Identification and Estimation of Forward-looking Behavior: The Case of Consumer Stockpiling*

Andrew T. Ching† and Matthew Osborne‡

February 15, 2017

Abstract

We develop a new empirical strategy for identifying the parameters of dynamic structural models in markets for storable goods, with a focus on identification of the discount factor. The identification strategy rests on an exclusion restriction generated by discontinuities in package sizes: In storable goods product categories where consumption rates are exogenous and small relative to package sizes, a consumer's current utility does not depend on inventory most of the time, his/her expected future payoff does. We demonstrate the feasibility of our identification strategy with an empirical exercise, where we estimate a stockpiling model using scanner data on laundry detergents. Our estimates suggest that consumers are not as forward-looking as most papers in the literature assumes; our estimates of weekly discount factors average at about 0.73, which is significantly lower than the value used in previous research (it typically is set at 0.99, using the market interest rate). We also find significant unobserved heterogeneity in discount factors across individuals.

Key words: Discount Factor, Exclusion Restriction, Stockpiling, Dynamic Programming

---

*We thank Stephan Seiler and seminar participants at the Summer Institute for Competitive Strategy, the Stanford Institute for Theoretical Economics, Boston College, McMaster University, University of Toronto, and University of Western Ontario for providing helpful feedback on earlier drafts of the paper.

†University of Toronto, Rotman School of Management. email: andrew.ching@rotman.utoronto.ca.

‡University of Toronto, Institute for Management and Innovation and Rotman School of Management. email: matthew.osborne@rotman.utoronto.ca

1 Introduction

Forward-looking behavior is a critical component of many quantitative models of consumer behavior used by researchers in marketing and economics (Erdem and Keane (1996), Crawford and Shum (2005), Hendel and Nevo (2006a), Erdem, Imai, and Keane (2003), Seiler (2013) and Liu and Balachander (2014), Osborne (2011), Yang and Ching (2014)). When consumers are forward-looking, they also behave strategically when making their purchase decisions. For instance, in the context of new durable goods such as cameras or smart phones, consumers may wait to purchase a product if they expect the product's price to fall in the future. Similarly, in the context of storable packaged goods such as canned tuna or canned soup, forward-looking consumers may respond to a temporary price promotion today by stockpiling the product, since they understand that future prices are likely to be high (Erdem, Imai, and Keane (2003), Haviv (2014)). If all shoppers are extremely forward-looking and act in such a saavy way, durable goods producers would not be able to use a price skimming strategy (Coase 1972), and grocery stores or supermarkets would never sell their carried items at a regular price. This is not the case, since in reality price skimming and hi-lo pricing are both prevalent, and consumers do make purchases of products when their prices are high. Prior research on periodic promotions in economics (Hendel and Nevo (2013), Hong, McAfee, and Nayyar (2002), Pesendorfer (2002), Sobel (1984)) has recognized that firms can use periodic promotions to price discriminate between patient and impatient consumers. Thus, the extent to which consumers consider the future clearly has implications for firms' optimal pricing strategies. Additionally, forward-looking behavior has important implications for public policy. One such area is in estimating price indexes, such as the Consumer Price Index. Price indexes are constructed by government agencies to measure inflation, and are used by businesses to index contracts. Standard price indexes will correctly measure changes in the cost of living if consumers are myopic; however, if consumers substitute purchases across time, recent research has suggested that standard indexes overstate growth in the cost of living (Feenstra and Shapiro (2003), Reis (2009), Osborne (2017)).

In these models, the strength of forward-looking behavior is captured by a parameter called a *discount factor*: the closer the discount factor is to 1, the more weight consumers put on future payoffs when making current decisions. However, instead of estimating the discount factor, much research focuses on estimated models of dynamic consumer behavior exercises the "rational expectations" assumption and uses the prevailing interest rate to fix the discount factor accordingly. Depending on the length of a period, this calibration approach would lead to a value of weekly discount factor about 0.99.[1] Interestingly, Frederick, Loewenstein, and O'Donoghue (2002) sur-

[1] At a yearly interest rate of 5%, a rational consumer would discount utility in the following year at a rate of about

vey prior experimental work in measuring discount factors: They document a significant amount of heterogeneity in the estimates those studies obtained, ranging between close to 0 and close to 1. Additionally, in stated choice experiments performed by Dubé, Hitsch, and Jindal (2014), consumers appear to be much less forward-looking than economic theory implies, with average discount rates of 0.43. Dubé, Hitsch, and Jindal (2014) also find substantial heterogeneity in discount factors across individuals.

The reason why the discount factor is typically not estimated in structural econometric work stems from an identification problem: Most problems under study do not provide natural exclusion restrictions that could help identify this parameter, and so any estimate of the discount factor would be heavily reliant on functional form assumptions. Roughly speaking, to address this problem one would need to have at least one state variable that impacts a consumer's future payoffs, but not her current payoffs. In econometrics terminology, such a variable provides exclusion restrictions that helps to identify the discount factor (since it is excluded from current payoffs but not future payoffs). The intuition is that if a consumer is completely myopic, then the consumer's choice should be independent of that variable. The extent to which consumer's choice is influenced by the exclusion restriction provides information about how forward-looking consumer is.

