Consumer demand for variety: intertemporal effects of consumption, product switching and pricing policies

Ribeiro, Ricardo

Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD) London School of Economics (LSE)

15 August 2010

Consumer Demand for Variety: Intertemporal Effects of Consumption, Product Switching and Pricing Policies

Ricardo Ribeiro*

London School of Economics and Political Science

August 2010

Abstract

The concept of diminishing marginal utility is a cornerstone of economic theory. The consumption of a good typically creates satiation that diminishes the marginal utility of consuming more. Temporal satiation induces consumers to increase their stimulation level by seeking variety and therefore substitute towards other goods (substitutability across time) or other differentiated versions (products) of the good (substitutability across products). The literature on variety-seeking has developed along two strands, each focusing on only one type of substitutability. I specify a demand model that attempts to link these two strands of the literature. This issue is economically relevant because both types of substitutability are important for retailers and manufacturers in designing intertemporal price discrimination strategies. The consumer demand model specified allows consumption to have an enduring effect and the marginal utility of the different products to vary over consumption occasions. Consumers are assumed to make rational purchase decisions by taking into account, not only current and future satiation levels, but also prices and product choices. I then use the model to evaluate the demand implications of a major pricing policy change from hi-low pricing to an everyday low pricing strategy. I find evidence that consumption has a lasting effect on utility that induces substitutability across time and that the median consumer has a taste for variety in her product decisions. Consumers are found to be forward-looking with respect to the duration since the last purchase, to price expectations and product choices. Pricing policy simulations suggest that retailers may increase revenue by reducing the variance of prices, but that lowering the everyday level of prices may be unprofitable.

1 Introduction

"*Rachel*: Hi!

Chandler: Another cheesecake came! They delivered it to the wrong address again!

Rachel: So just bring it back downstairs, what's the problem?

Chandler: I can't seem to say goodbye.

Rachel: Are you serious?! Chandler, we ate an entire cheesecake two days ago and
     you want more?"

Friends, Episode 7-11, The one with All The Cheesecakes

The concept of diminishing marginal utility is a cornerstone of economic theory. The consumption of a good typically creates satiation that diminishes the marginal utility of consuming more. The length of time that the marginal utility is diminished is likely to vary across goods and, as Rachel and Chandler's cheesecake episode illustrates, across consumers. While it only took Chandler two days for his cheesecake marginal utility to return to pre-consumption levels, Rachel seemed still satiated (suggesting that the utility provided by her prior cheesecake consumption had not yet faded).

Temporal satiation induces consumers to increase their stimulation level by seeking variety and therefore substitute towards other consumption alternatives. In this paper, I define a forward-looking dynamic discrete choice model of demand that, similarly to Hartmann (2006), allows consumption to have an intertemporal effect: consuming produces a consumption capital stock that provides utility over time until it gradually depreciates. However, unlike Hartmann (2006), I consider a differentiated products setting that allows consumers to switch not only towards other goods, but also towards other differentiated versions (products) of the good.

In each shopping trip, consumers decide whether or not to purchase the good, and in case they decide to purchase, which quantity and product to buy. Consumers are assumed to make rational purchase decisions by taking into account, not only current and future satiation levels, but also prices and product choices. Price expectations are an important determinant of intertemporal substitution. If prices are expected to be higher in the future, consumers may anticipate their purchase decisions and vice-versa. Another important feature is consumer product choice. The marginal utility of different products is allowed to change over consumption occasions, depending on the switching costs of the individual. Consumers with low switching costs may exhibit shorter interpurchase durations than those that incur

in high switching costs whenever they alter their product choice.

This paper relates to the literature on variety-seeking developed from Jeuland (1978) and McAlister (1982). Jeuland (1978) explains variety-seeking behaviour by proposing that prior experience with a good decreases the consumer's utility for that good, which constitutes a direct application of the diminishing marginal utility concept. This explanation is predictive of the consumer's tendency to switch away from the most recently consumed good. McAlister (1982) then refined the explanation by proposing that prior experience with the *attributes* of a good decreases the consumer's utility for goods with similar attributes, which refocuses the diminishing marginal utility concept over attributes rather than goods. These two explanations have governed the subsequent development of an extensive literature centered on the implications of switching costs for consumer choice (see, for example, Keane (1997)). Recently, Hartmann (2006) extended the variety-seeking literature by allowing intertemporal effects of consumption, that is, by allowing consumption to have a lasting effect that diminishes the marginal utility of future consumption. However, the homogeneous nature of the good studied, golf, did not allow him to focus on product switching.

This paper attempts to link the two strands of the literature on variety-seeking by allowing consumers to substitute towards other goods (substitutability across time), as well as to other differentiated products of the same good (substitutability across products). This issue is economically relevant because both types of substitutability are important for retailers and manufacturers in designing intertemporal price discrimination strategies. I specify a consumer demand model which allows consumption to have an enduring effect and allows the marginal utility of the different products to vary over consumption occasions. The main contribution of the paper is to study how different pricing policies affect consumer demand for goods with such enduring effects of consumption and that are characterized by a high degree of differentiation. The model can then be used to simulate the demand implications of major pricing policy changes like a shift from hi-low pricing to everyday low pricing. To my knowledge, there is only one study that structurally addresses consumer response to such major policy changes, Erdem *et al.* (2003). However, they studied storable goods and do not allow for switching costs. I find similar patterns deriving from an entirely different source of dynamics, the stock of past consumption.

The state space implied by a dynamic problem where forward-looking consumers make optimal decisions in light of current and future satiation levels, prices and product choices is, in a differentiated products setting, extremely large for practical estimation. In order to reduce the dimensionality of the state space, I adopt a multi-stage budgeting approach that decomposes the consumers decision into a quantity choice and a product choice (see

3

Aguirregabiria (2002) and Hendel and Nevo (2006a) for similar dynamic applications of Gorman (1971)'s approach). Under the set of assumptions discussed below, I show that the consumer product choice *conditional on the quantity purchased* does not depend on dynamic considerations. This simplifies the estimation of many of the parameters of the demand model, while the remaining ones are estimated solving a simplified dynamic problem that involves only quantity and timing decisions.

I estimate the different stages of the model by maximum-likelihood and solve the dynamic programming problem by using value function parametric approximation with policy function iteration in the lines of Benitez-Silva *et al.* (2000). In order to control for unobserved heterogeneity, I incorporate a rich specification. This is important since in both stages unobserved consumer heterogeneity may confound the inference of true state-dependence effects. As Heckman (1981) points out, if households have different preferences "and if these differences are not properly controlled, previous experience may appear to be a determinant (...) of future experience solely because it is a proxy for temporally persistent unobservables that determine choices." The state dependence in the product decision arises because the marginal utility of different products is allowed to change over consumption occasions, while the state dependence in the quantity decision is induced by the consumption capital stock. For reasons I discuss below, I incorporate *observable* heterogeneity in the product choice and a *continuous distribution* of consumer heterogeneity in the quantity decision.

I apply the model to an indulgence good: ice cream (and related frozen desserts). The reason is twofold. First, ice cream constitutes the textbook illustration of the diminishing marginal utility concept and the industry is characterized by a high degree of product differentiation. Second, the temptation nature of the good can (and in fact does, for the empirical application considered) make stockpiling limited in relevance (and in particular duration). This is important because in a context where temporary price promotions are a key marketing tool, if consumers respond to temporary price cuts by accelerating (anticipating) purchases and hold inventories for future consumption (i.e. stockpile), the separate identification of satiation and stockpiling would be somewhat problematic. I show below that, even though consumers do anticipate purchases in response to temporary price promotions, they do not stockpile, maybe because of the temptation feature of the good.

I find evidence that consumption has a lasting effect on utility that induces substitutability across time and that the median consumer has a taste for variety in her product decisions. Consumers are found to be forward-looking with respect to the duration since the last purchase, to price expectations and product choices. Pricing policy simulations suggest that retailers may increase revenue by reducing the variance of prices, but that lowering the

4

everyday level of prices may be unprofitable.

2   Data Description and Preliminary Analysis

I use Information Resources Inc. (IRI) scanner data collected from June 1991 to May 1993 in two separate submarkets of a large Midwest city. The dataset covers 24 different product categories at both the store and household levels. The former includes weekly sales, prices, and promotional activities for each universal product code (UPC) in nine supermarkets, belonging to different chains, while the latter tracks the store visits of 548 households and includes when and how much each household spent in her shopping trips.

I estimate the model for an indulgence good category: ice cream and related frozen desserts. Frozen desserts are offered in four segments: regular ice cream, diet ice cream, frozen yoghurt and ice milk. Regular ice creams account for 67% of the volume purchased, with diet ice creams and frozen yoghurt roughly splitting the remaining of the market. The market share of ice milk is less than one percent. Ice creams come in a limited number of package sizes, with the top four sizes accounting for more than 99% of the market: 64 $oz.$ (72.3%), 16 $oz.$ (11.5%), 160 $oz.$ (10.8%) and 32 $oz.$ (4.8%). The choice set available to the households is substantial. The average supermarket in the sample carries 170 different frozen dessert products (from 20 brands) on a weekly basis. I defined a product as a segment-brand-flavour combination so that, for example, Häagen-Dazs Vanilla Ice Cream, Häagen-Dazs Chocolate Ice Cream, and Häagen-Dazs Vanilla Frozen Yoghurt are classified as distinct products.

Kemps is the dominant brand with 23% volume market share, followed by Breyers and Wessanen's Value Pack (both with 12%), Dreyer's (10%) and Häagen-Dazs (6%). Store private labels account for 3.5% of the market. The most popular flavours are vanilla (21%), chocolate (9%), neapolitan (7%), strawberry (5%) and chocolate chip (5%), although a typical supermarket would carry an average of 84 different flavours, each week. In contrast with the moderated brand and flavour concentration, there is substantial market fragmentation at the product level. Breyers Vanilla Ice Cream is the market leader with a 2.6% volume market share.

The median household has two members, an income between $25,000$ and $35,000$ dollars. I conduct the subsequent analysis using a subset of the sampled households selected based on three criteria. First, I eliminated consumers recorded purchasing in supermarkets for which no price data is available. An alternative approach could have been to include those

TABLE 1
*Volume Market Shares**

Panel A: Product level		Household				Store	
		All		Subset			
		S	CS	S	CS	S	CS
1	Wessanen's Value Pack NY Vanilla IC	3.85	3.85	4.14	4.14	2.03	2.03
2	Wessanen's Value Pack Vanilla IC	2.71	6.56	1.96	6.10	2.58	4.61
3	Wessanen's Value Pack Neapolitan IC	2.06	8.62	2.59	8.69	1.78	6.39
4	Fieldcrest Vanilla IC	2.05	10.67	2.81	11.50	1.22	7.61
5	Kemps Vanilla FY	1.97	12.64	5.26	16.76	1.13	8.74
6	Kemps Vanilla IC	1.69	14.33	1.37	18.13	2.15	10.89
7	Breyers Vanilla IC	1.51	15.84	1.62	19.75	2.60	13.49
8	Wessanen's Value Pack Chocolate IC	1.41	17.25	1.26	21.01	1.14	14.63

Panel B: Brand level		Household				Store	
		All		Subset			
		S	CS	S	CS	S	CS
1	Kemps	21.22	21.22	19.59	19.59	23.37	23.37
2	Wessanen's Value Pack	13.69	34.91	12.71	32.30	11.73	35.10
3	Breyers	9.22	44.13	10.04	42.34	11.86	46.96
4	Dreyer's	8.89	53.02	5.34	47.68	9.88	56.84
5	Sealtest	6.90	59.92	10.70	58.38	4.27	61.11
6	Fieldcrest	4.10	64.02	5.66	64.04	3.43	64.54
7	Dean Foods	3.75	67.77	3.85	67.89	2.71	67.25
8	Häagen-Dazs	3.62	71.39	1.53	69.42	5.81	73.06

Panel C: Flavour level		Household				Store	
		All		Subset			
		S	CS	S	CS	S	CS
1	Vanilla	22.20	22.20	26.38	26.42	21.10	21.10
2	Chocolate	8.00	30.02	9.04	35.46	8.59	29.69
3	Neapolitan	7.21	37.41	11.47	46.93	6.50	36.19
4	New York (NY) Vanilla	5.49	42.90	5.52	52.45	3.63	39.82
5	Strawberry	5.36	48.26	4.48	56.93	4.99	44.81
6	Butter Pecan	4.01	52.27	3.21	60.14	3.79	48.60
7	Chocolate Chip	2.69	54.96	2.59	62.73	4.63	53.23
8	Pistachio	2.39	57.35	2.94	65.67	1.31	54.54

* Columns labeled S denote market shares and columns labeled CS denote cumulative market shares. IQ stands for an ice cream product and FY for a frozen yoghurt product.

households and either (i) eliminate the purchases in unsampled stores as if they never happened, or (ii) assume some cross-store price pattern and generate price data to be imputed for those purchases. All solutions potentially could introduce bias in the analysis. I opted for the elimination after ensuring the subset sample was representative, an issue I discuss below. Second, computational barriers compelled me to eliminate consumers that purchased more than two items of ice cream in a shopping visit or bought non-representative package sizes. Their inclusion would increase the dimensionality of the state space to a degree that made the structural estimation computationally infeasible. Finally, I eliminated households that made less than 10 purchases of ice creams over the total sample period since they are likely to be either (i) not regularly in the market, or (ii) purchasing in alternative stores. This reduced the sample size from 548 to 115 consumers, who made a total of 17,899 supermarket trips and 2,822 ice-cream purchases.

An important question is obviously whether the subset sample is representative of the whole population buying at these supermarkets. Table 1 addresses this question by reporting, for the different samples, the top-8 products, brands and flavours in terms of their volume market share. The simple comparison of the columns show that, with minor exceptions, the product, brand and flavour market shares in the different samples are very similar, which is suggestive that the subset sample is reasonably representative.

2.1   Substitutability Across Time

In this section, I examine the shopping behaviour of consumers and the frequency of their purchasing patterns for the ice cream category as a whole. Table 2, Panel A presents summary statistics for the consumers supermarket trips. Although there is evidence of substantial heterogeneity across consumers with regard to their shopping behaviour, the median consumer in the sample visits a supermarket every three days to a total of 98 times over the observed sample. This consumer shops in two different supermarkets, but concentrates her purchases on a single one. In order to compute the consumers intertrip duration, I use the first six months in the sample to generate an initial trip for each household. I will discuss below that these first six months will also be instrumental in generating an initial product choice for each consumer to avoid spurious switching.

Table 2, Panel B displays some summary statistics of households ice-cream purchasing patterns. The results suggest substantial heterogeneity also at this level, with the median consumer making a single-item purchase of 64 $oz.$ of ice-cream every 16 days to a total of 13 purchases over the sample period.

TABLE 2
*Consumer Category Purchasing Patterns**

Panel A: Supermarket trips

	Mean	Median	Std	Min	Max
Number of Trips	114	98.0	59.9	36.0	317
Days from Previous Trip	4.75	3.00	5.05	0.00	75.0
Number of Stores Visited	1.95	2.00	0.79	1.00	4.00
Store HHI	0.86	1.00	0.20	0.36	1.00

Panel B: Ice cream purchases

	Mean	Median	Std	Min	Max
Number of Purchases	17.7	13.0	15.0	3.00	126
Volume	63.3	64.0	31.5	16.0	160
Multiple-item Purchases	0.12	0.00	0.32	0.00	1.00
Days from Previous Purchases	30.0	16.0	37.8	0.00	311

* For Number of Trips, Number of Stores Visited and StoreHHI, an observation is a household. For Days from Previous Trip, an observation is a trip instance. For all other statistics, an observation is a purchase instance. Store HHI denote the household's Herfindahl-Hirschman index for ice-cream volume purchases.

