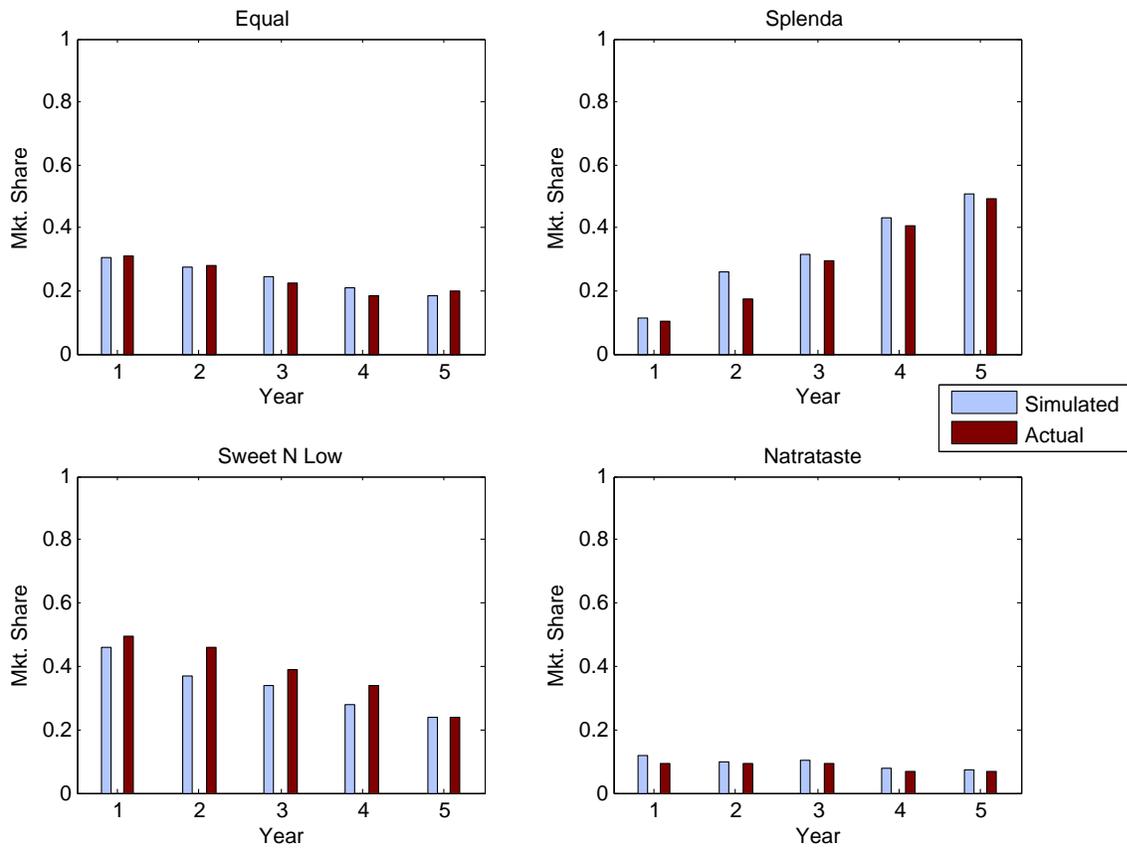|		Mean	Std. Dev.
Marg. Utility of Consumption	α	0.8372	0.4510
Cost of Inventory (linear)	β_1	5.1e-4	2.9e-4
Cost of Inventory (quadratic)	β_2	6.9e-4	2.3e-4
Marg. Utility of Income	ζ	0.9919	0.1549
Std. Dev. of Shocks	σ	0.2532	0.1159
Preference for Equal	γ_1	2.2037	1.0655
Preference for Splenda	γ_2	4.9183	0.8005
Preference for Sweet'N Low	γ_3	1.1041	0.6388
Preference for Natrataste	γ_4	0.8224	0.9650

Figure 3: Simulated and Actual Market Shares by Brand.

market share for Splenda is a bit too high in some years and the results for Sweet'N Low tend to come out a bit low the overall fit is very good. This suggests that my model is a good description of reality and also supports the estimation routine as a well-working method.