We contribute to this literature by arguing that one of the key state variables of the stockpiling problem, consumer inventory, provides exclusion restrictions that can help identify the discount factor. Our key insight is that for most inventory levels, consumer's current payoff does not depend on it because the storage cost does not change until a package runs out. We illustrate how using an example drawn from the laundry detergent market. Suppose a consumer is down to her last bottle of laundry detergent. She washes one full load of clothes per week (and such need is driven by her habit of wearing clean clothes every day). As she keeps consuming the laundry detergent, she may worry that if she does not buy another bottle soon when the price is low, she may be forced to buy it at a higher regular price when she uses it up in the near future. This sense of urgency becomes stronger as inventory (i.e., the amount of detergent in the bottle) runs down, and her demand would appear to become more sensitive to price cuts. Moreover, for any amount of inventory remaining, the more forward-looking a consumer is, the more intense this feeling of urgency will get.

However, note that if a consumer is totally myopic, then inventory should not affect her behavior, unless she runs out. A myopic consumer will only care about having enough detergent to do the current week's laundry, and her current utility will only be affected by the storage cost, which *does not change* since she still has a single bottle taking up the same amount of space. The example illustrates that inventory can provide exclusion restrictions to help identify the consumer's discount

0.95, and would have a weekly discount rate of about 0.999.

factor, because inventory impacts the consumer's expected future payoff, but not her current payoff.

Intuitively, if inventory is observed, the reseacher can compute the probability a consumer makes a purchase at each level of inventory. For a forward-looking consumer, this probability should rise smoothly as her inventory drops; for a myopic consumer, it will not change as inventory drops, until she runs out. Hence, if inventory is observed, the change in the purchase probability that is observed as inventory decreases can help identify the discount factor.[2]

A complication for our approach to identification is that much research that estimates structural stockpiling models uses supermarket scanner data, which does not track consumer inventory. Hence, the main state variable of interest, inventory, is *unobserved* to the researcher. We argue that when inventory is unobserved, the identification can still be achieved by the observed purchase hazard.[3] The exclusion restrictions (reduction in inventory does not affect storage costs most of the time) generate overidentifying restrictions, which help us to identify the parameters of our stockpiling model, including the discount factor.

To demonstrate our identification arguments, we first show that we can recover the true parameter values of a stockpiling model even when inventory is unobserved using artificial data experiments. We then estimate a more complicated stockpiling model using IRI scanner data for laundry detergents, allowing for continuously distributed unobserved heterogeneity in most of the model parameters, particularly the discount factor. We find that consumer discount factors range from about 0.6 and 0.85, and average at about 0.73. The values of the discount factors for most consumers are significantly lower than the value of 0.95 or 0.99 that many papers assume when estimating dynamic discrete choice models of consumer behavior. Additionally, we find that higher income, older and larger households seem to be more forward-looking, although the effect of demographics on the discount factor is small; most of the heterogeneity in discount factors we estimate seems to be driven by unobserved factors (this result is also consistent with Dubé, Hitsch, and Jindal (2014)). Our results could have strong substantive implications in answering the questions that the literature has examined (e.g., short-term vs. long-term responses to temporary and permanent price cuts).

An outline of the paper is as follows. In Section 2, we discuss related work. Section 3 introduces

[2]We note that in standard formulations of stockpiling models, researchers assume that storage costs increase continuously as inventory increases. This assumption creates an identification problem, since a myopic individual's purchase probability will increase smoothly as inventory drops, making identification of forward-looking behavior difficult.

[3]It is defined to be the average probability of a purchase occurring τ periods after a purchase occured in period t, with no purchase occuring in the intervening time.

a simple stylized model of stockpiling behavior, and Section 4 contains proofs of the important properties of the model. Section 5 presents conditions for identification of the discount factor when inventory is observed. Section 6 describes how the discount factor can be identified when inventory is unobserved. Section 7 describes the results of our artificial data experiments. Section 8 describes our empirical application and the estimates, and Section 9 concludes.

2 Review of Literature

Proofs of identification of the discount factor often build on the conditional choice probability approach introduced in Hotz and Miller (1993). In the Hotz and Miller (1993) approach, the researcher assumes that the same state variables that are observed to the consumer are observed to the researcher, and there is no unobserved heterogeneity across consumers. In this setting, under a set of regularity conditions on the error term, one can flexibly estimate a consumer's *choice specific value*, which is the sum of the current period flow utility and the discount factor multiplied by the value function. The choice specific values are identified conditional on a normalization of the utility of one alternative (typically called the reference alternative), and given the functional form of the error distribution. With no restrictions on the functional form of the flow utility, the discount factor is not identified: in the conditional choice probability approach, one can think of each estimating equation as the probability of a consumer choosing each alternative at each value of all the state variables. A fully flexible model would allow the utility function to have a parameter that was unique for each alternative and each state. Hence, if the discount factor were fixed the number of equations and unknowns would be equal, and the model would be exactly identified. Formally, to identify the discount factor, some restriction must be put on the functional form of the utility function. Such a restriction will reduce the number of parameters in the model to be smaller than the number of equations, allowing the discount factor to be identified.