I now move on to examine the hypothesis that consumption has a lasting effect that diminishes the marginal utility of future consumption. If the magnitude of this effect is such that induces consumers to vary their choice of dessert, the probability of purchase will be related to how long it has been since their last purchase. If, on the other hand, the magnitude of the effect is small, then the probability of purchase will not depend on the interpurchase duration. Figure 1 displays the purchase hazard rate by no-purchase spell duration in days. The hazard rate denotes here the probability that you purchase if you have not purchased up to now. The pattern illustrated provides some support for the duration dependence argument: there is evidence of a non-linear relationship between the probability of purchase and the duration since the last purchase. The hazard rate is quite low immediately after a purchase, then gradually increases until day 7, after what it exhibits a gradual, although rather jagged, downward trend (an interesting aspect of this hazard rate relates to its recurring spikes, an issue I address below). However, the downward trend of the hazard rate suggests that, in contrast to the initial argument of this paper, the probability of purchase seems to decrease (and not increase) with the duration since last purchase. There are two possible explanations for this behaviour (and that illustrate the well known problem that unobserved heterogeneity can be confounded with state-dependence). Either the utility from consuming ice cream does in fact decrease with duration (positive state-dependence), or alternatively, the utility increases with duration from last purchase (negative state-dependence), but there exists a group of low-demand consumers who are more likely to exhibit longer interpurchase durations (heterogeneity). In order to evaluate the degree

of consumer unobserved heterogeneity, I re-compute the hazard rate at the *consumer-level*. The consumer-level hazard rate denotes here the probability that a consumer purchases if she has not purchased up to now. Figure 2 displays, for each duration spell, the mean of the probability of purchase across consumers, as well as the interval limited by that mean value plus and minus one standard deviation (subject to a non-negativity constrain). The high standard deviation around the mean indicates substantial heterogeneity across consumers, which is suggestive of the importance of controlling for unobserved heterogeneity in the structural estimation.

While the duration dependence of ice cream purchases implied by Figures 1 and 2 can be consistent with variety-seeking behaviour induced by a diminishing marginal utility, it can also be consistent with the main alternative theory: if consumers respond to temporary price cuts by accelerating (anticipating) purchases and hold inventories for future consumption (i.e. stockpile), the probability of purchase will also be duration-dependent. Temporary price promotions are an important marketing tool in the pricing strategy of many nondurable goods and ice creams are no exception. The ice cream prices in the sample display a classic high-low pattern: products have a "regular (modal) level" that remains constant for long periods of time with occasional temporary reductions. Figure 3 displays, as an illustration, the price of Dreyer's Vanilla Ice Cream 64 *oz.* over the sample weeks in a typical supermarket. The price is at the "regular level" ($4.59) for 57% of the weeks. Defining a sale to be a price reduction of at least 5% (below the modal level), it is on sale for 31% of the time, with the average price discount being $1.61. If we consider the sample as a whole, untabulated statistics show that prices are, on average, 66% at the "regular level" and 26% on sale (with an average discount of $0.70). In such an environment, consumers may respond to temporary price cuts by accelerating (anticipating) purchases and stockpile.

Table 3 addresses the purchase acceleration effect by comparing household level sale and nonsale purchasing patterns. The first column displays averages during nonsale purchases. The following columns examine the difference towards a sale purchase, decomposing the total difference into a within and a between households effects. As before, a sale is defined as any price at least 5% below the model price of a store-UPC combination over the observed period. I focus the analysis on the within column,[2] that compares the household purchasing patterns over time. The evidence seems to indicate that consumers do respond to temporary price cuts. Unsurprisingly, the results suggest that households tend to shorten their duration from

---

[2]The results from the between column in Table 2.3 suggest substantial heterogeneity in how consumers respond to temporary price cuts, with households that purchase more frequently on sale, buying larger volumes and less frequently. This reinforces the need for the structural model to control for consumer heterogeneity.
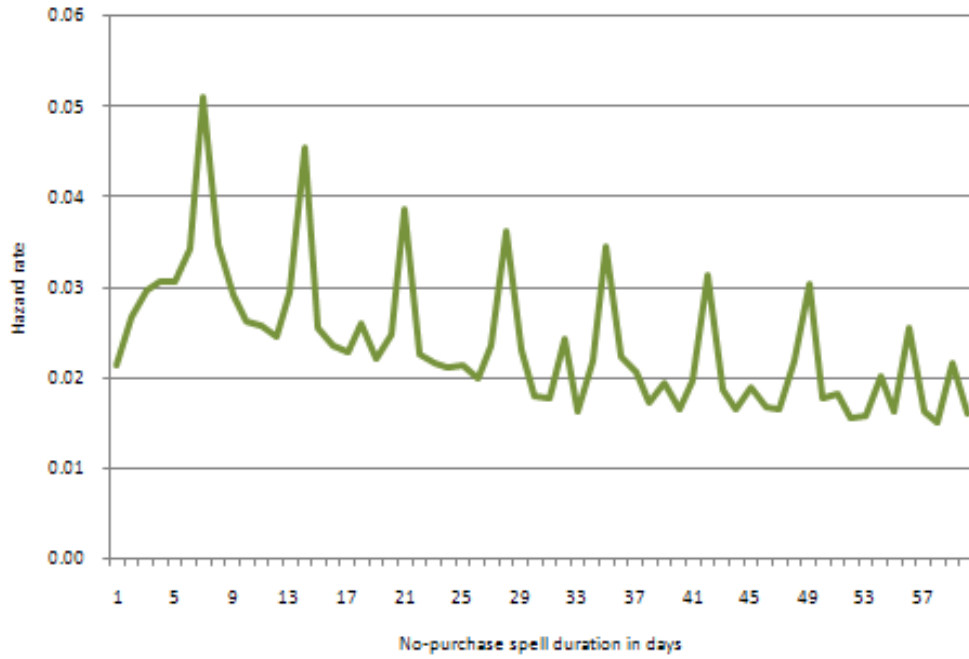
Figure 1
Purchase Hazard

Figure 2
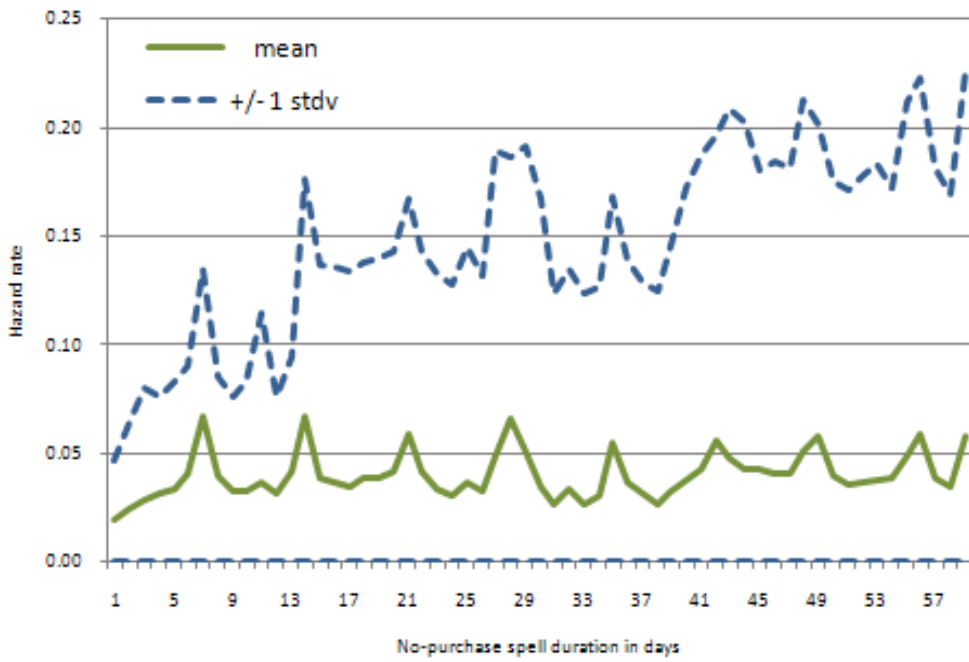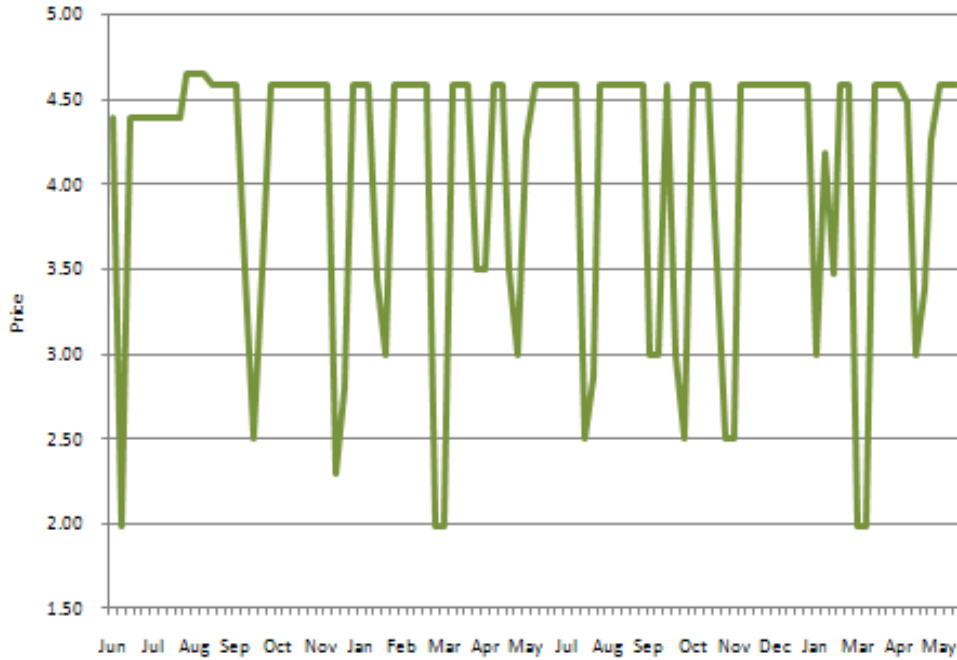*Consumer-level Purchase Hazard*

10

FIGURE 3
Price Example: Dreyer's Vanilla Ice Cream

previous purchase (between 3-4 days) and to increase their volume purchases (by roughly 17%), when buying on sale.

The interesting question is whether this response by consumers translates into a consumption effect or merely represents a demand-anticipation effect with households stockpiling for future consumption. In order to examine this question, I follow Hendel and Nevo (2006a) and examine households interpurchase duration to the next purchase. The idea here is that if consumers do stockpile, then the duration to next purchase is expected to be longer for large volume purchases (like purchases on sale). The results from Table 3 show that there is no significant difference in the duration forward to next purchase between sale and non-sale purchases. The comparison of the quantity and duration effects seem to indicate that stockpiling may not be a relevant feature of ice-cream demand and that the quantity effect induced by temporary price reduction substantiates a consumption effect. In order to examine the robustness of this conclusion, I also compared the duration forward to next purchased when consumers buy an above average volume. The results of those regressions (which are untabulated) are consistent with the above conclusion. The difference in interpurchase duration is again not significantly different from zero. The alternative theory that state-dependence in the probability of purchase is due to stockpiling can not explain these results. Furthermore, the analysis of how the additional quantity is bought is consistent with

	Average during Nonsale	Difference during Sale			
			Consumers		Week F.E.
		Total	Within	Between	
Volume (oz.)	57.7 (3.08)	11.8 (2.26)	9.91 (1.42)	23.3 (7.70)	No
	–	12.1 (2.30)	10.1 (1.51)	22.5 (12.7)	Yes
Units	1.05 (0.01)	0.16 (0.02)	0.17 (0.02)	0.13 (0.04)	No
	–	0.16 (0.02)	0.17 (0.02)	0.07 (0.08)	Yes
Average Package Size	56.1 (2.95)	2.27 (1.94)	-0.31 (0.91)	14.3 (6.91)	No
	–	2.22 (1.97)	-0.18 (0.90)	13.0 (12.2)	Yes
Days from Previous Purchase	27.0 (2.40)	0.22 (1.83)	-3.69 (1.54)	17.6 (7.04)	No
	–	-0.13 (1.77)	-4.00 (1.51)	19.8 (9.84)	Yes
Days to Next Purchase	25.9 (2.31)	2.55 (1.65)	-0.14 (1.39)	9.23 (7.03)	No
	–	2.59 (1.64)	-0.14 (1.44)	28.3 (11.5)	Yes

* An observation denotes a purchase instance. Standard errors clustered by household in parentheses (except for the between analysis).

the variety-seeking theory. When purchasing on sale, consumers do not significantly change their average package size. Instead, they purchase more units of ice cream. This supports the variety-seeking story since if the increased volume translates into increased consumption, then purchasing multiple-items is a sensible strategy to deal with the diminishing marginal utility from consumption.

Having addressed the issue of eventual stockpiling behaviour, I now move on to address another somewhat problematic issue. In this paper, I model consumption to have a lasting effect that diminishes the marginal utility of future consumption. However, I do not observe the actual time and magnitude of consumption. So, I am forced to infer it from purchase choices. The results from Table 3, Panel A provide some evidence that, not only consumers do not anticipate purchases to hold inventories for future consumption, as already discussed, but also that utility does not depend on the stock of past consumption. If consumption did create a stock, then duration to next purchase would increase with the size purchased, which it does not. That said, following Table 3, the only relevant variable that may affect the marginal utility of future consumption is the timing of current consumption. As I discuss below, due to the temptation nature of the good, assuming that the time of consumption coincides with the time of purchase is not unreasonable. At least for most people, in line with what Erdem *et al.* (2003) argue, ice creams are technologically, but not practically storable over more than a few days. It may seem inconsistent to assume that consumption has a lasting affect that induces intertemporal substitution in purchases while assuming that the good held in inventory has a temptation feature. These assumptions are however consistent

Table 4
Category Purchasing Patterns: Seasonality*

	Average during Nonsummer	Difference during Summer		
			Consumers	
		Total	Within	Between
Volume (oz.)	63.1 (0.78)	0.76 (1.74)	1.86 (1.25)	2.54 (18.7)
Units	1.12 (0.01)	0.00 (0.02)	0.00 (0.02)	-0.17 (0.10)
Average Package Size	57.0 (0.63)	0.94 (1.39)	1.58 (0.86)	12.1 (16.4)
Days from Previous Purchase	27.0 (0.86)	0.74 (1.86)	-0.72 (1.68)	28.8 (13.8)
Days to Next Purchase	27.1 (0.86)	0.09 (1.86)	-1.01 (1.68)	15.2 (16.4)

* An observation denotes a purchase instance. Standard errors clustered by household in parentheses (except for the between analysis).

with observed behaviour, since consumers seem to depreciate the costs of goods they have in inventory (see Gourville and Soman (1998) and Prelec and Loewenstein (1997)).

I now move on to describe two other timing aspects of consumers ice cream category purchasing patterns. I begin by addressing seasonality. If the decisions of consumers are seasonal, then the structural model must reflect this feature. Table 4 addresses this question by comparing household level summer and nonsummer purchasing patterns. The first column displays averages during nonsummer purchases while the following columns examine the difference towards a summer purchase, again decomposing the total difference into a within and a between consumers effects. The results suggest that summer does not induce a significant difference in the purchasing patterns of households, at any dimension: volume, units, average package size, days from previous purchase or days to next purchase. This holds both within and across consumers. In order to examine the robustness of this conclusion, I replicated this analysis to compare the consumers purchasing patterns in winter and nonwinter seasons. The results of those regressions (which are untabulated) are consistent with the above conclusion. They show no significant difference in the associated purchasing patterns. Surprisingly, seasonality does not seem therefore to be an important feature in the purchasing decision of ice cream and related frozen desserts.