6   Counterfactual Study

In this section I study how the success of market entrants depends on timing. To do this, I look at the actual market entry of Dixie Crystals Sweet Thing (DCST) in 2004. This is complicated by the fact that I don't observe DCST in the household panel of my data. Therefore, I cannot simply re-estimate the model with an additional brand available. However, I do observe DCST in the store panel. This divide stems from DCST entering only some of the markets in my data together with the fact that the store panel covers many more markets than the household panel. In order to uncover the preference distribution for DCST (γ_5) I calibrate my model using the store panel for the Raleigh-Durham metropolitan area. Market shares in this market are closest to what I observe in the household panel. More precisely, I simulate households with the estimated parameters from the household panel and find the mean and standard deviation for γ_5 which best explain the market shares observed in the Raleigh-Durham store panel. This methods leads to a γ_5-distribution with mean 1.1912 and standard deviation 0.8214. Thus, the average consumer likes DCST slightly better than Natrataste and Sweet'N Low but significantly less than Equal and Splenda.

After this I can study the effect of the timing of market entry. I simulate household behavior (using prices observed in the Raleigh-Durham) market several times while varying the entry year for DCST. Before I can get to the actual computation I have to overcome two hurdles: First, since DCST entered the market in 2004 I do not observe any prices from 2001 to 2003, so I have to make a price choice for those simulations with earlier entry. Second, pricing is a strategic game and I need to consider how the incumbent brands would react to the market entry. For both of those issues I find the solution in my data. For the first problem I note that the real prices of DCST observed in my data are nearly constant over time. Thus, I think it is a good approximation to set the real price from 2001 to 2003 to the mean of the prices observed between 2004 and 2005. Secondly, I do not find any significant reaction of incumbent brands to DCST's market entry. This is probably due to the relatively small scale of the entry. (In my data DCST achieved only about 0.7% and 1.6% market share in 2004 and 2005, respectively.) Given this I find it reasonable to use observed prices from the data for this counterfactual study.

The analysis shows a strong effect of timing on the success of the market entry. As depicted in figure 4, DCST's market share in the year of entry would have been close to 10% had the entry occured in 2001. It falls continuously in entry year until the market share upon entry in 2005 is below 3.5%.[7] This is consistent with findings in previous studies that a brand achieved a higher market share the earlier it enters. However, this result adds to the literature in two ways: First, it offers learning as an explanation for this relationship. The later DCST enters the market the more consumers have already learned about Splenda's apparent superiority and are therefore hard to convince to buy the new product. Second, even when keeping the time since the market entry constant the timing of the entry matters. All market shares shown in figure 4 are for the respective first year on the market but nonetheless there are large differences between the various scenarios.

In order to uncover how important learning is for this result I run the same counterfactual simulation in a world without learning, i.e. in a setting where households perfectly know their product preferences. Figure 5 shows that this leads to significantly different results. In fact, in this case DCST's market share in its first year on the market is almost independent the timing of the entry. It always captures between 12.3% and 13.5% of the market with the exception of 2005, where an entry leads to only 11.4% market share. This slight drop is presumably because the real prices of some of the other brands decreased slightly in that year. Differences in outcomes between the best and the worst time of entry are only about 15% in this case compared to almost two-thirds in the scenario with learning. Thus, learning about preferences seems to play a major role in finding the optimal launch time for new products.

A second interesting question is who benefits from DCST's late entry. Figure 6 shows market shares in 2005 for DCST and Splenda. It is obvious that the Splenda's position is much stronger when DCST enters late. In fact, the combined market share of the two brands stays almost perfectly constant at roughly 56% when DCST's entry varies between 2001 and 2004 and only drops off a bit if the year of entry is 2005. The reason for this might be that Splenda itself was a relatively new product in 2001, having been introduced in 1999. Thus, in the early years of my panel many households might still be learning about Splenda's characteristics.[8] Were DCST to enter the market early then many households might make a choice about *which* of the new brands to sample.

---

[7]These market shares are higher than the ones reported for the actual entry in the Raleigh-Durham market. This is partially because in the counterfactual scenario I assume that every store carries DCST while after the actual entry many stores did not sell the new brand.