One such type of restriction that has been proposed to help identify the discount factor is called an *exclusion restriction*. Fang and Wang (2015) show that one can identify the discount factor in the conditional choice probability setting if a dynamic model has at least two values of a state variable where, for each alternative, flow utilities are the same for both values, but the value functions differ.[4]

[4]Magnac and Thesmar (2002) is widely cited as the first paper which shows how exclusion restrictions can identify the discount factor. However, it should be pointed out that their exclusion restriction is defined in a way that is quite different from Fang and Wang (2015) and what we use here. It is difficult to give economic interpretations to the exclusion restriction used in Magnac and Thesmar (2002).

To our knowledge, there are only a handful papers that explore such an identification argument to estimate consumer's discount factor or her incentive to consider future payoffs (Ishihara and Ching (2012), Chung, Steenburgh, and Sudhir (2013), Lee (2013), Ching, Erdem, and Keane (2014), Chevalier and Goolsbee (2009)).[5] Moreover, as far as we know, the previous structural models on consumer stockpiling models all assume that the storage cost is an increasing and continuous function of inventory. This simplifying assumption, though convenient, has ruled out the exclusion restrictions that we use in our identification arguments. As a result, all of the previous structural works in consumer stockpiling fix the discount factor according to the interest rate, instead of estimating it.[6]

3 A Stylized Stockpiling Model

In this section we describe a model that is simplified somewhat from the model we will use for our empirical application, but contains its most important features. The econometrician observes a market containing N consumers making purchase decisions over T periods. Consumers are forward-looking and discount the future at a discount rate $\beta_i < 1$. In this stylized model, we assume that a single product is available to consumers in some discrete package size. Each decision period t is broken up into two phases: a purchase phase and a consumption phase. In the purchase phase, consumer i observes the price of a package of the product (p_{it}), an exogenous consumption need (c_{it}), and a choice-specific error (ε_{ijt}). The consumer's choice is her decision of how many packages of the product to buy, which we denote as $j \in \{0, 1, ..., J\}$. After making her purchase, the consumer receives her consumption utility.

We denote the size (or volumn) of a package as b, and for simplicity of exposition we assume that b is an integer (we will relax this assumption in the empirical application). We denote the consumer's

[5] Additionally, on-going research by Akça and Otter (2015) describes an alternative mechanism by which inventory can be used to identify the discount factor. They argue that if inventory is observed by researchers and consumers consume the inventory in a last-in-last-out order, then the discount factor can be identified. Our approach, which focuses on package size discontinuities, rather than the order in which brands are purchased, can also handle unobserved inventory. Geweke and Keane (2000) and Yao, Mela, Chiang, and Chen (2012) explore another identification strategy which requires making assumptions that the current payoffs are either observed or can be recovered from a static environment first.

[6] Note that with the assumption that the storage cost is an increasing and continuous function of inventory, a consumer has an incentive to wait longer before buying a new bottle, since the storage cost keeps dropping as the inventory shrinks. This has the opposite effect of the increase in expected stock-out cost as the inventory drops. Therefore, the models in the previous works do not have clear implications about consumer purchase behavior as inventory drops.

inventory (which will also be integral) at the beginning of the period as I_{it}. Consumption rates c_{it} will be in the set $\{0, 1, 2, ..., \bar{c}\}$. If the consumer's inventory at the end of the purchase phase, which we denote as $I_{it} + b \cdot j$, is above the consumption need c_{it} then she receives consumption utility γ_i. If she cannot cover her consumption need then she incurs a stockout cost ν_i.[7] At the end of the period, the consumer incurs a storage cost $s(\cdot; \boldsymbol{\omega}_i)$. Here we formally introduce our first assumption about $s(\cdot; \boldsymbol{\omega}_i)$, which would allow the inventory variable to generate exclusion restrictions.

First Model Assumption Related to Exclusion Restrictions, X1

1. The storage cost function s is only a function of the number of packages held at the end of the period, B, rather than inventory I, and the package size $b > 1$.

The number of packages held can be written as the following function of inventory $B_{i,t+1}(j, I, c_{it}) = \lceil \max\{(I_{it} + b \cdot j - c_{it})/b, 0\} \rceil$.[8] The assumption that $b > 1$ ensures that X1 is meaningful. $\boldsymbol{\omega}_i$ is a vector of parameters determining how storage costs vary with the number of packages held. We will parameterize the storage cost function as flexibly as possible:

$$s(B; \boldsymbol{\omega}_i) = \omega_{i,B}. \tag{1}$$

This functional form is nonparametric in the sense that there is a different parameter, $\omega_{i,B}$ for each possible number of packages held. In practice, one may consider imposing a functional form on s, such as quadratic. We will assume that the cost of storing 0 packages is 0.