Another timing aspect of consumers ice cream category choice patterns relates to the purchase day. If consumers are more likely to purchase on a particular day of the week or weekend, then the structural model must somehow incorporate it. Untabulated statistics show no evidence of a clear preference towards a given day of the week, when comparing *across* consumers. However, Figure 1 illustrated an interesting pattern. The probability of purchase spikes at every seven days (and *exactly* every seven days), which suggests that even though no preference exists *across* consumers, each consumer seems to have a preferred day of the week to purchase ice creams - maybe at their main weekly shopping trip. This

TABLE 5
Consumer Product Choice Behaviour*

Panel A: Product level					
	Mean	Median	Std	Min	Max
CR1	0.38	0.33	0.19	0.11	0.95
CR5	0.83	0.84	0.16	0.46	1.00
HHI	0.26	0.20	0.17	0.07	0.91
Number of different products	8.27	8.00	4.18	2.00	26.00
Probability of successive product switching	0.77	0.82	0.22	0.15	1.00
Probability of product exploration switching	0.45	0.42	0.26	0.00	1.00
Panel B: Brand level					
	Mean	Median	Std	Min	Max
CR1	0.57	0.53	0.23	0.19	1.00
CR5	0.97	1.00	0.06	0.73	1.00
HHI	0.46	0.38	0.25	0.13	1.00
Number of different brands	4.43	4.00	2.10	1.00	11.00
Probability of successive brand switching	0.57	0.62	0.27	0.03	1.00
Probability of brand exploration switching	0.33	0.30	0.20	0.03	1.00

* An observation is a household. CRm and HHI denote the household's m-product (brand) volume concentration ratio, and Herfindahl-Hirschman index, respectively.

constitutes a feature of consumer behaviour that must be incorporated into the structural model.

2.2   Substitutability Across Products

Having described the shopping behaviour and purchasing patterns of consumers for the ice cream category as a whole, I now move on to examine their product choice patterns. Table 5 displays household-level concentration and variety-seeking measures for ice cream product and brand choices. Table 5, Panel A displays the descriptive statistics for the product level measures. The median consumer buys 8 different products over the sample period and fragments her volume purchases considerably as the relatively low household-level concentration ratios (CRm) and Herfindahl-Hirschman index (HHI) suggest. Nevertheless, there is evidence of substantial heterogeneity across consumers as indicated by the large range intervals and standard deviation of the several concentration measures. So, although some households show evidence of considerable product fragmentation, others concentrate their purchases on a relatively small number of products.

Having examined product choice concentration, I now move on to examine a measure of product switching, following Menon and Kahn (1995). The probability of *successive* product

switching denotes the proportion of consumer purchases that involved switching, where a switch is defined as occurring each time the product(s) chosen on a purchase occasion is different from those chosen on the immediately preceding purchase instance. This is consistent with Faison (1977) and Venkatesan (1973). Counting switching from the beginning of the sample period would generate spurious switching. Therefore, as discussed previously, I use the first six months in the sample to generate an initial product choice for each consumer. This approach is similar to Shum (2004) and Pozzi (2009). The descriptive statistics for the probability of successive product switching suggest substantial heterogeneity across consumers, with the median household switching from the immediately preceding products in 82% of her purchases.

An alternative approach would be to define a switch as occurring each time the product chosen on a purchase occasion is different from any of the preceding choices, following Faison (1977) and Pessemier (1985). The two definitions differ in the idea of variety-seeking that may capture. While the latter definition implicitly assumes that the level of stimulation of a household can only be increased by exploring new products, i.e. products that the consumer never tried before, the successive switching definition assumes that the level of stimulation of a household can be increased by alternating from one product to another, even if the products in the switching set are all familiar. A simple comparison of the probability of switching according to the two definitions suggests that the proportion of switching involving familiar products should not be neglected.

Table 5, Panel B presents descriptive statistics for the same concentration and variety-seeking measures, but aggregated at the brand level. The purchases of the median consumer show a higher degree of concentration and a lower probability of switching when compared with her product choice patterns, which may be suggestive of the relative importance of different flavours and product types in increasing the level of stimulation of a household.

One problem with inferring variety-seeking from product switching is that unobserved heterogeneity can be confounded with product state-dependence. The identification problem arises because a consumer may exhibit high product switching by repeatedly alternating products in her purchases either because of a weak unobserved, idiosyncratic preference for the different products or because she has a taste for variety. In order to evaluate the importance of product preferences, I examine the association between product switching behaviour and product choice. The dependent variable is a product preference measure in the lines of Simonson and Winer (1992). Each consumer purchase is associated with a score equal to the volume market share of the corresponding product (or products in

15

TABLE 6
Product State-Dependence vs Product Preference*

	(1)	(2)	(3)	(4)
Product State-Dependence	-0.28	-0.31	-0.17	-0.10
	(0.06)	(0.06)	(0.03)	(0.01)
Marketing-Mix				
Price		0.04	0.07	0.01
		(0.01)	(0.01)	(0.01)
Feature		0.05	0.07	0.01
		(0.02)	(0.02)	(0.01)
Display		-0.00	0.01	-0.02
		(0.03)	(0.02)	(0.02)
Household F.E.	No	No	Yes	Yes
Product F.E.	No	No	No	Yes
R2	0.26	0.28	0.64	0.85

* An observation is a purchase instance by a household. Standard errors clustered by households in parentheses.

case it is a multiple-item purchase) in the consumer's shopping history.[3] Products with high consumer-level market shares are assumed to correspond to products for which the consumer has a strong preference, given their weight in the household shopping basket. This assumption is of course problematic, but it allows me to illustrate the high degree of unobserved product heterogeneity. Table 6 presents the OLS results of regressing the product preference score of each purchase on a product state-dependence variable that keeps track of the number of product switches from the immediately preceding purchase instance. I include marketing-mix variables as covariates: price and two types of promotional activities: feature (defined as any type of retailer product advertising) and display (defined as any type of special store display).[4] The different specifications vary on the degree of controls included. The results suggest a significant negative association between product switching and the product preference score, implying that consumers tend to switch more towards less preferred products. This association weakens as additional heterogeneity is incorporated, which is indicative of the importance of controlling for unobserved product heterogeneity in the structural estimation.

A final problem with inferring variety-seeking from product switching is that product

---

[3]As an illustration consider the hypothetical example of a consumer that, over her shopping history, purchases 100 *oz.* of product A, 50 *oz.* of product W and 50 *oz.* of product B. Each item purchase of product A will receive a product preference score of 0.50 (100/200).

[4]The data includes several categories of feature and display. I aggregate across the different categories to feature/no feature and display/no display.

unavailability may generate spurious switching. In order to address this concern, I need to separate *true* product switching from switching *induced* by product unavailability. However, as the IRI dataset does not include information on product availability, I have to infer it from store product sales. I consider that a product was available in a given week and supermarket combination if the store sold at least one unit of the product in that week.[5] I then compute, for each consumer, the proportion of *true* product switching, i.e. switching that occurs despite the products chosen on the immediately preceding purchase instance being available (as measured by the proxy). This analysis (which is untabulated) seems to suggest that most product switching is not induced by unavailability: the median proportion of *true* product switching is 83%, the average is 80%, the 25th percentile is 66%, and the 75th percentile is 98%.

Having described the important features of the purchasing behaviour of consumers, I now move on to specify the demand model and the proposed estimation procedure.

3   Demand Model

This section introduces the utility function and the assumptions of the model. I study the demand for a temptation good in a setting similar to Hartmann (2006) where consumption creates a stock that diminishes over time. This creates in the consumer an incentive to variety-seek and thus intertemporal substitute consumption for the good. Unlike Hartmann (2006), I extend the analysis to address the differentiated nature of the good and examine not only substitution across time, but also substitution across products. In order to do so, I adapt Aguirregabiria (2002) and Hendel and Nevo (2006a) multi-stage budgeting approach.

3.1   The Setup

There are I consumers who are indexed by i. In each shopping trip t, consumer i chooses whether or not to purchase the good, and in case she decides to purchase, which product and size to buy. Let $j = 1, \ldots, J$ index the inside product alternatives to the consumer, with each product alternative being (possibly) offered in a variety of different sizes x. Multiple-item purchases are included by expanding the choice set to allow for bundles. If in a particular trip a consumer buys, for example, both 64 *oz.* of Häagen-Dazs Vanilla Ice Cream *and* 16 *oz.* of Häagen-Dazs Chocolate Ice Cream, the purchase size is given by $x = 80$ *oz.* and product

---

[5]The product availability proxy will obviously overestimate the induced product switching.

j denotes the bundle of the two products. The no purchase choice (outside alternative) is indexed by $j = 0$.

3.2 Consumer Flow Utility

The consumer flow utility is expressed in terms of the indirect utility from each of the available alternatives. I begin by specifying the indirect utility from not purchasing (the outside option). I follow Hartmann (2006) and relax the common assumption of additively-separable utility in consumption by considering a frequency of purchase model where past choices affect current utility. In particular, I assume the utility of the outside option to be a function of the depreciated stock of past consumption:

$$u_{i0t}\left(y_{it}, \varepsilon_{i0t}\right) = z_{it} + \varepsilon_{i0t}, \tag{1}$$

where z_{it} denotes the stock of past consumption of individual i at time t and ε_{i0t} is a random shock to consumer choice. The depreciated stock of past consumption will, in full generality, depend on both the time elapsed since the previous purchase and the size of past purchases. However, because the stock of past consumption is intangible and unobservable, I am required to infer it from (observed) past purchase choices. In order do so, I make the following assumptions.

Assumption 1 *Consumption takes place at the time of purchase.*

Assumption 1 is motivated by the temptation nature of the good. Since the data description analysis has shown that consumers do not anticipate purchases to hold inventories for future consumption, inferring that individuals do consume their purchased ice cream *before* their next purchase occasion is not unreasonable. However, the actual time of consumption is unobserved. Due to the temptation nature of the good, I assume that the time of consumption coincides with the time of purchase. At least for most people, in line with what Erdem *et al.* (2003) argue, ice creams are technologically, but not practically storable over more than a few days. As discussed above, it may seem inconsistent to assume that consumption has a lasting affect that induces intertemporal substitution in purchases while assuming that the good held in inventory has a temptation feature. These assumptions are, however, consistent with observed behaviour, since consumers seem to depreciate the costs of goods they have in inventory (see Gourville and Soman (1998) and Prelec and Loewenstein (1998)).

Assumption 2 *The stock of past consumption fully depreciates after a new consumption occasion.*

Assumption 2 implies that the stock of past consumption does not accumulate across multiple consumption occasions and that only the last consumption occasion is relevant. The motivation behind this assumption is twofold. First, it significantly reduces the state space since, instead of keeping track of all past choices, only the (observable) last consumption is relevant to the decision of consumers. An alternative approach would consist of constructing an accumulated stock index, a strategy that would also have the advantage of a simplified state space. However, it carries a disadvantage related to the second justification for Assumption 2. Under this assumption, the initial stock of past consumption is observable and does not need to be inferred, which would not be true if I allowed the stock to accumulate across multiple consumption occasions.

One concern with the simplification implied by Assumption 2 is that it comes at a cost: the stock of past consumption is measured with error. Although measurement error is a potentially troublesome problem, it may not be too problematic here. The error introduced will, at best, underestimate the incidence of intertemporal substitution, rather than falsely induce finding intertemporal substitution. In order to understand why this the case, note that because the estimated stock of past consumption will not exceed the *true* stock of past consumption, the utility of the outside option will be underestimated (in a setting where this alternative is the most common choice made by consumers). As a consequence, intertemporal substitution is underestimated.

Assumption 3 *The stock of past consumption is independent of the quantity purchased.*

Assumption 3 relates to a previous discussion since the descriptive analysis of the data has shown that the consumer interpurchase duration is not affected by the quantity purchased. Even though consumers respond to price promotions by increasing their purchased size, the effect on the duration to the next purchase is not significantly different from zero.

A consequence of Assumptions 1 to 3 is that only the time elapsed from the last purchase is relevant to infer the depreciated stock of past consumption. The utility of the outside option can then be specified as:

$$u_{i0t}\left(H_{it}, \varepsilon_{i0t}\right) = \varphi\left(H_{it}\right) + \varepsilon_{i0t}, \tag{2}$$

where H_{it} is the number of days since the last purchase occasion and $\varphi(H_{it})$ denotes the function that allows me to infer the (unobserved) stock of past consumption from (observed) past purchase choices.

I now move on to specify the indirect utility from choosing an inside alternative. I assume the utility to individual i in time period t from choosing a product j of size $x > 0$ that belongs to h_{it-1} is:

$$
\begin{aligned}
u_{ijxt}(p_{jxt}, a_{jxt}, h_{it-1}, \varepsilon_{ijxt}) &= \bar{u}_{ijxt}(p_{jxt}, a_{jxt}, h_{it-1}) + \varepsilon_{ijxt} \qquad (3) \\
&= \gamma_{ix} + \alpha_i p_{jxt} + \xi_{ijt} + \beta_i a_{jxt} + \lambda_i m_{jxt} + \eta_i y_{ijt-1} + \varepsilon_{ijxt},
\end{aligned}
$$

where h_{it-1} indicates the set of products purchased by consumer i in her previous purchase event, γ_{ix} denotes the (dis)utility from making a purchase of size x, which could be interpreted as a carrying cost associated with that particular purchase, p_{jxt} is the price of product j in size x, ξ_{ijt} is consumer i taste for for product j that could be a function of product characteristics (like, for example, size), a_{jxt} denotes a vector of indicator variables that control for other promotional activities, m_{jxt} is an indicator variable that takes the value 1 if product j denotes a multiple-item purchase, and ε_{ijxt} is a random shock to consumer choice. The variable y_{ijt-1} keeps track of the number of products that do not belong to the set h_{it-1} if consumer i purchases product j in purchase event t.

The term η_i accounts for state-dependence effects. A positive η_i implies that consumer i has a taste for variety-seeking, since switching to products different from those included in the h_{it-1} set increases the consumer's utility (see McAlister and Pessemier (1982)). The marketing literature provides several explanations for such variety seeking behaviour. Consumers may have an internal desire for change due to satiation or need for stimulation, or they may be balancing the different tastes within the household (see Kahn, 1995, for a comprehensive review of the variety seeking literature). A negative η_i, on the other hand, implies consumer i incurs in a switching cost, since switching to products that do not belong to the set h_{it-1} decreases the consumer's utility (see Pollack (1970) and Spinnewyn (1981)). Klemperer (1995) provides a number of possible reasons for switching costs. Consumers may have shopping search costs and, therefore, do not reoptimize the set of products purchased at every purchase occasion, or they may keep repurchasing the same product as part of a learning process. I do not attempt to distinguish here between these alternative explanations. Rather, I focus on whether state dependence in fact exists and can be identified from observed purchasing behaviour. This approach is similar to Osborne (2007).

3.3 Consumer Dynamic Optimization Problem

Consumers in each period decide if or not to purchase, and in case they opt to purchase, which product or products to choose. I make the following assumptions about how consumer expectations of the future affect current period decisions.

Assumption 4 *Consumers are forward-looking with regard to their purchase decisions, but myopic with respect to their product choices.*

The myopic assumption implies that consumers maximize their per-period expected utility when making their product choices and is motivated by solely pragmatism. A forward-looking consumer, who experiences state-dependence in her choices of product, considers the future consequences of those choices. The state space of the dynamic problem without this assumption would be extremely large, making the structural estimation computationally infeasible. The development of a framework that incorporates such forward-looking behaviour into a feasible computational estimation procedure is a very interesting potential area for future research. That said, the myopic assumption seems a reasonable assumption about consumer formation of expectations with regard to product choice. While some consumers may plan the whole sequence of product decisions accounting for the consequence of state-dependence in future periods, I tend to believe such forward-looking behaviour to be rare. I should note, however, that Assumption 4 does not imply that dynamics are absent from product choice. As I discuss below, current product choices impact the expected future flow utility of the different inside alternatives and, as a consequence, influence the purchase size decision. In other words, even though consumers are myopic with regard to product choice, their decisions have dynamic implications for current and future purchase size choices.