[8]This is consistent with the dramatic rise in Splenda's market share over time.
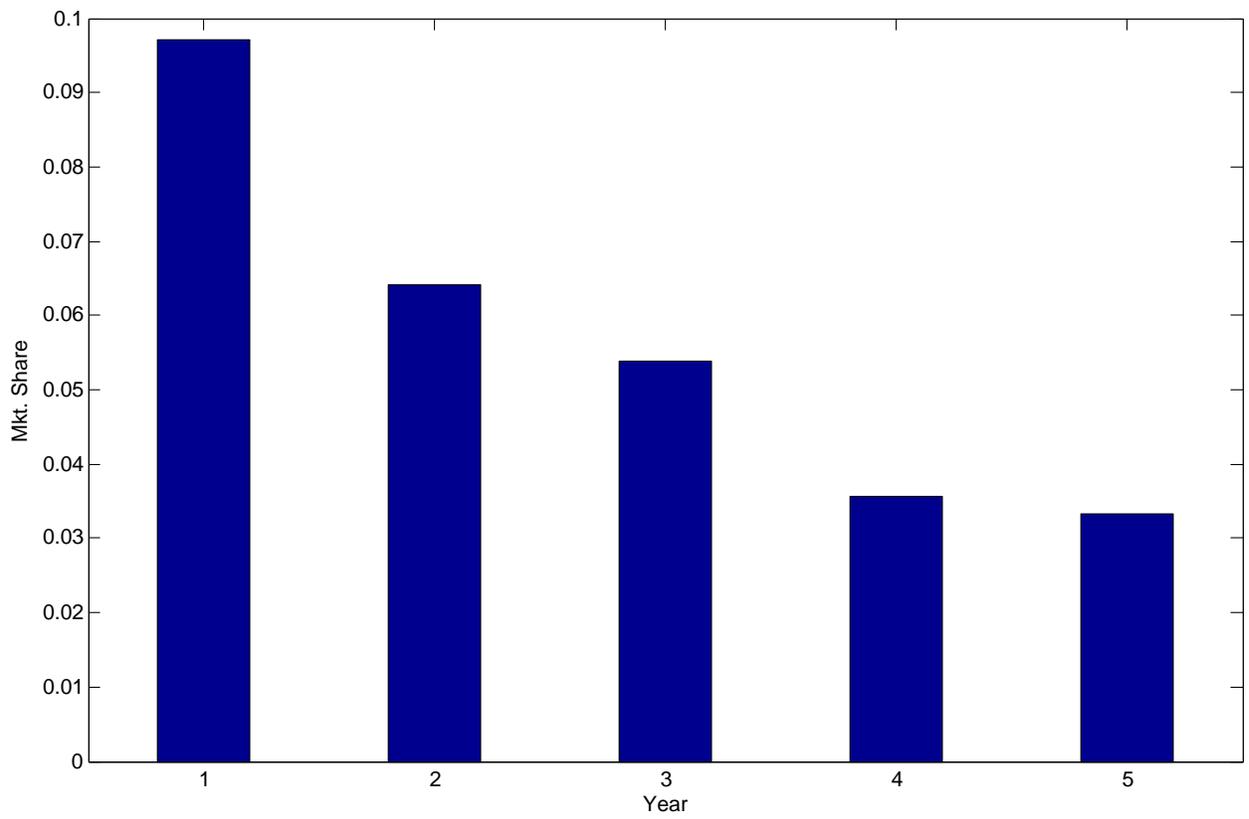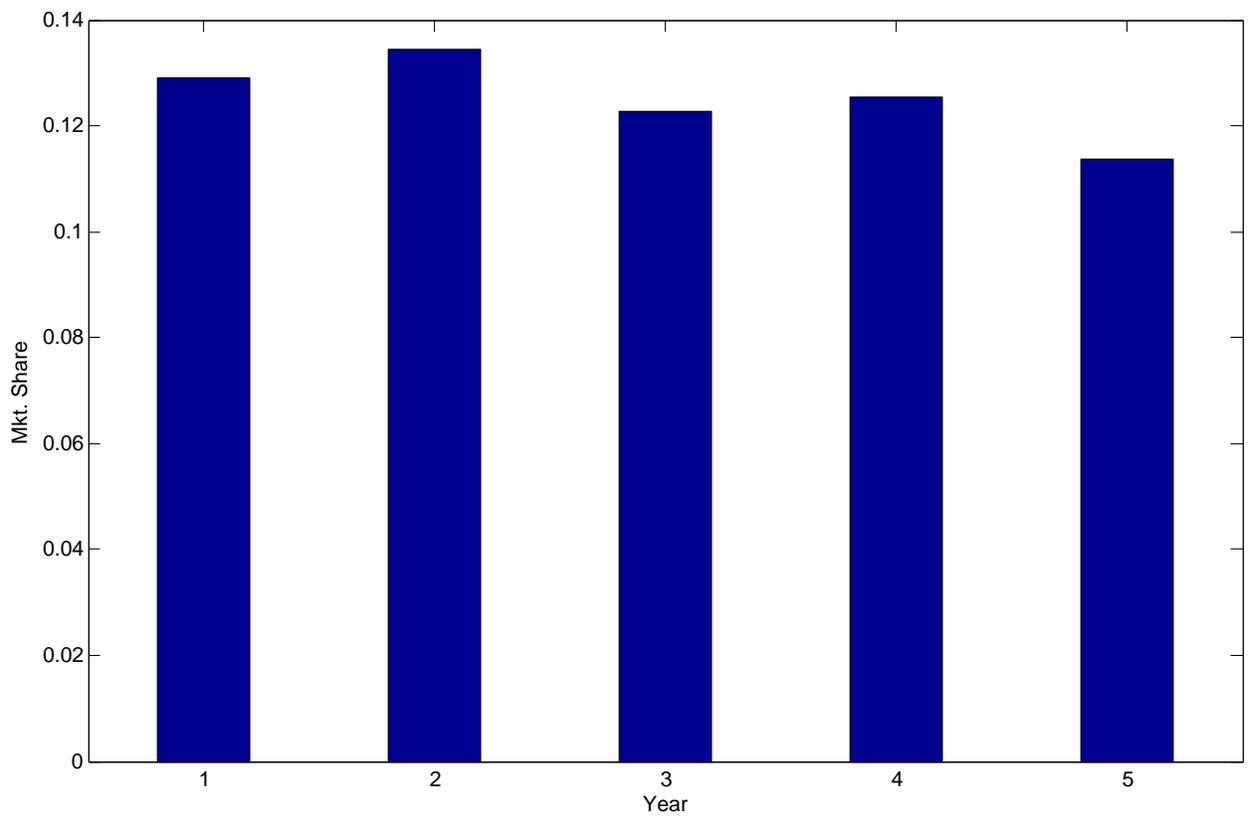
Figure 4: DCST Market Share in Year of Entry.