The assumption that a consumer's storage cost depends on the number of packages held is valid for many product categories. For example, products that are sold in bottles or boxes such as laundry detergent or breakfast cereal will likely satisfy this assumption. The cost to storing laundry detergent depends on the amount of space taken up by the bottle, but not the amount of liquid within the bottle.

Another crucial assumption is that consumption rate is exogeneous in the sense that it does not depend on one's inventory. This assumption will ensure that the current payoff does not vary with inventory continuously.

Second Model Assumption Related to Exclusion Restrictions, X2

2. The consumption need is exogenous (i.e., it is not a function of inventory).

[7] We assume that the stockout cost does not depend on the consumption need but this assumption is innocuous. We could also assume that the stockout cost is proportional to the difference between inventory and the consumption shock, and our identification results will be unaffected.

[8] The ceiling function $\lceil \cdot \rceil$ returns the smallest integer that is greater than or equal to its argument.

Intuitively, this assumption says consumers receive no additional utility from consuming more than their consumption needs. This assumption should also be largely applicable to products like laundry detergent, breakfast cereal, etc. But for products such as snacks, this assumption might not hold (e.g., Sun (2005))

We should note that our exclusion rerstriction argument would fail if the realizations of consumption need, c_{it}, are discrete and take on values that are multiples of the package size, as storage costs would always change when inventory changes. To be a bit more precise, we would like consumption need to be significantly less than the package size.

Third Model Assumption Related to Exclusion Restrictions, X3

3. Consumption needs are much smaller than the package size: $c_{it} << b$.

This assumption ensures that for most values of inventory levels faced by consumers, the number of packages held remains unchanged (and hence storage costs remain unchanged). For product categories such as laundry detergent, ketchup, etc., this assumption will likely be satisfied. But for categories like canned tuna (or canned soup), this assumption will likely be violated.

Given this information, we can write down the consumer's flow utility as follows:

$$\begin{aligned} & u_{it}(j, I_{it}, \varepsilon_{ijt}, p_{it}, c_{it}; \boldsymbol{\theta}_i) \\ & = \begin{cases} \gamma_i - s(B_{i,t+1}(j, I, c_{it}); \boldsymbol{\omega}_i) - \alpha_i p_{it} j + \eta \varepsilon_{ijt} & \text{if } I_{it} + b \cdot j \geq c_{it} \\ -\nu_i - \alpha_i p_{it} j + \eta \varepsilon_{ijt} & \text{otherwise} \end{cases} . \end{aligned} \tag{2}$$

The vector $\boldsymbol{\theta}_i = (\alpha_i, \beta_i, \gamma_i, \nu_i, \boldsymbol{\omega}_i)$ is a vector of the consumer utility coefficients and the discount factor. The parameter α_i is the price coefficient. The parameter η is an error term weight that will be relevant for some of the theoretical proofs. In estimation we will normalize this parameter to 1.

We assume that consumers believe that the product's price follows a stochastic Markov process with a transition density $F(p_{i,t+1}|p_{it})$. The consumption shocks c_{it} are i.i.d over time for each consumer and are drawn from a discrete distribution where the probability of receiving consumption shock level l as π_c^l. The consumer's Bellman equation is as follows:

$$V_{it}(I_{it}, p_{it}) = \sum_{l=0}^{\bar{c}} E_{\varepsilon_{it}} \max_{j=0,\ldots,J} \{u_{it}(j, I_{it}, \varepsilon_{ijt}, p_{it}, l; \boldsymbol{\theta}_i) + \beta_i E_{p_{i,t+1}|p_{it}} V(I_{i,t+1}, p_{i,t+1})\} \pi_c^l. \tag{3}$$

The transition process for the inventory state variable I_{it} is

$$I_{i,t+1} = \max\{I_{i,t} + b \cdot j - c_{it}, 0\}.$$

We also put an upper bound on the number of packages a consumer can carry, which we denote M. We assume that if a consumer makes a purchase when her inventory is above $Mb - c_{it}$, then her inventory is set to the upper bound Mb. Intuitively, this is consistent with a situation where a consumer's storage space is used up, but if she purchases another bottle she takes the one that is already open and gives it away or otherwise disposes of it.

4 Model Properties

In the derivations below we will normalize the consumption utility, γ, to 0. We can do this because if a consumer runs out, her decision can be written in terms of $\gamma + \nu$. If she does not, γ appears in both utility of purchasing and of not purchaseing, and so it does not affect the purchase decision.

In this section we will derive some useful properties of the model above, which will help us to understand intuitively what type of variation in the data will help identify the discount factor. The basic idea behind the identification of β is that there are exclusion restrictions in our model: Consumer inventory enters the expected future payoffs in a continuous way, and it almost never directly affects consumer's current payoffs.[9] Intuitively, unless a consumer is very close to using up a bottle/package of laundry detergent, washing an extra load of laundry will lower the inventory level, but it does not change the storage cost. Therefore, if a consumer cares about her future payoffs, her incentive to purchase should depend on the inventory. Consider the situation where a consumer's inventory is down to the last bottle. As she continues to consume, she gets closer and closer to stocking out. At the same time, her incentive to avoid the stock out cost ν also gets stronger and stronger if she is more forward-looking (i.e., $\beta > 0$). In other words, the functional relationship between consumer's purchase incidence and inventory should depend on the value of her discount factor. Therefore, a consumer's discount factor should be identified if we observe consumer's choice at different inventory levels.