Assumption 4 implies a multi-stage budgeting approach to model the purchase and product decisions of consumers (see Aguirregabiria (2002) and Hendel and Nevo (2006a) for similar dynamic applications of Gorman (1971)'s approach). The consumer's expected discounted utility in purchase occasion t can therefore be represented as:

$$V\left(s_{it}\right) = \max_{\Pi_i} \sum_{\tau=t}^{\infty} \delta^{\tau-t} E\left[\sum_{x>0} \max_{\Phi_i} \sum_{j} d_{ix\tau} d_{ij/x\tau} u_{ijx\tau}\left(p_{jx\tau}, a_{jx\tau}, h_{i\tau-1}, \varepsilon_{ijx\tau}\right) \quad (4)\right.$$
$$\left. + d_{i0\tau} u_{i0\tau}\left(H_{i\tau}, \varepsilon_{i0\tau}\right) | s_{it}, \Pi_i, \Phi_i\right],$$

where s_{it} denotes the state at time t and $\delta > 0$ the discount factor. The state s_{it} in each period consists of the vector of current prices and promotional activities for all products and

sizes, the set of of products purchased by consumer i in her previous purchase event, the stock of past consumption as measured by the time since the last purchase, and the vector of random shocks to consumer choices, $s_{it} \equiv (p_t, a_t, h_{it-1}, H_{it}, \varepsilon_{it})$. For convenience, I define also the state space s_t^* that consists only of the vector of current prices and promotional activities for all products and sizes, and the set of of products purchased by consumer i in her previous purchase event, $s_{it}^* \equiv (p_t, a_t, h_{it-1})$.

Π_i and Φ_i denote a set of decision rules mapping states, s_{it}, to choices, d_{ixt} and $d_{ij/xt}$, respectively, where d_{ixt} is an indicator variable equal to 1 if the choice of consumer i is a purchase of size x (with $x = 0$ standing for no purchase) and $d_{ij/xt}$ is an indicator variable equal to 1 if consumer i chooses to buy product j when purchasing size x. The product of the two indicator variables, $d_{ijxt} = d_{ixt}d_{ij/xt}$, denotes the purchase of product j and size x. I assume that $\sum_{x,j} d_{ijxt} = 1$.

At every state, s_{it}, the consumer faces the same infinite-horizon maximization problem. The value function $V(s_{it})$ defined in equation (4) above is, therefore, the solution to the following Bellman's equation:

$$
\begin{aligned}
V(s_{it}) \;=\; \max_{d_{ixt}} \Bigg\{ & \sum_{x>0} \max_{d_{ij/xt}} \sum_j d_{ixt} d_{ij/xt} u_{ijxt}(p_{jxt}, a_{jxt}, h_{it-1}, \varepsilon_{ijxt}) \qquad (5) \\
& + d_{i0t} u_{i0t}(H_{it}, \varepsilon_{i0t}) + \delta E\left[V(s_{it+1}) \,|\, s_{it}, d_{ixt}, d_{ij/xt}\right] \Bigg\}.
\end{aligned}
$$

In order to complete the specification of the demand model, I make the following assumptions about the beliefs of consumers regarding the uncertain future prices (and promotional activities) and future utility random shocks.

Assumption 5 *Consumers have rational expectations.*

The assumption of rational expectations implies that consumers take all available information into account in forming expectations. Though expectations may turn out incorrect, they will not be systematically wrong. In particular, Assumption 5 implies that consumers know both the true transition probability of prices and promotional activities, and the true distribution of the utility random shocks.

Assumption 6 *The transition probability of prices and promotional activities are exogenous from the point of view of consumers. Furthermore, they follow a first-order Markov process.*

Assumption 6 is consistent with the view that retailers inter-temporal price discriminate by playing mixed strategies that are exogenous from the point of view of consumers (Conslik *et al.* (1984), Sobel (1984), Varian (1980), Pesendorfer (2002)). This assumption implies that, conditional on the control variables, price and promotional activities are independent of the unobserved random shocks, which might be unreasonable if consumers stockpile and inventories are not accounted for. If prices are persistent over time and consumers anticipate purchases in order to hold inventories for future consumption, then unobserved inventories will be correlated with current prices causing an endogeneity problem. Another concern with this assumption might be seasonality. If the likelihood of a temporary promotion is affected by seasonality and it is not accounted for into the transition probability, then unobserved random shocks will be correlated with current prices causing (again) an endogeneity problem. However, as discussed in the previous section, both issues are probably not a concern here.

The first-order Markov process assumption reduces the state space and, although probably inconsistent with equilibrium prices, it is not unreasonable with regard with observed consumers' memory and formation of expectations. The assumption can be relaxed to allow higher order processes, with an increase in the associated computational burden.

Assumption 7 ε_{ix_jt} *is independently and identically distributed extreme value type 1.*

Assumption 7 is motivated by pragmatism as it significantly reduces the computational burden. The main concern with this type of assumption might be to preclude correlation between products. This is not probably a concern here since the model accounts for product heterogeneity and product state-dependence. Incorporating correlation between the unobserved random shocks of different products can, in principle, be allowed, but at a significant increase in the computational costs of the estimation procedure.

4   Maximum Likelihood Estimation

This section presents the estimation details. I estimate the parameters of the model via maximum likelihood. The standard approach would begin by specifying the probability of observing consumer i's choices at time t, with this probability being given by the following likelihood function:

$$l_{it}\left(d_{it}|s_{it}\right) = \prod_{x,j} \left[\Pr\left(d_{ijxt} = 1\right)|s_{it}\right]^{d_{ijxt}}, \tag{6}$$

where $d_{it} \equiv \{d_{ijxt}\}$ denotes the vector of her choices. The likelihood of consumer i choices across all time periods would then be:

$$L_i\left(s_{i1}, \ldots, s_{iT}, d_{i1}, \ldots, d_{iT} | s_{i0}, d_{i0}\right) = \int \prod_t l_{it}\left(d_{it} | s_{it}\right) dF\left(s_{it} | s_{it-1}, d_{it-1}\right), \qquad (7)$$

where s_{i0} and d_{i0} denote the initial conditions, which are observed, and $F\left(s_{it} | s_{it-1}, d_{it-1}\right)$ is the transition probability.

The problem with the standard approach relates to the computation of $\Pr\left(d_{ijxt} = 1\right)$, the probability of observing consumer i purchasing product j and size x in period t, due to the dimensionality of the state space. In order to understand why this is the case, note that given the extreme value assumption on the unobserved utility random shocks (Assumption 7) this probability can be defined as:

$$\Pr\left(d_{ijxt} = 1 | s_{it}\right) = \frac{\exp\left\{\bar{u}_{ijxt}\left(p_{jxt}, a_{jxt}, h_{it-1}\right) + \delta E\left[V\left(s_{it+1}\right) | s_{it}, d_{ixt}, d_{ij/xt}\right]\right\}}{\sum_{y,k} \exp\left\{\bar{u}_{ikyt}\left(p_{kyt}, a_{kyt}, h_{it-1}\right) + \delta E\left[V\left(s_{it+1}\right) | s_{it}, d_{ixt}, d_{ij/xt}\right]\right\}}, \qquad (8)$$

where the summation is over all products from all sizes. The state space includes the vector of current prices and promotional activities for all products and sizes, the set of products purchased by consumer i in her previous purchase event, the stock of past consumption as measured by the time since the last purchase, and the vector of random shocks to consumer choices. Given the multitude of products and sizes available to consumers, the state space is extremely large for practical estimation of $\Pr\left(d_{ijxt} = 1 | s_{it}\right)$.

In order to simplify the estimation procedure, I propose a three-stage budgeting approach in the lines of Aguirregabiria (2002) and Hendel and Nevo (2006a).

Step 1 *Estimation of Product Preferences*

I begin by noting that $\Pr\left(d_{ijxt} = 1 | s_{it}\right)$ can, in full generality, be decomposed into the product of two components: the probability of choosing product j conditional on the size x purchased *and* the probability of choosing a purchase of size x:

$$\Pr\left(d_{ijxt} = 1 | s_{it}\right) = \Pr\left(d_{ij/xt} = 1 | s_{it}, d_{ixt}\right) \Pr\left(d_{ixt} = 1 | s_{it}\right). \qquad (9)$$

The myopia of consumers with regard to product choice (Assumption 4), implies that consumers maximize their per-period expected utility when making their product decisions. As a consequence, $\Pr\left(d_{ij/xt} = 1 | s_{it}, d_{ixt}\right)$ can be computed without solving the full dynamic problem. Furthermore, given the extreme value assumption on the utility shocks (Assumption

7), that probability can be defined as:

$$
\begin{aligned}
\Pr\left(d_{ij/xt} = 1 | s_{it}, d_{ixt}\right) &= \frac{\exp\left[\bar{u}_{ijxt}\left(p_{jxt}, a_{jxt}, h_{it-1}\right)\right]}{\sum_k \exp\left[\bar{u}_{ikxt}\left(p_{kxt}, a_{kxt}, h_{it-1}\right)\right]} \tag{10}\\
&= \frac{\exp\left(\gamma_i^m + \alpha_i p_{jxt} + \xi_{ijt} + \beta_i a_{jxt} + \lambda_i m_{jxt} + \eta_i y_{ijt-1}\right)}{\sum_k \exp\left(\gamma_i^m + \alpha_i p_{kxt} + \xi_{ikt} + \beta_i a_{kxt} + \lambda_i m_{kxt} + \eta_i y_{ikt-1}\right)}\\
&= \Pr\left(d_{ij/xt} = 1 | p_{xt}, a_{xt}, h_{it-1}\right),
\end{aligned}
$$

where the summation is now only over the products of size x.

The parameters in $\bar{u}_{ijxt}\left(p_{jxt}, a_{jxt}, h_{it-1}\right)$ - with the exception of γ_{ix} that cancels out - can therefore be recovered by maximizing the likelihood of consumer product choice conditional on the size purchased. Let $L_i^{step1}\left(h_{i1}, \ldots, h_{iT-1}, d_{ij/x1}, \ldots, d_{ij/xT} | h_{i0}\right)$ denote the likelihood of consumer i's conditional choices across all time periods:

$$
L_i^{step1}\left(h_{i1}, \ldots, h_{iT-1}, d_{ij/x1}, \ldots, d_{ij/xT} | h_{i0}\right) = \prod_t \prod_{j/x} \left[\Pr\left(d_{ij/xt} = 1 | p_{xt}, a_{xt}, h_{it-1}\right)\right]^{d_{ij/xt}}, \tag{11}
$$

where h_{i0} denotes the initial set of products purchased by consumer i, which is observed. Taking the product of this likelihood function across consumers yields the likelihood function to be maximized in step 1:

$$
L = \prod_i L_i^{step1}\left(h_{i1}, \ldots, h_{iT-1}, d_{ij/x1}, \ldots, d_{ij/xT} | h_{i0}\right). \tag{12}
$$

In making the utility of choosing a given product state-dependent from the set of products bought in the previous purchase occasion, h_{it-1}, I introduce an identification problem since unobserved consumer heterogeneity may confound the inference of true state-dependence effects. As Heckman (1981) points out, if households have different preferences "and if these differences are not properly controlled, previous experience may appear to be a determinant (...) of future experience solely because it is a proxy for temporally persistent unobservables that determine choices."

State-dependence is usually identified by testing the null hypothesis that the current choice, after accounting for consumer-level heterogeneity, is independent of the previous choice. One approach to introduce heterogeneity is to include *observed* consumer heterogeneity. This approach assumes the existence of a finite number of types or segments, with each type consisting of a set of consumers with identical overall choice preferences (Kamakura and Russell (1989)). As the number of types assumed increases, so will the degree of heterogeneity accounted for under this approach. Goldfarb (2006) presents the extreme

25

case where the number of types exactly coincides with the number of consumers. He makes use of a rich dataset containing nearly $1,000$ observations per household to estimate a fully flexible model of consumer preferences, by allowing for consumer-specific regressions.

Another approach is to introduce heterogeneity by considering consumer preferences to be realizations of *random variables.* These random variables are assumed in the literature to follow a multitude of distributional assumptions. For example, Chintagunta *et al.* (1991), Gonul and Srinivasan (1993) and Keane (1997) consider preferences to follow a continuous probability distribution, while Jain *et al.* (1994) consider a discrete probability distribution approximation. An intermediate assumption is presented by Dubé *et al.* (2006) that considers a flexible semi-parametric, but continuous model of consumer heterogeneity.

The estimation procedure in step 1 can not allow for random effects in the lines of the latter approach. If consumer preferences are assumed to be realizations of random variables that follow a probability distribution (either parametric or semi-parametric), then computing $\Pr\left(d_{ij/xt} = 1|p_{xt}, a_{xt}, h_{it-1}\right)$ requires integration over the assumed distribution. Although conditional on the type of consumer, this probability will still be independent of the dynamic purchase decision, computing this probability unconditional on the type of consumer requires integration over the distribution of types conditional on the size bought. And working out this distribution requires solving the dynamic problem.

Consumer-level heterogeneity can, however, be allowed in the lines of the former approach: either by using observable household demographics to segment consumers into types or, in the lines of Goldfarb (2006) and Hendel and Nevo (2006a), by considering household-level product and state-dependence fixed effects. One concern with the latter solution might be the dimensionality of the parameters to estimate. However, since the likelihood function in equation (12) is well behaved, the estimation of a considerable number of consumer-level fixed effects is feasible and involves very slight increases in computational costs. Furthermore, the consumer-product fixed effects need only to include those products that belong to each consumer shopping history. Consumer-level product preferences can not be estimated for products never purchased by the household. This reduces the number of consumer-product fixed effects substantially since each household typically purchases a relatively small number of products when compared with the full supermarket assortment. Another concern might be the standard incidental parameters problem. However, given the large number of consumer shopping trips in the typical scanner panel datasets, this issue is probably not a concern and therefore assuming T grows asymptotically is not unreasonable.

Step 2 *Estimation of the Inclusive Values Transition Process*

Having outlined the procedure to estimate the probability of choosing product j conditional on the size x purchased, I now move on to specify the two remaining steps required to estimate the probability of choosing a purchase of size x.

The consumer decision with regard to purchase size (whether and what quantity to purchase) is the solution to the dynamic problem characterized by Bellman's equation (5). However, instead of solving this problem, I follow Hendel and Nevo (2006a) and consider a simplification of the state space that makes use of the extreme value assumption on the utility shocks (Assumption 7). This simplification involves summarizing the consumer state space, $s_{it}^* \equiv (p_t, a_t, h_{it-1})$, into a single index per size, an index representing the utility expected by the consumer, *before* seeing the realization of the utility shocks, from all products of each size. Under Assumption 7, this expected utility is given by the inclusive value w_{ixt} (McFadden (1981a)):

$$w_{ixt} = \log \left[\sum_k \exp \left(\alpha_i p_{kxt} + \xi_{ikt} + \beta_i a_{kxt} + \lambda_i m_{kxt} + \eta_i y_{ikt-1} \right) \right], \tag{13}$$

which can be computed with the parameter estimates from step 1.

In order to show that the original dynamic problem can be written in terms of the simplified state space, I make the following additional assumption, where w_{it} denote the vector of inclusive values at time t:

Assumption 8 $F\left(w_{it} | s_{t-1}^* \right)$ can be summarized by $F\left(w_{it} | w_{it-1} \right)$.