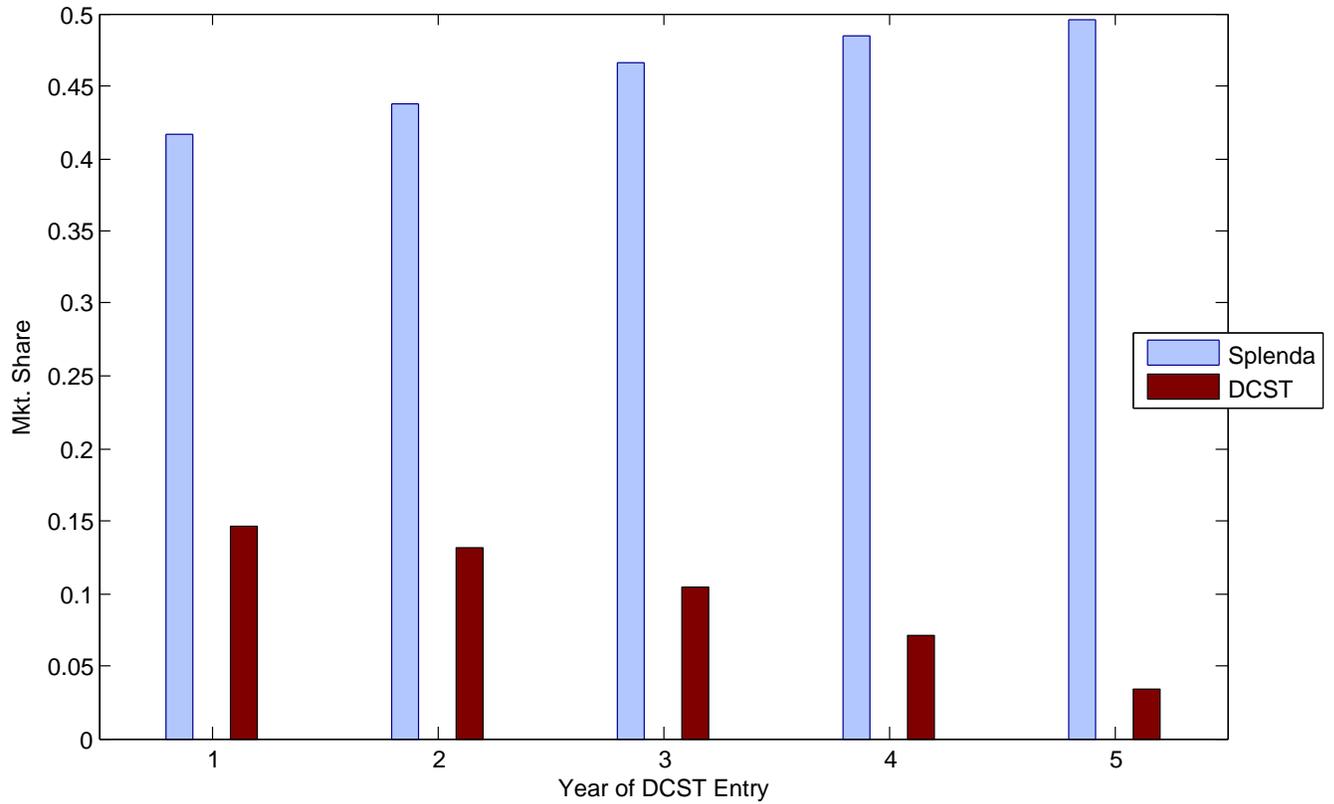Figure 5: DCST Market Share in Year of Entry without Learning.