Formally, we can express the above intuition by deriving two key properties of the value function. First, the expected future value of a purchase should increase as inventory drops, at least for sufficiently low values of inventory. Second, as consumers get more forward-looking, the expected future value of a purchase should rise. We will prove these statements are true in a simple setting with no price variation, and where the error term has zero variance ($\eta = 0$). We will then

[9] As we describe in more detail below, the two exceptions are at the exact point a consumer stocks out, and at the point a consumer uses up a package. We will argue that when inventory is observed, it is sufficient to observe some range of inventory where storage cost does not change in order to guarantee identification. When inventory is unobserved, changes in storage cost occur rarely enough that the discount factor will still be identified.

demonstrate that the value function is continuous in η under some regularity conditions on the error term, and if payoffs can be bounded, which will demonstrate that for small values of η the same properties of the value function will hold.

We will make some simplifications to the model which we will relax in later sections. To simplify notation we will normalize the price coefficient, α, to 1. We will also assume that all the model parameters are homogeneous across the population, and thus will drop the i subscript on everything except the state variables and the error term.

The simplifying assumptions that we will maintain for the remainder of this section are listed below:

Assumptions A1-A6

1. The consumption need is constant across time and individuals: $c_{it} = 1$ for all i, t.
2. In a given purchase occasion, the maximum number of packages a consumer is allowed to buy is 1.
3. Prices are fixed over time at a level $p > 0$.
4. The package size $b \geq 2$.
5. Purchasing at 0 inventory is better than running out: $\omega_1 + p < \nu$.
6. The storage cost function is weakly increasing and weakly convex, and storage costs are weakly positive.

Assumption A1 is made for convenience; the propositions proved in this section will still hold under stochastic consumption rates, which we demonstrate in Online Appendix 11. Assumption A4 is a more precise statement of the exclusion restriction assumption X3 from the prior section. Assumption A5 will imply that the stockout cost is positive, as $p > 0$ and the storage cost parameters are weakly positive by A6.

Lemma 1 *If A1-A5 holds and $\eta = 0$, then it is optimal to purchase only when $I = 0$, for all β.*

Proof.

Case 1: $\beta = 0$. This is given by A5.

Case 2: $\beta > 0$.

First note that if a consumer only buys when $I = 0$, then $I \geq 0$. Then by A1 and A4, a consumer always receive $\gamma, \forall t$. Hence, the discounted sum of utility, $U_1 = \frac{\gamma - \omega_1}{1-\beta} - \frac{p}{1-\beta^b}$.

Claim 2a: It is not optimal for a consumer to choose not to buy when $I = 0$.

If a consumer chooses not to buy when $I = 0$, then the discounted sum of utility, $U_2 = \frac{-\nu}{1-\beta}$.

Note that

$$U_1 = \frac{\gamma - \omega_1}{1-\beta} - \frac{p}{1-\beta^b} > \frac{\gamma - \omega_1 - p}{1-\beta}. \tag{4}$$

Since $\gamma > 0$, it follows from A6 that $\frac{\gamma-\omega_1-p}{1-\beta} > \frac{-\nu}{1-\beta} = U_1$. It then follows from the above equation that $U_2 > U_1$. This shows Claim 2a.

Claim 2b: It is not optimal for a consumer to choose to buy when $I > 0$.

Note that $U_1 = \frac{\gamma}{1-\beta} - \frac{\omega_1}{1-\beta} - \frac{p}{1-\beta^b}$.

If a consumer makes a purchase when $I > 0$, it will only make the second and third components in U_1 more negative. This proves the claim.

■

Lemma 2 *If A1-A6 holds and $\eta = 0$, then $V(I)$ is increasing in I, for $I < \bar{I}$ and sufficiently small storage costs.*

Proof. Denote as x the number of packages held by a consumer, and $n = I - b(x-1)$ as the number of units left in the package currently begin consumed by an individual. The value function for I units of inventory can be written as:

$$V(I) = \frac{1-\beta^I}{1-\beta}\gamma - \frac{1-\beta^{n-1}}{1-\beta}\omega_x - \sum_{k=1}^{x-1} \beta^{n-1+b(k-1)} \frac{1-\beta^b}{1-\beta}\omega_{x-k} + \beta^I V(0) \tag{5}$$

Case 1: Suppose an increase in I does not change the number of packages.

Then it follows from Lemma 1 that $V(I)$ increases with I. This is because the future stream of storage costs and consumption do not change, and the future stream of payments is being postponed.

Case 2: Suppose that an increase in I leads to an increase in the number of packages being stored.