Solving the consumer dynamic programming decision requires solving the associated Bellman's equation, which in turn involves working out the expectation of the value function. In order to compute such expectation, I need to specify the transition probabilities for the different state variables. Assumption 8 simplifies these processes. The motivation is twofold. First, the transition probabilities of prices and promotional activities from a multitude of different products of the same size are summarized into the transition probability of a single index. Second, it also simplifies the transition probabilities of product state-dependence. Although consumer product choice is, by Assumption 4, myopic (which means that current product choices do not impact future product choices), it does not mean that dynamics are absent. Current product choices impact the expected future flow utility of the different inside alternatives and, as a consequence, impact the expected future inclusive values that influence the purchase size decision. In other words, even though consumers are myopic with regard to product choice, their decisions have dynamic implications for current and future purchase size choices. Assumption 8 summarizes the transition probabilities regarding product choice

into the inclusive values processes. Because product-choices are consumer-specific, the inclusive values and their transition processes will necessarily be consumer-specific, requiring that the Bellman's equation is solved separately for each consumer

One concern with Assumption 8 might be that shopping trips involving different prices, promotional activities and/or previous period product choices can be reflected in a same inclusive value, which in turn yields the same future transition probabilities. This restriction can, to some extent, be relaxed, although at a substantial computational cost.

Step 3 *Estimation of the Intertemporal Effects of Consumption*

Step 3 addresses the computation of the probability of choosing a purchase of a given size x. Not by solving the dynamic problem characterized by Bellman's equation (5), but by solving a simplified problem, where the consumer state space, s_{it}^*, is summarized into the vector of single size indexes, w_{it}. In this simplified problem, the consumer observes only H_{it} and w_{it} and decides whether and how much to purchase.

I now move on to specify the details of this simpler problem. The utility of consumer i in time period t is given by:

$$
\begin{aligned}
u_{i0t}^{step3}\left(H_{it}, \varepsilon_{i0t}\right) &= \varphi\left(H_{it}\right) + \varepsilon_{i0t}, & \text{if } x = 0 \qquad (14)\\
u_{ixt}^{step3}\left(w_{ixt}, \varepsilon_{ixt}\right) &= \gamma_{ix} + w_{ixt} + \varepsilon_{ixt}, & \text{if } x > 0,
\end{aligned}
$$

where, as before, $x = 0$ stands for no purchase. The consumer is assumed to be forward-looking and, therefore, to maximize the expected discounted utility:

$$
\begin{aligned}
V_{step3}\left(H_{it}, w_{it}, \varepsilon_{it}\right) &= \max_{\Pi_i^{step3}} \sum_{\tau=t}^{\infty} \delta^{\tau-t} E\Bigg[\sum_{x>0} d_{ix\tau} u_{ix\tau}^{step3}\left(w_{ix\tau}, \varepsilon_{ix\tau}\right) \qquad (15)\\
&\quad + d_{i0\tau} u_{i0\tau}^{step3}\left(H_{i\tau}, \varepsilon_{i0\tau}\right)\big|H_{it}, w_{it}, \varepsilon_{it}, \Pi_i^{step3}\Bigg],
\end{aligned}
$$

where Π_i^{step3} denotes a set of decision rules mapping states to choices, d_{ixt}. The Bellman's equation associated with the consumer's simpler dynamic problem is given by:

$$
\begin{aligned}
V_{step3}\left(H_{it}, w_{it}, \varepsilon_{it}\right) &= \max_{d_{ixt}}\Bigg\{\sum_{x>0} d_{ixt} u_{ixt}^{step3}\left(w_{ixt}, \varepsilon_{ixt}\right) + d_{i0t} u_{i0t}^{step3}\left(H_{it}, \varepsilon_{i0t}\right) \qquad (16)\\
&\quad + \delta E\left[V_{step3}\left(H_{it+1}, w_{it+1}, \varepsilon_{it+1}\right)\big|H_{it}, w_{it}, \varepsilon_{it}, d_{ixt}\right]\Bigg\}.
\end{aligned}
$$

It remains to be shown that the probability of purchasing size x computed from the

simplified problem is equivalent to the one computed from the original problem. Establishing this equivalence involves two steps. In the first step, I show that the Bellman's equations associated with the original and simplified problems have the same solution. The second step involves actually showing the equivalence of the probability of purchasing size x from the two problems.

The two-step proof adapts the one presented in Hendel and Nevo (2006a) to this variety-seeking framework.

Proposition 1 *The Bellman's equations associated with the original and simplified problems have the same solution.*

Proof. I begin by addressing the original dynamic problem. The Bellman's equation associated with this problem is given in equation (5), reproduced here for convenience:

$$V(s_{it}) = \max_{d_{ixt}} \left\{ \sum_{x>0} \max_{d_{ij/xt}} \sum_j d_{ixt} d_{ij/xt} u_{ijxt} \left(p_{jxt}, a_{jxt}, h_{it-1}, \varepsilon_{ijxt} \right) \right.$$
$$\left. + d_{i0t} u_{i0t} \left(H_{it}, \varepsilon_{i0t} \right) + \delta E \left[V \left(s_{it+1} \right) | s_{it}, d_{ixt}, d_{ij/xt} \right] \right\}.$$

Given Assumption 7, the expected value of $V\left(s_{it+1}|s_{it}, d_{ixt}, d_{ij/xt}\right)$ will be a function of $H_{it}, s_{it}^*, d_{ixt}$ and $d_{ij/xt}$. Recall that s_{it}^* denotes the state space that consists only of the vector of current prices and promotional activities for all products and sizes, and the set of of products purchased by consumer i in her previous purchase event. Let $V^e\left(H_{it}, s_{it}^*\right)$ denote such function to simplify notation: $V^e\left(H_{it}, s_{it}^*\right) = E\left[V\left(s_{it+1}\right)|s_{it}, d_{ixt}, d_{ij/xt}\right]$.

Computing the expected value of $V\left(s_{it}\right)$ conditional on the information available at time $t-1$ yields:

$$V^e\left(H_{it-1}, s_{it-1}^*\right) = \int \left\{ \max_{d_{ixt}} \left\{ \sum_{x>0} \max_{d_{ij/xt}} \sum_j d_{ixt} d_{ij/xt} u_{ijxt} \left(p_{jxt}, a_{jxt}, h_{it-1}, \varepsilon_{ijxt} \right) \right\} \right.$$
$$\left. + d_{i0t} u_{i0t} \left(H_{it}, \varepsilon_{i0t} \right) + \delta V^e \left(H_{it}, s_{it}^* \right) \right\} dF\left(s_{it}|s_{it-1}, d_{it-1}\right).$$

The myopic assumption with regard to consumer product choice together with the extreme-value assumption allows this expected value to be re-written in terms of the inclusive values defined in equation (13):

$$V^e\left(H_{it-1}, s_{it-1}^*\right) = \int \log \left\{ \sum_{x>0} \exp\left(\gamma_{ix} + w_{ixt} + \delta V^e\left(H_{it}, s_{it}^*\right)\right) \right.$$
$$\left. + \exp\left[\varphi\left(H_{it}\right) + \delta V^e\left(H_{it}, s_{it}^*\right)\right] \right\} dF\left(s_{it}^*|s_{it-1}^*, d_{it-1}\right).$$

where the expression inside the integral represents integration over the vector of random utility shocks. From the analysis of the above equation, it is possible to conclude that $V^e\left(H_{it-1}, s_{it-1}^*\right)$ can, under Assumption 8, be iterated using the following Bellman's equation rewritten in terms of w_t instead of s_{it-1}^*:

$$
\begin{aligned}
V^e\left(H_{it-1}, w_{it-1}\right) \; = \; & \int \log \left\{\sum_{x>0} \exp\left[\gamma_{ix} + w_{ixt} + \delta V^e\left(H_{it}, w_{it}\right)\right]\right. \\
& \left. + \exp\left[\varphi\left(H_{it}\right) + \delta V^e\left(H_{it}, w_{it}\right)\right]\right\} dF\left(w_{it}|w_{it-1}, d_{it-1}\right).
\end{aligned}
$$

I now address the simplified problem. The Bellman's equation associated with this problem is given in equation (16). After substituting for $u_{i0\tau}^{step3}\left(H_{it}, \varepsilon_{i0t}\right)$ and $u_{ixt}^{step3}\left(w_{ixt}, \varepsilon_{ixt}\right)$ yields:

$$
\begin{aligned}
V_{step3}\left(H_{it}, w_{it}, \varepsilon_{it}\right) \; = \; & \max_{d_{ixt}} \left\{\sum_{x>0} d_{ixt}\left(\gamma_{ix} + w_{ixt} + \varepsilon_{ixt}\right) + d_{i0t}\left[\varphi\left(H_{it}\right) + \varepsilon_{0t}\right]\right. \\
& \left. + \delta E\left[V_{step3}\left(H_{it+1}, w_{it+1}, \varepsilon_{it+1}\right)|H_{it}, w_{it}, \varepsilon_{it}, d_{ixt}\right]\right\}.
\end{aligned}
$$

Taking expectations given the information available at time $t-1$ and integrating out the utility random shocks making use of the extreme-value assumption (Assumption 7) allows me to write the expected value, $V_{step3}^e\left(H_{it-1}, w_{it-1}\right)$, as:

$$
\begin{aligned}
V_{step3}^e\left(H_{it-1}, w_{it-1}\right) \; = \; & \int \log \left\{\sum_{x>0} \exp\left[\gamma_{ix} + w_{ixt} + \delta V^e\left(H_{it+1}, w_{it+1}\right)\right]\right. \\
& \left. + \exp\left[\varphi\left(H_{it}\right) + \delta V^e\left(H_{it}, w_{it}\right)\right]\right\} dF\left(w_{it}|w_{it-1}, d_{it-1}\right).
\end{aligned}
$$

Thus, as the proposition claims, the solution to the Bellman's equations associated with the original and simplified problems is the same. ∎

I now address the second step of the proof, showing the equivalence between the probability of purchasing size x computed from the original problem to the one computed from the simplified problem.

Proposition 2 $\Pr\left(d_{ixt} = 1|H_{it}, s_{it}^*\right) = \Pr\left(d_{ixt} = 1|H_{it}, w_{it}\right).$

Proof. The probability of purchasing size x computed from the simplified problem is given by:

$$
\Pr\left(d_{ixt} = 1|H_{it}, w_{it}\right) = \frac{\exp\left[\gamma_{ix} + w_{ixt} + \delta V^e\left(H_{it+1}, w_{it+1}\right)\right]}{M_{i0t} + \sum_{y>0} \exp\left[\gamma_{iy} + w_{iyt} + \delta V^e\left(H_{it+1}, w_{it+1}\right)\right]},
$$

where for notational simplicity $M_{i0t} = \exp\left[\varphi\left(H_{it}\right) + \delta V^e\left(H_{it}, w_{it}\right)\right]$. If, on the other hand,

this probability is computed from the original problem, it is given by:

$$\Pr\left(d_{ixt}=1|H_{it},s_{it}^*\right)=\frac{\sum_j \exp\left\{\bar{u}_{ijxt}\left(s_{it}^*\right)+\delta V^e\left(H_{it+1},s_{it+1}^*\right)\right\}}{M_{i0t}^*+\sum_{y,k}\exp\left\{\bar{u}_{ikyt}\left(s_{it}^*\right)+\delta V^e\left(H_{it+1},s_{it+1}^*\right)\right\}},$$

where $M_{i0t}^*=\exp\left[\varphi\left(H_{it}\right)+\delta V^e\left(H_{it},s_{it}^*\right)\right]$. The summation in the numerator is over all products of size x, and the summation in the denominator is over all products of all sizes. As discussed in Proposition 1, $V^e\left(H_{it+1},s_{it+1}^*\right)$ can, under Assumption 8, be re-written in terms of w_{it+1} instead of s_{it+1}^*. This implies that the expected value function depends on the purchase size chosen, but not on the particular product choice. Furthermore, $M_{i0t}=M_{i0t}^*$. As a consequence, the above probability can be decomposed and simplified as follows:

$$\Pr\left(d_{ixt}=1|H_{it},s_{it}^*\right)=$$

$$=\frac{M_{ixt}\sum_j\exp\left(\alpha_i p_{jxt}+\xi_{ijt}+\beta_i a_{jxt}+\lambda_i m_{jxt}+\eta_i y_{ijt-1}\right)}{M_{i0t}+\sum_{y>0}M_{iyt}\sum_k\exp\left(\alpha_i p_{kyt}+\xi_{ikt}+\beta_i a_{kyt}+\lambda_i m_{kyt}+\eta_i y_{ikt-1}^e\right)}$$

$$=\frac{M_{ixt}\exp\left\{\log\left[\sum_j\exp\left(\alpha_i p_{jxt}+\xi_{ijt}+\beta_i a_{jxt}+\lambda_i m_{jxt}+\eta_i y_{ijt-1}\right)\right]\right\}}{M_{i0t}+\sum_{y>0}M_{iyt}\exp\left\{\log\left[\sum_k\exp\left(\alpha_i p_{kyt}+\xi_{ikt}+\beta_i a_{kyt}+\lambda_i m_{kyt}+\eta_i y_{ikt-1}^e\right)\right]\right\}}$$

$$=\frac{\exp\left(w_{ixt}+M_{ixt}\right)}{M_{i0t}+\sum_{y>0}\exp\left(w_{iyt}+M_{iyt}\right)}$$

$$=\frac{\exp\left[\gamma_{ix}+w_{ixt}+\delta V^e\left(H_{it+1},w_{it+1}\right)\right]}{M_{i0t}+\sum_{y>0}\exp\left[\gamma_{iy}+w_{iyt}+\delta V^e\left(H_{it+1},w_{it+1}\right)\right]}$$

$$=\Pr\left(d_{ixt}=1|H_{it},w_{it}\right),$$

where $M_x=\exp\left[\gamma_{ix}+\delta V^e\left(H_{it+1},w_{it+1}\right)\right]$.

Thus, as the proposition claims, the probabilities computed from the original and simplified problems are equivalent. ■

Having established the equivalence of the probability of a purchase of size x between the two problems, I move on to specify the estimation procedure. I estimate the remaining parameters by maximizing the likelihood of consumer purchase choices. Let the likelihood of consumer i's purchase choices across all time periods be denoted by:

$$L_i^{step3}\left(H_{i1},\ldots,H_{iT},d_{ix1},\ldots,d_{ixT}|H_{i0}\right)=\prod_t\prod_x\left[\Pr\left(d_{ixt}=1|H_{it},w_{it}\right)\right]^{d_{ixt}}dF\left(w_{it}|w_{it-1},d_{it-1}\right),$$

$$(17)$$

where H_{i0} denotes the initial stock of past consumption as measured by the time since the initial purchase.

In making the utility of purchasing a given size state-dependent from the duration since

the last purchase, I introduce an identification problem similar to the one discussed previously for product switching: unobserved consumer heterogeneity may confound the inference of true state-dependence effects. The identification problem arises from the fact the interpurchase duration may be long either due to a low taste for the good or a strong state-dependence effect. In order to control for unobserved heterogeneity, I assume preferences vary across consumers using a random effects specification:

$$\lambda_i = \lambda + \Gamma v_i, \tag{18}$$

where λ_i denotes the vector of the remaining parameters to be estimated (and includes the duration dependence and the size-specific (dis)utilities parameters) and v_i is a independently and identically distributed standard normal. The vector λ denotes the mean values of the different coefficients, while Γ denotes the Cholesky decomposition of the variance-covariance matrix, Σ, which for computational simplicity is assumed diagonal.