Figure 6: DCST and Splenda Market Shares in 2005 for various Years of DCST Market Entry.

7 Conclusion

I develop and estimate a structural model that combines learning with stockpiling. I use this model to study how the timing of entry into a maturing market influences the success of the entrant. I find that timing plays an important role: The earlier the entry the more successful is the new brand. The reason for this seems to be that delaying the entry gives consumers additional time to learn about incumbent brands; thus they might already have found a brand they like when the entrant product finally becomes available. Additionally, most of the gains from early entry come directly from other recent entrants while older incumbent brands are barely affected. This again is consistent with learning by consumers.

My findings also might be one reason — besides more obvious arguments like demand shifts technological progress — why market entry tends to occur in bunches instead of uniformly over time.[9] Once a new product is in the market other producers have to enter as quickly as possible before consumers "settle" on that brand.

On the methodological side, this paper contributes to the literature by combining learning and stockpiling into a single model. This has been challenging due to the computational complexity of the task. However, new techniques allow me to estimate the model despite these difficulties. The model and its estimation can be used for studying numerous other questions regarding storable goods.

---

[9]See for instance Geroski (1995) for this stylized fact.

References

Ackerberg, D.A. 2003. "Advertising, learning, and consumer choice in experience good markets: an empirical examination." *International Economic Review*, 44(3): 1007–1040.

Ackerberg, D.A. 2009. "A new use of importance sampling to reduce computational burden in simulation estimation." *Quantitative Marketing and Economics*, 7(4): 343–376.

Arcidiacono, P., P. Bayer, F.A. Bugni, and J. James. 2012. "Approximating high-dimensional dynamic models: Sieve value function iteration."

Bayus, B.L., S. Jain, and A.G. Rao. 1997. "Too little, too early: Introduction timing and new product performance in the personal digital assistant industry." *Journal of Marketing Research*, 50–63.

Bronnenberg, B.J., S.K. Dhar, and J.P.H. Dubé. 2009. "Brand history, geography, and the persistence of brand shares." *Journal of Political Economy*, 117(1): 87–115.

Ching, A.T. 2010. "Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration." *International Journal of Industrial Organization*, 28(6): 619–638.

Ching, A., T. Erdem, and M. Keane. 2011. "Learning models: An assessment of progress, challenges and new developments." *Working Paper*.

Crawford, G.S., and M. Shum. 2005. "Uncertainty and learning in pharmaceutical demand." *Econometrica*, 73(4): 1137–1173.

Erdem, T., and M.P. Keane. 1996. "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets." *Marketing Science*, 15(1): 1–20.

Erdem, T., S. Imai, and M.P. Keane. 2003. "Brand and quantity choice dynamics under price uncertainty." *Quantitative Marketing and Economics*, 1(1): 5–64.

Geroski, P.A. 1995. "What do we know about entry?" *International Journal of Industrial Organization*, 13(4): 421–440.

Hartmann, W.R., and H.S. Nair. 2010. "Retail competition and the dynamics of demand for tied goods." *Marketing Science*, 29(2): 366–386.

Hendel, I., and A. Nevo. 2006. "Measuring the implications of sales and consumer inventory behavior." *Econometrica*, 74(6): 1637–1673.

Katz, M.L., and C. Shapiro. 1992. "Product introduction with network externalities." *The Journal of Industrial Economics*, 55–83.

Krider, R.E., and C.B. Weinberg. 1998. "Competitive dynamics and the introduction of new products: The motion picture timing game." *Journal of Marketing Research*, 1–15.

Osborne, M. 2011. "Consumer learning, switching costs, and heterogeneity: A structural examination." *Quantitative Marketing and Economics*, 9(1): 25–70.

Regibeau, P., and K.E. Rockett. 1996. "The timing of product introduction and the credibility of compatibility decisions." *International Journal of Industrial Organization*, 14(6): 801–823.

Urban, G.L., T. Carter, S. Gaskin, and Z. Mucha. 1986. "Market share rewards to pioneering brands: an empirical analysis and strategic implications." *Management Science*, 32(6): 645–659.

Wang, E. 2012. "Estimating the Distributional Impact of Taxes on Storable Goods: A Dynamic Demand Model with Random Coefficients." *Working Paper, Duke University.*