The only difference between here and Case 1 is that the future stream of storage costs is incresaed by $\omega_B - \omega_{B-1}$ for a finite number of periods (by one period if we consider the increase in I is 1). As long as $(\omega_B - \omega_{B-1})$ is sufficiently small, $V(I)$ is still increasing in I.

■

Proposition 1 *If A1-A6 hold and $\eta = 0$, then $\beta * [V(I+b) - V(I)]$ is decreasing in I, for $I \geq c$.*

Proof. Denote as x the number of packages held by a consumer. Note that

$$\begin{aligned}
V(I+b) &= (\gamma-\omega_x)*\frac{(1-\beta^b)}{(1-\beta)}+\beta^{b+1}*V(I)\\
V(I+b)-V(I) &= (\gamma-\omega_x)*\frac{(1-\beta^b)}{(1-\beta)}+\beta^{b+1}*V(I)-V(I)\\
V(I+b)-V(I) &= (\gamma-\omega_x)*\frac{(1-\beta^b)}{(1-\beta)}+(\beta^{b+1}-1)*V(I)
\end{aligned}$$

Because $\beta^{b+1}-1<0$, and $V(I)$ is increasing in I by Lemma 2, it follows that $\beta*[V(I+b)-V(I)]$ is decreasing in I. ■

Proposition 2 *If A1-A6 hold and $\eta=0$, then $\beta*[V(I+b)-V(I)]$ is decreasing in β, for $I\geq 0$ and sufficiently small storage costs.*

Proof.

To start note that we need to show

$$\frac{\partial\beta(V(I+b)-V(I))}{\partial\beta}=V(I+b)-V(I)+\beta\left(\frac{\partial(V(I+b)-V(I))}{\partial\beta}\right)>0. \tag{6}$$

A sufficient condition for the above inequality to hold is $V(I+b)-V(I)>0$, which we can show in Proposition 1 for sufficiently small storage costs. Next, we want tosign the derivative $\frac{\partial(V(I+b)-V(I))}{\partial\beta}$. It is possible to show that this derivative is positive if storage costs are sufficiently small.

To start, because lemma 1 implies a consumer does not buy until she runs out, it has to be the case that if $I\leq b+1$ (note that if $I=b+1$, the individual holds 2 packages at the beginning of the period, but one of the packages is used up during the period, and storage costs are paid at the end of the period) then

$$V(I)=\frac{1-\beta^I}{1-\beta}\gamma-\frac{1-\beta^{I-1}}{1-\beta}\omega_1+\beta^I V(0) \tag{7}$$

The first term is the consumption utility discounted over I periods. The second is the discounted storage cost. Storage costs are paid for $I-1$ periods, since in the $I-1$st period an individual has only 1 unit at the beginning of the period, uses up that unit, and has no packages at the end of the period; hence no storage cost is paid I periods from now.

Lemma 1 implies that an individual purchases when beginning the period with 0 inventory, which means we can derive the following formula for $V(0)$:

$$\begin{aligned}
V(0) &= \gamma-\omega_1-p+\beta V(b-1)\\
&= \frac{1-\beta^b}{1-\beta}\gamma-\frac{1-\beta^{b-1}}{1-\beta}\omega_1-p+\beta^b V(0),
\end{aligned}$$

Where the last line follows by substituting in equation (7). We can solve the above equation to derive an explicit formula for $V(0)$:

$$\begin{aligned} V(0) &= \frac{1}{1-\beta^b}\left(\frac{1-\beta^b}{1-\beta}\gamma - \frac{1-\beta^{b-1}}{1-\beta}\omega_1 - p\right) \\ &= \frac{1}{1-\beta}\gamma - \frac{1}{1-\beta^b}\frac{1-\beta^{b-1}}{1-\beta}\omega_1 - \frac{p}{1-\beta^b} \qquad (8) \end{aligned}$$

The derivations above can be used to derive a formula for $V(I)$, for all inventory values of I. The general formula is a little more complicated since we have to account for the fact that storage costs may change. To derive a general formula that includes storage costs, we denote as x the number of packages held by a consumer, and $n = I - b(x-1)$ to be the number of units left in the package currently begin consumed by an individual. The formula for the value function will be

$$V(I) = \frac{1-\beta^I}{1-\beta}\gamma - \frac{1-\beta^{n-1}}{1-\beta}\omega_x - \sum_{k=1}^{x-1}\beta^{n-1+b(k-1)}\frac{1-\beta^b}{1-\beta}\omega_{x-k} + \beta^I V(0) \qquad (9)$$

We will split up the value function difference $V(I+b) - V(I)$ into three terms:

$$V(I+b) - V(I) = \Delta_1 + \Delta_2 + \Delta_3$$

The first term, Δ_1, we define to be the difference in the first term from equation (9):

$$\begin{aligned} \Delta_1 &= \frac{1-\beta^{I+b}}{1-\beta}\gamma - \frac{1-\beta^I}{1-\beta}\gamma \\ &= \frac{1-\beta^{I+b}-(1-\beta^I)}{1-\beta}\gamma \\ &= \frac{\beta^I - \beta^{I+b}}{1-\beta}\gamma \\ &= \beta^I\frac{1-\beta^b}{1-\beta}\gamma \qquad (10) \end{aligned}$$