I should note that although one may argue that consumer-specific inclusive-values already control for unobserved consumer heterogeneity in their taste for the good, a random effects specification for the size-specific (dis)utilities is required in practice. The justification relates to an unfortunate property of the conditional logit model used in step 1. The model is not able to estimate an intercept since it plays no role in determining the product-choice probability conditional on the size purchased. As a consequence, in order to estimate the consumer-level product preferences in step 1, a normalization is required for each consumer and size (since including dummy variables for all products in the conditional choice would amount to estimate a size-specific intercept). The random effects specification for the size-specific (dis)utilities in step 3 is instrumental in making the inclusive values across sizes and consumers, each estimated using a different normalization, comparable.

With the introduction of the random effects, the likelihood of consumer i's purchase choices across all time periods is now given by:

$$L_i^{step3}(H_i, d_{ix}|H_{i0}) = \int \prod_t \prod_x \left[\Pr(d_{ixt} = 1|H_{it}, w_{it}, v_i) \right]^{d_{ixt}} dF(v_i) \, dF(w_{it}|w_{it-1}, d_{it-1}). \tag{19}$$

I follow Pakes (1986), Pakes and Pollard (1989), and McFadden (1989) and draw ns pseudo-random consumers to approximate the integral using a (smooth) simulator estimator:

$$L_i^{step3}(H_i, d_{ix}|H_{i0}) = \int \prod_t \prod_x \left[\frac{1}{ns} \sum_s \Pr(d_{ixt} = 1|H_{it}, w_{it}, v_s) \right]^{d_{ixt}} dF(v_i) \, dF(w_{it}|w_{it-1}, d_{it-1}). \tag{20}$$

Taking the product of this likelihood function across consumers yields the likelihood function to be maximized in step 3:

$$L = \prod_i L_i^{step3} \left(H_i, d_{ix} | H_{i0} \right). \tag{21}$$

4.1 Bellman's Equation Solution

The structural estimation is based on Rust (1987)'s algorithm that nests the solution of the consumer's dynamic programming within the estimation parameter search. In this section, I address the computational details of the strategy used to solve the functional equation (16) associated with the consumer's simpler dynamic problem. One strategy to solve dynamic programming problems is by *discrete approximation*. In this type of approach, the value function is solved for numerically by discretizing continuous state spaces into a finite number of n grid points. However, in high-dimensional problems, discretization results in a curse of dimensionality, since n increases exponentially fast in the dimension of the state space. Another approach is to solve dynamic programming problems by *parametric approximation*, where the value function is approximated by a smooth parametric function with k unknown parameters. The latter approach is superior to the former whenever the number of parameters k required to obtain a good global approximation (according to some metric) under parametric approximation is smaller than the value n of grid points required to obtain a comparable fit by discrete approximation.

I follow Hendel and Nevo (2006a) and solve the functional equation (16) by using value function parametric approximation with policy function iteration in the lines of Benitez-Silva *et al.* (2000). Policy function iteration consists of an alternating sequence of *policy improvement* and *policy valuation* steps:

Policy Valuation

The policy valuation step computes the value function, $V_{step3} \left(H_{it}, w_{it}, \varepsilon_{it} \right)$, for a given initial guess of the consumer decision, d_{it}. Under a parametric approximation approach, the value function is approximated by a linear combination of k basis functions $(\rho_1, \ldots, \rho_k)$:

$$V_{step3} \left(H_{it}, w_{it}, \varepsilon_{it} \right) \simeq \sum_k \theta_k \rho_k \left(H_{it}, w_{it}, \varepsilon_{it} \right). \tag{22}$$

Substituting $V_{step3} \left(H_{it}, w_{it}, \varepsilon_{it} \right)$ in functional equation (16) by the polynomial approximation

yields a linear equation of k unknown parameters θ:

$$\sum_k \theta_k \rho_k \left(H_{it}, w_{it}, \varepsilon_{it}\right) = \sum_x d_{ixt} u_{ixt}^{step3} \left(w_{ixt}, \varepsilon_{ixt}\right) + d_{i0t} u_{i0t}^{step3} \left(H_{it}, \varepsilon_{i0t}\right) \tag{23}$$
$$+ \delta \int \sum_k \theta_k \rho_k \left(H_{it+1}, w_{it+1}, \varepsilon_{it+1}\right) dF\left(\varepsilon_{it+1}\right) dF\left(w_{it+1}|w_{it}, d_{it}\right).$$

This can be solved by ordinary least squares when evaluated at a finite set of $m \geq k$ sample points in the state space $(H_{it}, w_{it}, \varepsilon_{it})$. In order to understand why this is the case, define the $(m \times k)$ matrices r and Er, as well as the $(m \times 1)$ vector u with the following elements:

$$\begin{aligned} r_{mk} &= \rho_k \left(H_m, w_m, \varepsilon_m\right) \tag{24}\\ Er_{mk} &= \int \rho_k \left(H_{m+1}, w_{m+1}, \varepsilon_{m+1}\right) dF\left(\varepsilon_{m+1}\right) dF\left(w_{m+1}|w_m, d_m\right) \\ u_m &= \sum_x d_{xm} u_{xm}^{step3} \left(w_{xm}, \varepsilon_{xm}\right) + d_{0m} u_{0m}^{step3} \left(H_m, \varepsilon_{0m}\right). \end{aligned}$$

Equation (23) can then be re-written as a system of linear equations: $u = X\theta$, where $X = (r - \delta Er)$. The solution to this system of equations, which is given by $\hat{\theta} = (X'X)^{-1} X'u$, can then be used to evaluate the approximated value function.

Policy Improvement

The policy improvement step updates the guess of the consumer decision, d_{it}, using the value function approximation from the policy valuation step. The updated consumer decision (purchase size) can be performed analytically by maximizing the sum, evaluated at the same m sample points, of current utility and the expected discounted value of the value function:

$$d_{xm} = \arg\max \left\{ \sum_{x>0} d_{xm} u_{xm}^{step3} \left(w_{xm}, \varepsilon_{xm}\right) + d_{0m} u_{0m}^{step3} \left(H_m, \varepsilon_{0m}\right) \right. \tag{25}$$
$$\left. + \delta \int \sum_k \hat{\theta}_k \rho_k \left(H_{m+1}, w_{m+1}, \varepsilon_{m+1}\right) dF\left(\varepsilon_{m+1}\right) dF\left(w_{m+1}|w_m, d_m\right) \right\}.$$

The two steps are then iterated until convergence of the parameters of the value function approximation. The consumer decision that it converges to, and the corresponding value functions are approximated solutions to the Bellman's equation. See Puterman and Shin (1978) for sufficient conditions for policy iteration to converge in continuous state spaces.

4.2 Identification

In this section, I provide an informal discussion of identification. I begin by addressing the identification of step 1 parameters. The identification of the non-dynamic product preference parameters is standard, with the coefficients being identified through the effect of *current* period's variation in those exogenous variables on *current* period's probability of choosing a given product. Temporary price and non-price promotions provide variation to identify sensitiveness to price and other promotional activities. The (dis)utility from multiple-item purchasing is identified by the share of multiple-item purchases across trips. Consumer-level product effects are identified from variations in consumer shares across products.

Product choice state-dependence is identified, as argued by Chamberlain (1985), through the effect of *previous* period's variation in exogenous variables on *current* period's probability of choosing a given product. If a temporary promotion for product j at time $t-1$ decreases the probability of a given consumer choosing product j at time t, then the consumer may be a variety-seeker. If, on the other hand, such promotion increases that probability, the consumer may incur in switching costs. Given a long enough consumer-level price (and other promotional activities) time series, variation in previous period's promotions identifies product-choice state-dependence.

Step 2 parameters are identified through the effect of previous period's variation in each consumer inclusive values on her current period's inclusive values.

I now move on to address the identification of step 3 parameters. The purchase size coefficients help fit the infrequent incidence of purchase across observed trips and are naturally identified from each consumer's propensity to purchase the different sizes. The intertemporal effect of purchasing on the utility of the outside alternative is identified by each consumer's interpurchase duration in days. Because in a discrete choice demand model only the relative utilities are identifiable, an identifying normalization is required. I normalize the utility of the outside option to zero when $H_{it} \leq 1$. Finally, I note that in this frequency of purchase model, the discount factor is not identifiable. I assume it to equal 0.995.

5 Empirical Analysis

5.1 Step 1: Estimation of Product Preferences

Step 1 estimates product preferences by maximizing the likelihood of observing the sequence of household product choices, conditional on the size purchased. Therefore, the choice set includes only products of the same size as the actual purchase. Table 7 presents the results of this analysis, with the different columns reporting distinct specifications that vary on the covariates included. Specification (1) includes as explanatory variables price and a multiple-item purchase dummy variable. The price coefficient is of the expected sign and statistically significant suggesting that the average household is price sensitive. The multiple-item coefficient is not statistically different from zero which seems to indicate that consumers product choice pattern when purchasing a single-item does not significantly differ from when they purchase multiple-items. Specification (2) controls for promotional activities by including feature and display dummy variables as additional covariates. The coefficients on these controls are positive and statistically significant suggesting that consumers do respond to promotional activities. However, the comparison of the price coefficient in the two specifications is suggestive of an endogeneity issue. Prices are negatively correlated with promotional activities since promoted products sell at lower prices and, as a consequence, not including these controls will overestimate consumer price sensitiveness. In specification (3), I include product dummy variables in order to control for market-level unobserved product characteristics. The product dummy variables are interacted with size so that the preference for each specific product is proportional to the package size purchased. The effect of including these controls on the price coefficient is again suggestive of an endogeneity issue. Products with higher unobserved characteristics sell at higher prices inducing a positive correlation that will underestimate consumer price sensitiveness if not accounted for.

Specification (4) addresses the question of whether household have switching costs or are variety-seeking by including as covariate the number of products that, in each alternative choice, do not belong to the set of products bought in her previous purchase event. The coefficient is negative and statistically significant suggesting that the average consumer incurs in a cost when switching products in successive purchase occasions. The problem with this specification is that unobserved household heterogeneity will confound the inference of true state-dependence effects. The identification problem arises because a consumer may repeatedly purchase a particular product either because of a strong unobserved, idiosyncratic preference for it or because she dislikes switching. In order to identify true state dependence,

TABLE 7
Step 1: Estimation of Product Preferences*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Price	-0.67	-0.56	-0.67	-0.66	-0.54	-0.54	-0.54
	(0.06)	(0.07)	(0.08)	(0.08)	(0.12)	(0.12)	(0.12)
Price × Single					-0.06	-0.06	-0.07
					(0.16)	(0.16)	(0.16)
Price × Children					0.01	0.01	0.01
					(0.16)	(0.16)	(0.17)
Feature		0.54	0.52	0.54	0.78	0.78	0.79
		(0.13)	(0.13)	(0.12)	(0.15)	(0.15)	(0.16)
Display		0.61	0.63	0.63	0.72	0.73	0.76
		(0.15)	(0.15)	(0.15)	(0.16)	(0.16)	(0.17)
Multiple-Item	-0.47	-0.73	-1.17	-1.06	-3.51	-3.50	-3.94
	(0.40)	(0.45)	(0.57)	(0.56)	(1.00)	(0.99)	(1.18)
Multiple-Item × Single					2.71	2.74	2.88
					(1.30)	(1.30)	(1.55)
Multiple-Item × Children					2.65	2.62	3.01
					(1.20)	(1.21)	(1.39)
Product State-Dependence				-0.39	0.22	0.21	
				(0.08)	(0.09)	(0.11)	
Product State-Dependence × Single						-0.07	
						(0.21)	
Product State-Dependence × Children						0.16	
						(0.18)	
HH Product State-Dependence Dummy Variables							yes
Product Dummy Variables			yes	yes			
HH Product Dummy Variables					yes	yes	yes
Pseudo R2	0.04	0.05	0.15	0.15	0.24	0.24	0.26

* An observation is a purchase instance by a household. Standard errors clustered by households in parentheses.

I control for household heterogeneity in specification (5). I introduce heterogeneity in two ways. First, I interact price and multiple-item covariates with two observable household demographics: a dummy variable that takes the value 1 if the household is of a single person, and another if children under the age of 18 are present in the household.[6] Second, I introduce household-level product dummy variables. I assume, as before, preference for each specific product to be proportional to the package size purchased. No dimensionality problem arises with this introduction because (i) I only consider the products that belong to each household shopping history (I can not expect to estimate household product preferences for products never purchased by the household), (ii) each household buys a relatively small number of products, and (iii) has a relatively long time sequence of purchases. Most demographic interactions are statistically insignificant suggesting observable characteristics are not important in explaining price sensitiveness or taste for multiple-item purchases.[7] In contrast, most household product dummy variables are statistically significant. The introduction of household-level product heterogeneity generates substantial changes in the state-dependence coefficient. Households are now estimated to have an average positive taste for variety-seeking. These results seem to indicate that controlling for household heterogeneity matters. Specification (6) and (7) introduce heterogeneity in the variety-seeking/switching cost coefficient. In specification (6), I interact it with observable household demographics, with the interactions being statistically insignificant, while in specification (7) I allow for full household heterogeneity in the coefficient by interacting it with household-level dummy variables. Except for three households, all coefficients are statistically significant. Figure 4 plots the coefficient frequency distribution. Most of the households have a taste for variety, but the magnitude is relatively small. Approximately 18% of the consumers actually incurs in a cost when switching products in successive purchase occasions, while approximately 21% are heavily variety-seeking.

5.2 Step 2: Estimation of the Inclusive Values Transition Process

Step 2 estimates the transition process for the inclusive values, which were computed for the purchase sizes observed in the data (16 $oz.$, 32 $oz.$, 64 $oz.$, 80 $oz.$, 128 $oz.$ and 160 $oz.$) using step 1's estimates from specification (7) above. I follow Hendel and Nevo (2006a) and assume the following first-order Markov process for the transition probability of the inclusive

---

[6]I also estimated several specifications that included interactions with household income. Since the results were never significant, I do not consider them here.

[7]The only exception being, surprisingly, that one person households have a higher taste for multiple-item purchases.

FIGURE 4
Frequency Distribution for Product State-Dependence Coefficient
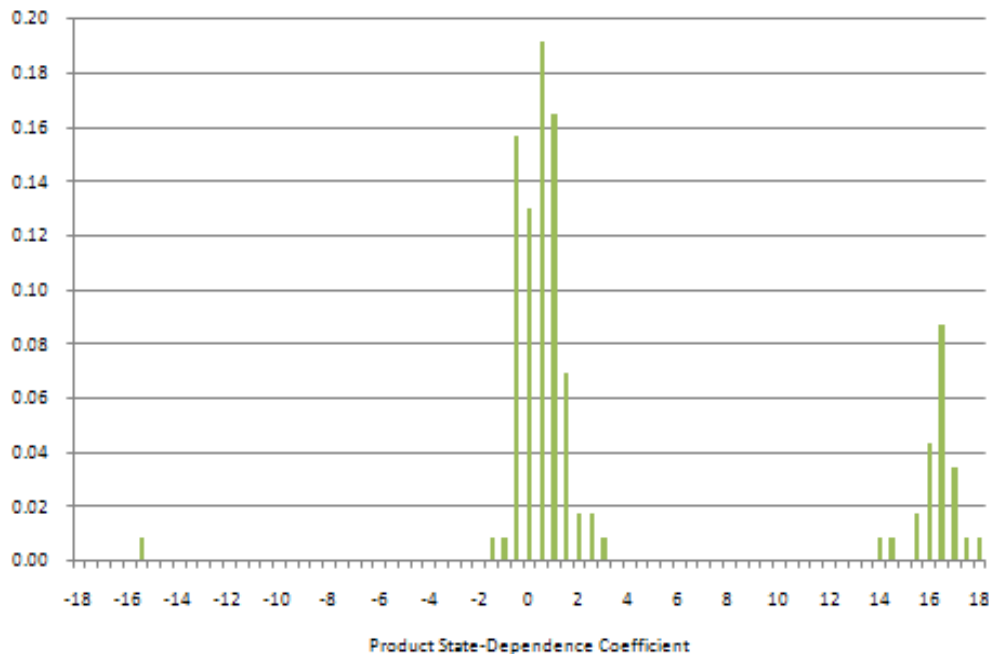
values:

$$w_{ixt} = \theta_{ix0} + \sum_{s \in \Im} \theta_{ixs} w_{ist-1} + \varsigma_{ixt}, \tag{26}$$

where the summation is over the set of package sizes $\Im = (16\ oz., 32\ oz., 64\ oz., 160\ oz.)$ and ς_{ixt} is distributed normal with mean zero and standard deviation σ_{ix}. Multicollinearity precludes the transition process of being defined over the set of all possible purchase sizes (since consumers that are observed purchasing 80 $oz.$ and 128 $oz.$, do so by buying multiple items: 80 $oz.$ = 16 $oz.$ + 64 $oz.$ while 128 $oz.$ = 64 $oz.$ + 64 $oz.$). Finally, the transition process parameters are index by i because the inclusive values are consumer-specific.

The assumption that the inclusive values are normally distributed may seem somewhat problematic given the evolution of the state variables they summarize. In order to test this assumption, I performed the Shapiro-Wilk W test for normality on the different consumer-level inclusive values. In the untabulated tests, the null hypothesis that the inclusive values are distributed normal is accepted only for a small fraction of consumers: 16 $oz.$ (3%), 32 $oz.$ (3%), 64 $oz.$ (17%), 80 $oz.$ (3%), 128 $oz.$ (13%) and 160 $oz.$ (1%). This assumption can be relaxed, although with a substantial increase in the computational burden.

Table 8 reports the estimated transition probabilities. Table 8, Panel A presents the point estimates (and associated standard errors) under the constraint that all consumers

TABLE 8

Table 8
Step 2: Estimation of the Inclusive Values Transition Process*

| | Panel A | | | | Panel B | | | |
| | Same Process for All Consumers | | | | Consumer-Level Process | | | |
	ω_{16t-1}	ω_{32t-1}	ω_{64t-1}	ω_{160t-1}	ω_{16t-1}	ω_{32t-1}	ω_{64t-1}	ω_{160t-1}
ω_{16t}	0.97	0.14	-0.05	0.54	0.21	-0.19	-0.20	0.36
	(0.09)	(0.06)	(0.03)	(0.08)	(0.34)	(2.61)	(6.51)	(3.20)
ω_{32t}	0.02	1.08	-0.09	0.15	0.31	0.08	0.08	0.11
	(0.11)	(0.07)	(0.04)	(0.10)	(1.06)	(0.21)	(5.30)	(4.25)
ω_{64}	0.60	-0.27	0.96	0.00	-0.02	-0.01	0.25	0.01
	(0.08)	(0.05)	(0.03)	(0.07)	(0.43)	(0.11)	(0.26)	(0.06)
ω_{80t}	2.21	-0.46	0.48	0.27	0.17	-0.19	-0.05	0.39
	(0.12)	(0.07)	(0.04)	(0.10)	(0.56)	(2.61)	(6.56)	(3.41)
ω_{128t}	0.63	-0.28	1.76	0.16	-0.03	-0.02	0.45	0.02
	(0.08)	(0.05)	(0.03)	(0.07)	(0.43)	(0.20)	(0.47)	(0.13)
ω_{160t}	0.33	-0.10	-0.03	1.24	-0.00	0.00	0.12	0.11
	(0.07)	(0.04)	(0.02)	(0.06)	(0.00)	(0.01)	(5.98)	(0.25)

* An observation is a shppping trip instance by a household. Also included are a constant and size indicator variables to control for unavailability of a package size at a given shopping trip. Panel A displays point estimates and standard errors in parentheses. Panel B displays the mean and standard deviation across the different consumer estimates.

face the same transition probabilities. The results suggest that the lagged inclusive value of own size (or of the two own sizes for those cases that involve multiple-item purchases) is the most important in predicting its future variation. In Table 8, Panel B the estimated transition probabilities are consumer-specific, with the results displaying the mean and standard deviation across the different consumer-level estimates. There is evidence of substantial heterogeneity across consumers, as suggested by the large standard deviations, which supports the option for the individual-level transition processes.

5.3 Step 3: Estimation of the Intertemporal Effects of Consumption

Step 3 maximizes the likelihood of observing the sequence of consumer purchase choices after solving the consumer-specific Bellman's equations associated with the simplified dynamic programming problem. Even though I solved the Bellman's equation separately for each consumer, the random effects specification for the parameters allowed me to pool the likelihoods across consumers. As discussed previously, I approximated the value function by a linear combination of k basis functions $(\rho_1, \ldots, \rho_k)$, with the approximation basis used being a polynomial in the natural logarithm of the duration in days since the consumer's

last purchase and in the levels of the remaining state variables.

In order to estimate the model, I have to specify a functional form for $\varphi(H_{it})$. I assume the following:

$$\varphi(H_{it}) = \kappa_0 \ln(H_{it}) + \kappa_1 H7_{it},$$

where $H7_{it}$ denotes an indicator variable that takes the value 1 if the shopping trip at time t corresponds to the consumer seven days cycle as suggested by the purchase hazard rate.

Table 9 reports the results for different specifications of step 3. In specification (1) I do not allow for heterogeneity or forward-looking behaviour. The state-dependence results suggest that the utility of the outside option decreases with the duration since the last purchase, which supports the intertemporal substitution argument. This result should not come as a surprise, despite the opposite suggestion from the raw data (recall the slight downward trend of the hazard rate), because the consumer-specific inclusive values do control in some extent for unobserved heterogeneity. The coefficient on the indicator variable $H7_{it}$ is negative, which suggests that once every 7 days, the value of the outside option decreases. I interpret this result as illustrating potentially reduced transaction costs of consumers purchasing in their main shopping trip. The estimates for the size-specific effects are statistically significant at standard significance levels. Econometrically, they help fit each size frequency of purchase. However, as discussed previously, the magnitude and ordering of these estimates can not be directly interpreted as they capture the different normalizations required for step 1 estimation.

Specification (2) introduces dynamic considerations into the consumers decisions, with this introduction substantively reducing state-dependence. The reason is that the static specification omits price expectations from the consumers purchase decisions. When facing a price promotion, the typical consumer expectation is that the price will go up in the future. This induces her, as I discussed in the descriptive analysis section, to typically take advantage of the price promotion by anticipating purchases. The static specification, by omitting price expectations, bias the results since it interprets this shorter interpurchase durations as stronger state-dependence. The addition of the forward-looking behaviour also impacts the coefficient on the indicator variable, which becomes (significantly) positive. This result is unexpected and hard to interpret. The estimates for the size-specific effects maintain the same ordering and magnitude.

Finally, specification (3) estimates the version of the model described in the previous sections that allows for both heterogeneity (via random coefficients) and forward-looking behaviour. The results for the mean estimates do not change substantively. However, the

41

TABLE 9
Step 3: Estimation of the Intertemporal Effects of Consumption*

	Static	Dynamic		
	Standard Logit Estimate (1)	Standard Logit Estimate (2)	Random Coefficients Mean Estimate	Standard Deviation (3)
No Purchase (Outside Alternative)				
κ_0	-2.42	-1.63	-1.83	4.57
	(0.00)	(0.00)	(0.02)	(0.03)
κ_1	-2.37	0.43	-0.22	0.26
	(0.02)	(0.02)	(0.19)	(0.15)
Purchase Size				
16 *oz.*	-36.87	-36.84	-38.14	8.76
	(0.02)	(0.02)	(0.34)	(0.18)
32 *oz.*	-40.52	-40.27	-40.56	6.08
	(0.02)	(0.02)	(0.22)	(0.12)
64 *oz.*	-25.56	-25.25	-35.88	13.95
	(0.01)	(0.01)	(0.29)	(0.21)
80 *oz.*	-43.39	-41.57	-41.61	4.29
	(0.20)	(0.04)	(0.27)	(0.43)
128 *oz.*	-49.43	-49.60	-42.89	6.09
	(0.02)	(0.02)	(0.13)	(0.08)
160 *oz.*	-39.42	-39.48	-39.52	5.42
	(0.05)	(0.05)	(0.26)	(0.40)
Log-likelihood	-45,501	-36,294	-14,244	

* An observation is a shopping trip instance by a household. Standard errors in parentheses.

fit of the model, as measured by the log-likelihood, increases significantly. This illustrates the importance of accounting for heterogeneity, not only to control for different degrees of state-dependence (the results suggest substantial heterogeneity at this level), but also to control for the different *consumer*-size normalizations required for step 1 estimation.

5.4 Simulation Algorithm and Goodness of Fit

In this section, I present an algorithm to simulate the several dimensions of the model and examine its fit. The need to specify a simulation algorithm arises because the estimation algorithm decomposes the likelihood of the consumer choices into two components: the choice of whether to purchase (and what size) and the decision of which product to buy if purchasing a positive amount. However, the choice of product influences the purchase

decision and vice-versa. In order to address this issue, I propose the following consumer-level algorithm:

1. Solve the simplified dynamic programming problem and simulate the sequence of purchase decisions (whether to purchase and not, and what size if they decide to purchase) conditional on the *observed* inclusive values.

2. For each shopping trip that the consumer decides to purchase a positive amount, simulate her product choice(s).

3. Using the sequence of simulated product choices, I then *simulate* the corresponding inclusive values and update the associated transition probabilities.

4. Iterate the previous steps until convergence of the coefficients of the inclusive values transition processes.

I do not provide here a covergence proof for this algorithm. I note, however, that convergence was, in practice, achieved for *all* consumers after a small number of iterations. The remaining section examines several dimensions of the fit between the simulated and observed sequence of consumer choices.

The simulated probability that a consumer makes a purchase in any given week is 13.65%, which fits the observed probability (15.55%) reasonably well. Figure 5 analyzes how the model fits the purchase decision dynamics, by comparing the simulated and observed distribution of inter-purchase duration in days. Overall, the fit is very good, although it slightly underestimates the frequency of purchases for duration spells between 3 and 4 days, at the expense of slightly overestimating the frequency of purchases for durations of 7 days. Other than that, the model is quite accurate in simulating this interpurchase duration. Figure 6 examines the hazard rate of purchasing by duration in days from the last purchase, i.e. the probability that the consumer purchases a positive amount given that she has not purchased up to now. Again, the model predicts the pattern of the hazard rate quite accurately only very slightly underestimating the purchase probabilities for longer durations of no purchase spells, due to the low frequency of purchases with such duration.

Having addressed the purchase decision dynamics, I now move one to examine how the model fits the product switching decision dynamics. The simulated probability that a consumer exhibits a product switch from a purchase instance to the next is 63.79%, which only slightly underestimates the observed probability (72.81%). Figure 7 presents the distribution of purchases with regard to product switching. Although there is a slight underestimation of product switching, the fit is reasonably good.

43

FIGURE 5
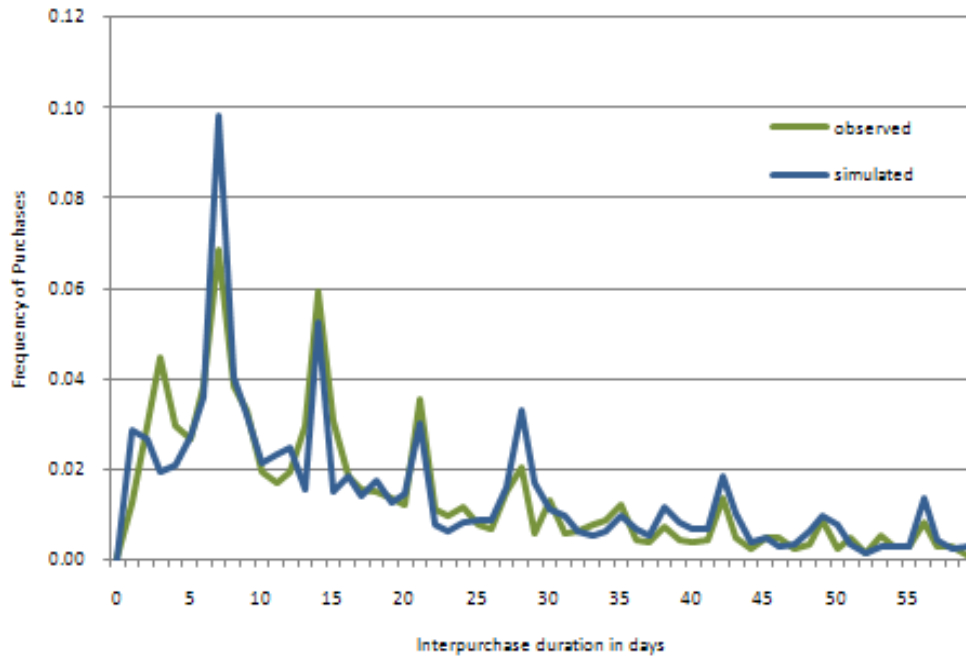
Observed and Simulated Interpurchase Duration Distribution

FIGURE 2.6

Observed and Simulated Purchase Hazard

FIGURE 7
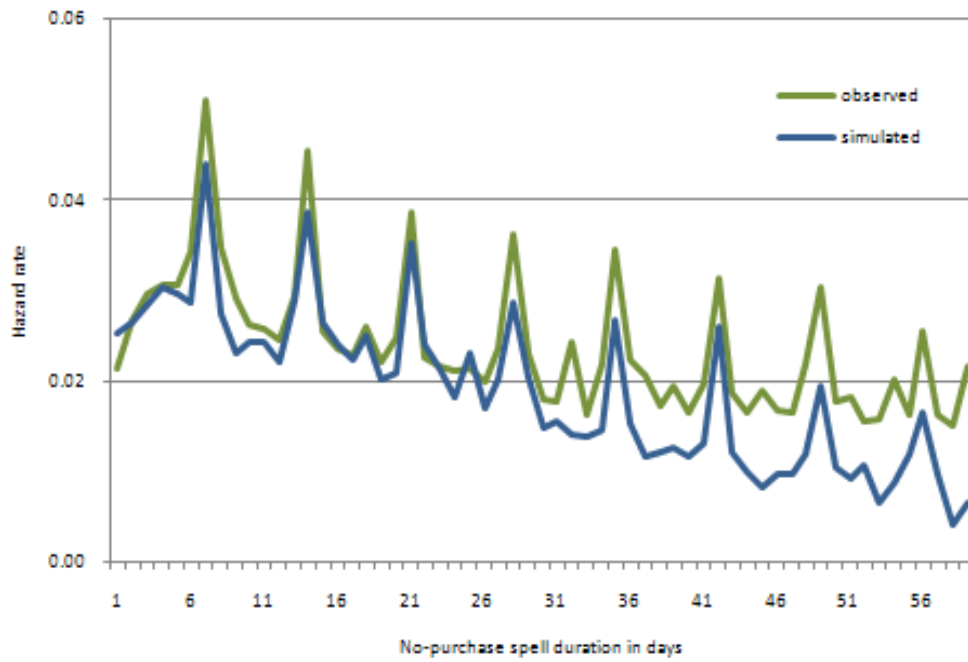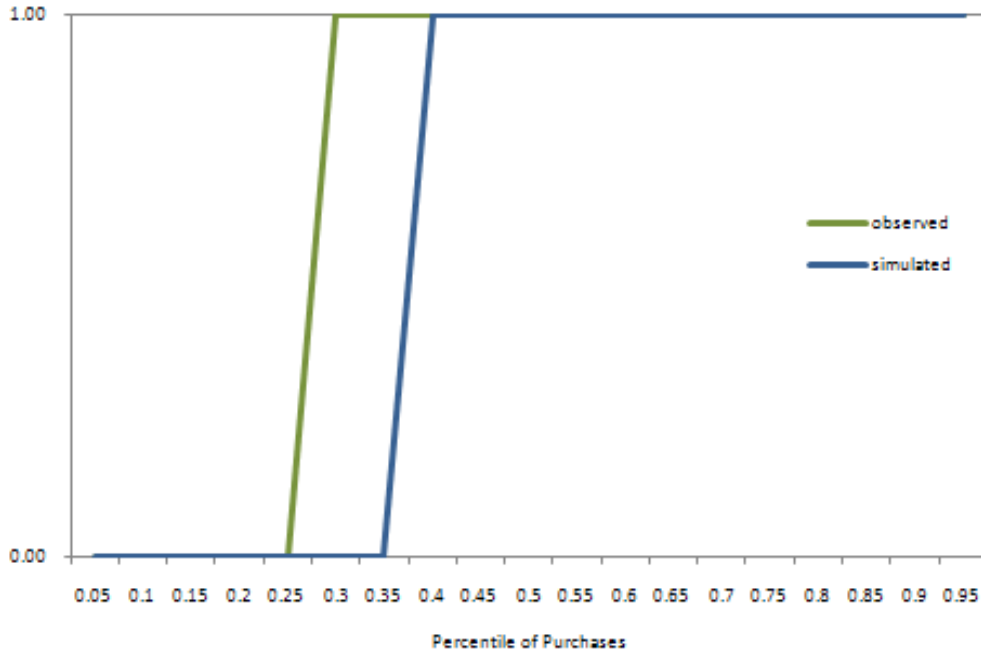*Observed and Simulated Product Switching Distribution*

Figure 7
Observed and Simulated Product Switching Distribution

6 Policy Implications

The pricing decision is one of the most critical for retailers. In this section, I discuss the implications of a major pricing policy change from hi-low pricing strategy to everyday low pricing (EDLP). In a pure EDLP policy, retailers charge a constant everyday price with no temporary price discounts. In contrast, in a hi-low pricing policy, prices have a higher regular level that remains constant for long periods of time, but then retailers run frequent promotions that lower the price below the EDLP level. In practice, however, pure EDLP strategies rarely exist (see Information Resources, Inc. (1993)). EDLP retailers typically charge lower prices on an everyday basis, but do engage in some temporary price discounts.