We define Δ_2 to be the difference in the terms in equation (9) that contain storage costs:

$$\Delta_2 = -\frac{1-\beta^{n-1}}{1-\beta}\Delta\omega_{x+1} - \sum_{k=1}^{x-1}\beta^{n-1+b(k-1)}\frac{1-\beta^b}{1-\beta}\Delta\omega_{x-k+1} - \beta^{n-1+b(x-1)}\frac{1-\beta^b}{1-\beta}\omega_1$$

The final term, Δ_3, is defined to be the difference in the final term of (9), which contains $V(0)$:

$$
\begin{aligned}
\Delta_3 &= (\beta^{I+b} - \beta^I)V(0) \\
&= \beta^I(\beta^b - 1)V(0) \\
&= \frac{\beta^I(\beta^b-1)}{1-\beta}\gamma - \frac{\beta^I(\beta^b-1)}{1-\beta^b}\frac{1-\beta^{b-1}}{1-\beta}\omega_1 - \frac{\beta^I(\beta^b-1)p}{1-\beta^b} \\
&= \frac{\beta^I(\beta^b-1)}{1-\beta}\gamma + \frac{\beta^I(1-\beta^{b-1})}{1-\beta}\omega_1 + \beta^I p \qquad (11)
\end{aligned}
$$

Note that Δ_1 cancels with the first term in Δ_3 from the last line of equation (11). As a result, cosumption utility does not affect the difference in value functions.

Second, consider the impact of storage costs on the value function difference. This difference can be written as Δ_3 plus the second term of equation (11):

$$
\Delta_2 + \frac{\beta^I(1-\beta^{b-1})}{1-\beta}\omega_1 = -\frac{1-\beta^{n-1}}{1-\beta}\Delta\omega_{x+1} - \sum_{k=1}^{x-1}\beta^{n-1+b(k-1)}\frac{1-\beta^b}{1-\beta}\Delta\omega_{x-k+1} - \beta^{n-1+b(x-1)}\omega_1
$$

Since this term is negative, if storage difference increase, the first term of equation (6) will decrease, decreasing the overall derivative. Second the derivative of the term $\Delta_2 + \frac{\beta^I(1-\beta^{b-1})}{1-\beta}\omega_1$ does not have a clear sign. For instance, the derivative of the first term is

$$
\frac{(n-1)\beta^{n-2}(1-\beta) - (1-\beta^{n-1})}{(1-\beta)^2}\Delta\omega_{x+1},
$$

which could be negative (ie for small β). The terms in the summation sign will have derivatives that look as follows:

$$
\left[-(n-1+b(k-1))\beta^{n-2+b(k-1)}\frac{1-\beta^b}{1-\beta} + \beta^{n+b(k-1)}\frac{b\beta^{b-1}(1-\beta)-(1-\beta^b)}{(1-\beta)^2}\right]\Delta\omega_{x-k+1},
$$

which again may be negative. The term $-\beta^{n-1+b(x-1)}\omega_1$ will be decreasing in β. Note that if storage costs are zero, then it will be the case that $V(I+b) - V(I) = \beta^I p$. In this case the derivative will be

$$
V(I+b) - V(I) + \beta\frac{\partial(V(I+b)-V(I))}{\partial\beta} = \beta^I(I+1)p,
$$

which is positive. Because the difference in value functions is continuous in storage costs, the above derivative will still be positive, as long as storage costs are sufficiently small.

■

The size of the storage costs clearly plays a role here. In reality one would expect storage costs to be relatively small for the first few packages, and increasing and convex after that. For instance, suppose that an individual who stores laundry detergent has some dedicated space for bottles of detergent, so that ω_1 is close to zero. Then for $I \leq b+1$, both Propositions 1 and 2 will hold. The propositions could still hold for larger values of I if the increase in storage costs from adding more bottles is not too large: this results from the fact that for $I > b+1$, the value function difference depends on storage cost differences.

Both proposition 1 and 2 assume that there is no error term in the utility function. We will use the following lemma to show that these two propositions still hold in a random utility framework. The argument relies on showing that the expected future value of purchase is continuous in η. Here we make some regularity assumptions on the error term, and put boundedness and sign restrictions on the payoffs:

Assumptions E1-E2

1. Continuity and support: The CDF of the difference in $\varepsilon_1 - \varepsilon_0$, F, is continuous, strictly increasing, and has support $(-\infty, \infty)$.
2. Value function: There exists a bound on η, $\overline{\eta}$, such that if $\eta < \overline{\eta}$ then the following hold

$$
\begin{aligned}
I - c \geq 0: &\quad -p - (\omega_{B+1} - \omega_B \mathbf{1}\{I > c\} + \beta(V(I+b-c) - V(I-c)) < 0 \\
I - c < 0: &\quad -p - (\omega_1 - \nu + \beta(V(I+b-c) - V(0)) > 0
\end{aligned}
$$

where the number of packages held at the end of the period, B, is $\lceil (I+b-1)/b \rceil$ if $I > 1$ and 0 if $I = 1$.