The pricing policy choice is an empirical question. Hi-low pricing policies have been prevalent in the industry since it allows retailers to price discriminate between consumers that are heterogenous in their price sensitiveness (Pigou (1920)), the degree of price information (Varian (1980)), the level of inventory costs (Blattberg *et al.* (1981), Jeuland and Narasimhan (1985)), or the extent of store loyalty (Narasimhan (1988)), just to mention a few dimensions. However, the success of retailers like Wal-Mart, Home Depot and Toys R Us has increased the popularity of EDLP policies. There are various rationales for adopting EDLP. On the supply side, EDLP is assume to lower operating costs through (*i*) better inventory control,

TABLE 10
Simulated Effects of Pricing Policy Changes *

Panel A: 0% Price Reduction	Variance Reduction		
	25%	50%	75%
Average Interpurchase Duration (Days)	-0.02	-0.13	-0.25
Proportion of Product Switching	-0.24	-1.35	-1.25
Total Volume Purchased	-0.07	-0.33	-0.43
Total Revenue	1.84	2.95	4.07
Panel B: 5% Price Reduction	Variance Reduction		
	25%	50%	75%
Average Interpurchase Duration (Days)	-0.28	-0.73	-0.77
Proportion of Product Switching	-0.30	-1.22	-1.48
Total Volume Purchased	0.14	0.45	0.62
Total Revenue	0.83	1.12	1.82
Panel C: 10% Price Reduction	Variance Reduction		
	25%	50%	75%
Average Interpurchase Duration (Days)	-0.47	-1.14	-1.31
Proportion of Product Switching	-0.36	-1.30	-1.30
Total Volume Purchased	0.35	1.01	1.22
Total Revenue	-0.17	-0.57	-1.31

* The table reports the percentage changes implied by the different policy changes
when compared to the actual pricing strategy.

warehouse handling and lower in-store personnel costs due to less variability in demand, and (*ii*) lower advertising expenses due to a focus on image rather than price. On the demand side, EDLP is assumed is to restore price credibility with consumers disenchanted with constant changing prices.

Table 10 examines this empirical question. I evaluate the *demand* implications of various degrees of a policy change from high-low pricing towards EDLP in four dimensions: the average interpurchase duration, the proportion of product switching, total volume purchased and total revenue. I should note that I do not compute market equilibrium prices, which is beyond the scope of this paper (although providing such a framework constitutes a very interesting potential area for future research). I consider only ad-hoc changes in the observed pricing strategy. The table reports the percentage changes implied by different policy changes when compared to the actual pricing strategy.

Table 10, Panel A addresses the implications of changing *only* the extent of the hi-low pricing policy by simulating a reduction in price variance, while keeping the mean price for each supermarket-product-size combination constant. The results suggest that a pricing policy that exhibits lower price variance slightly decreases the average interpurchase duration.

This is the outcome of two opposite effects. On one hand, the magnitude of the promotion price cuts is now smaller, which reduces the response and purchase acceleration of consumers. On the other hand, reducing price variance around the same mean price also decreases its regular level and consumers respond by increasing the frequency of purchases. In this particular case, the latter effect dominates the former, which induces a slight decrease in the average interpurchase duration.

The results also suggest that the proportion of product switching decreases. This is an expected outcome since lower price variance implies worst price deals, that reduce the promotional induced switching. Finally, the results also suggest that total volume sold drops while revenues increase. This is the outcome of a pricing policy that reduces price deals. Before, a proportional large share of the total volume sold was purchased in promotion. Under the new pricing policy, the attractiveness of temporary price promotions is reduced, which decreases the share of volume sold in promotion. The net effect is a decrease in the quantity sold, but an increase in revenues.

Table 10, Panels B and C address the implications of changing not only the extent of the hi-low pricing policy, but also the mean price offered. Here I simulate prices that have *both* a lower mean level and variance. The results with regard to average interpurchase duration and proportion of product switching are qualitatively similar to the ones in Panel A. Total volume sold increases, as expected, in response to the reduction in the mean prices. However this positive impact on quantity is not enough to compensate the drop in price, inducing a decrease in revenues when compared with Panel A pricing experiment.

In sum, the results suggest that the demand profitability of a major pricing policy change from hi-low towards EDLP is questionable, which supports the view that retailers are already maximizing profits. There is evidence that changing the extent of the hi-low pricing policy (by only reducing price variance, while keeping the mean price constant) may be revenue-increasing. However, not knowing the cost function, I can't determine the general impact on profits.

7   Conclusions

In this paper, I attempt to link two strands of the literature on variety-seeking: one focusing on substitutability across time and another on substitutability across products. This issue is economically relevant because both types of substitutability are important for retailers and manufacturers in designing intertemporal price discrimination strategies. I specify a

consumer demand model which allows consumption to have an enduring effect and allows the marginal utility of the different products to vary over consumption occasions. I then use the model to evaluate the demand implications of a major pricing policy change from hi-low pricing to a everyday low pricing.

I find evidence that consumption has a lasting effect on utility that induces substitutability across time and that the median consumer has a taste for variety in her product decisions. Consumers are found to be forward-looking with respect to the duration since the last purchase, to price expectations and product choices. Pricing policy simulations suggest that retailers may increase revenue by reducing the variance of prices, but that lowering the everyday level of prices may be unprofitable.

This paper leaves many estimation issues yet to be explored. The development of a framework that allows consumers to be forward-looking in their product choice or incorporates the supply side of the market to derive equilibrium pricing strategies seem very interesting potential areas for future research.

References

AGUIRREGABIRIA, V., 2002, "Sales promotions in supermarkets: Estimating their effects on profits and consumer welfare." Mimeo.

ANAND, B. N. and R. SHACHAR, 2004, "Advertising, the matchmaker." Mimeo.

BENITEZ-SILVA, H., G. HALL, G. J. HITSCH, G. PAULETTO, and J. RUST, 2000, "A comparison of discrete and parametric approximation methods for solving dynamic programming problems in economics." Mimeo.

BLATTBERG, R. C., G. D. EPPEN, and J. LIEBERMAN, 1981, "A theoretical and empirical evaluation of price deals for consumer nondurables." *Journal of Marketing*, 45, 116-129.

CHAMBERLAIN, 1985, "Heterogeneity, omitted variable bias, and duration dependence." In J. J. Heckman and B. Singer (Eds.), *Longitudinal analysis of labor market data* (pp. 3-38). N° 10 in Econometric Society Monographs series, Cambridge, New York and sidney: Cambridge University Press.

CHINTAGUNTA, P. K., D. C. JAIN, and N. J. VILCASSIM, 1991, "Investigation heterogeneity in brand preferences in logit models for panel data." *Journal of Marketing Research*, 28, 417-428.

CONSLIK, J., E. GERSTNER and J. SOBEL, 1984, "Cyclic pricing by a durable goods monopolist."

Quarterly Journal of Economics, 99, 489-505.

DAVIS, P., 2006, "The discrete choice analytically flexible (DCAF) model of demand for differentiated products." Mimeo.

DUBE, J. H., G. J. HITSCH and P. E. ROSSI, 2006, "Do switching costs make markets less competitive?" Mimeo.

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.

FAISON , E. W. J., 1977, "The neglected variety drive." *Journal of Consumer Research*, 4, 172-175.

GOLDFARB, A., 2006, "State dependence at internet portals". *Journal of Economics & Management Strategy*, 15(2), 317-352.

GONUL, F. and K. SRINIVASAN, 1993, "Modeling unobserved heterogeneity in multinomial to logit models: methodological and managerial implications," *Marketing Science*, 12, 213-229.

GORMAN , 1971, "Two stage budgeting." Mimeo.

GOURVILLE, J. T. and D. SOMAN, 1998, "Payment depreciation: the behavioral effects of temporally separating payments from consumption." *Journal of Consumer Research*, 25(2), 160-174.

HARTMANN, W. R. and V. B. Viard, 2008, "Do frequency reward programs create switching costs? A dynamic structural analysis of demand in a reward

program." *Quantitative Marketing and Economics*, 6, 109-137.

HARTMANN, W. R., 2006, "Intertemporal effects of consumption and their implications for demand elasticity estimates." *Quantitative Marketing and Economics*, 4, 325-349.

HECKMAN, J., 1981, "Heterogeneity and state dependence." In S. Rosen (Ed.), *Studies in labor markets* (pp. 91-139). Chicago: University of Chicago Press.

HENDEL, I. and A. NEVO, 2006, "Measuring the implications of sales and consumer inventory behavior." *Econometrica*, 74(6), 1637-1673.

—, 2006, "Sales and consumer inventory."*RAND Journal of Economics*, 37(3), 543-561.

HOCH, S. J., X. DREZE and M. E. PURK, 1994, "EDLP, hi-low, and margin arithmetic." *The Journal of Marketing*, 58(4), 16-27.

INFORMATION RESOURCES, INC., 1993, *Managing your Business in an EDLP Environment.* Chicago: IRI.

JAIN, D. C., N. J. VILCASSIM and P. K. CHINTAGUNTA, 1994, "A random-coefficients logit brand-choice model applied to panel data." *Journal of Business and Economic Statistics*, 12, 317–328.

JEULAND, A. P. and C. NARASIMHAN, 1985, "Dealing - temporary price cuts - by seller as a buyer discrimination mechanism." *Journal of Business*, 58(3), 295-308.

JEULAND, A. P., 1978, "Brand preference over time: a partially deterministic operationalization of the notion of variety seeking." In S. Jain (Ed.), *Research frontiers in marketing: dialogues and directions*. N° 42, AMA 1978 Educator's Proceedings, Chicago: American Marketing Association.

KAHN, B. E., 1995, "Consumer variety seeking among goods and services: an integrative review," *Journal of Retailing and Consumer Services*, 2(3), 139-148.

KAMAKURA, W. A. and G. J. RUSSELL, 1989, "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26, 379-390.

KEANE, 1997, "Modelling heterogeneity and state dependence in consumer choice behavior." *Journal of Business and Economic Statistics*, 15(3). 310-327.

KLEMPERER, P., 1995, "Competition when consumers have switching costs: an overview with applications to industrial organization, macroeconomics, and international trade." *Review of Economic Studies*, 62, 515-539.

LAL, R. and R. RAO, 1997, "Supermarket competition: the case of every day low pricing." *Marketing Science*, 16(1), 60-80.

MCALISTER, L., 1982, "A dynamic attribute satiation model of variety seeking behavior." *Journal of Consumer Research*, 9, 141-150.

— and E. PESSEMIER, 1982, "Variety-seeking behaviour: an interdisciplinary review." *The Journal of Consumer Research*, 9(3), 311-322.

MCFADDEN, 1989, "A method of simulated moments for estimation of discrete response models without numerical integration". *Econometrica*, 57: 995-1026.

—, 1981, "Econometric models of probabilistic choice". In C. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data* (pp, 198-272). MA: MIT Press.

MENON, S. and B. E. KAHN, 1995, "The impact of context on variety seeking in product choices." *Journal of Consumer Research*, 22, 285-295.

NARASIMHAN, C., 1988, "Competitive promotional strategies." *Journal of Business*, 61(4), 427-429.

OSBORNE, M., 2007, "Consumer learning, switching costs, and heterogeneity: a structural examination." Mimeo.

PAKES, A., 1986, "Patents as options: some estimates of the value of holding European patent stocks". *Econometrica*, 54, 755-784.

— and D. POLLARD, 1989, "Simulation and the asymptotics of optimization estimators". *Econometrica*, 57, 1027-1057.

PESENDORFER, M., 2002, "Retail sales: a study of pricing behavior in supermarkets". *Journal of Business*, 75(1). 33-66.

PESSEMIER, E. A., 1985, "Varied individual behavior: some theories, measurement methods and models." *Multivariate Behavioral Research*, 20, 69-94.

PIGOU, A. C., 1920, *The Economics of Welfare.* London: MacMillan.

POLLACK, R., 1970, "Habit formation and dynamic demand functions." *Journal of Political Economy*, 78(4), 745-763.

POZZI, A., 2009, "Shopping cost and brand exploration in online grocery." Mimeo.

PRELEC, D. and G. LOEWENSTEIN, 1997, "Beyond Time Discounting." *Marketing Letters*, 8, 97-108.

PUTERMAN, M. L. and M. C. SHIN, 1978, "Modified policy iteration algorithms for discounted Markov decision problems." *Management Science*, 24, 1127-1137.

RUST, J., 1987, "Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher". *Econometrica*, 55(5). 999-1033.

SEETHARAMAN, P. B. and H. CHE, 2009, "Price competition in markets with consumer variety seeking." *Marketing Science*, 28(3), 516-525.

SHUM, M., 2004, "Does advertising overcome brand loyalty? Evidence from the breakfast-cereals market." *Journal of Economics and Management Strategy*, 13, 91-109.

SIMONSON, I. and R. S. WINER, 1992, "The influence of purchase quantity and display format on consumer preference for variety." *The Journal of Consumer Research*, 19(1), 133-138.

SOBEL, J., 1984, "The timing of sales." *Review of Economic Studies*, 53(3), 353-368.

SPINNEWYN, F., 1981, "Rational habit formation." *European Economic Review*, 15, 91-109.

VARIAN, H., 1980, "A model of sales." *American Economic Review*, 70(4), 651-659.

VENKATESAN, M., 1973, "Cognitive consistency and novelty seeking." In S. Ward and T. S. Robertson (Eds.), *Consumer Behavior: Theoretical Sources*, Englewood Cliffs, NJ: Prentice Hall, 355-384.

WERTENBROCH, K., 2001, "Self-rationing: self-control in consumer choice." Mimeo.