The two assumptions above, along with Assumptions A1-A4, will imply that the value function is continuous in η, which we summarize in the following lemma:

Lemma 3 *If assumptions A1-A6 and E1-E2 hold then the expected future value of purchase from an increase in inventory, $\beta\left[V(I+b) - V(I)\right]$, is continuous in η.*

Proof. For $I \geq c$, the probability of a purchase can be written as

$$P(I,c) = F((-p - (\omega_{B+1} - \omega_B \mathbf{1}\{I > c\} + \beta(V(I+b-c) - V(I-c)))/\eta), \tag{12}$$

while the probability of purchase for $I < c$ is

$$P(I,c) = F((-p - (\omega_1 - \nu + \beta(V(I+b-c) - V(0)))/\eta). \tag{13}$$

The value function for $I \geq c$ can be written as

$$V(I) \;=\; P(I,c)(-p - \omega_{B+1} + \beta V(I+b-c)) + (1 - P(I,c))(-\omega_B \mathbf{1}\{I > c\} + \beta V(I-c)). \tag{14}$$

Under E1 and E2 it is the case that if $I - c \geq 0$ then

$$\lim_{\eta \to 0^-} P(I,c) = 0,$$

and otherwise

$$\lim_{\eta \to 0^-} P(I,c) = 1.$$

The limits above are taken from the left as we assume that $\eta \geq 0$. If we consider the first limit, we know that if $I \geq c$ then for η sufficiently close to zero, the net value of buying becomes negative, which we assume in E2 (and Lemma 1 implies this inequality holds for $\eta = 0$). For η arbitrarily small and positive the term $(-p - (\omega_{B+1} - \omega_B \mathbf{1}\{I > c\} + \beta(V(I+b-c) - V(I-c)))/\eta$ will be negative and will approach $-\infty$. E1 guarantees that the probability in (12) will approach 0. A similar argument applies to the second limit in the context of equation (13). As a result, it is clear that the limit as η approaches zero of equation (14) will equal the value function that is obtained when $\eta = 0$, which we derive in the proofs of Lemma 1 and Lemma 2. Similar findings will be obtained for the value function when $I < c$.

■

Lemma 3 further shows that these properties hold in a random utility framework, suggesting that the magnitude of the incentive to purchase can be measured by choice probabilities. With Lemma 3, it is clear that both proposition 1 and 2 hold even if we introduce an error term in the utility function, as long as η is sufficiently small. To provide some more intuition in Figure 1, we plot the expected future value of a purchase for different values of the discount factor, for a low and high value of the stockout cost ν. For a forward-looking consumer, the expected future value of a purchase rises as inventory drops because purchasing delays the likelihood of a future stockout. Additionally, when inventory is sufficiently low, increasing the discount factor increases the expected future payoff from purchase. When storage costs are positive and the stockout cost is low, at sufficiently high levels of inventory the expected future value of purchase can be decreasing in the discount factor, since adding inventory will increase storage costs in the future, which will counterbalance the gain from delaying the stockout cost.

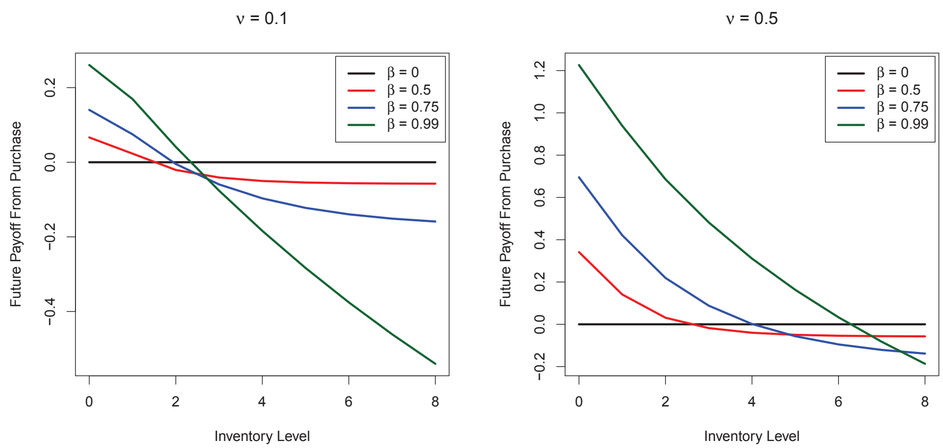

Figure 1: Expected future payoff from purchase, $\beta(V(I+b)-V(I))$, as a function of I and β. Parameter values $\omega_1 = 0, \omega_2 = 0.05, \omega_3 = 0.15, \eta = 1$, $M = 3$, $p = 2$, and logit error term.

5 Identification with Observed Inventory

In this section, we will discuss how the discount factor and other parameters of the model can be identified in the situation in which inventory is observed. Although in most empirical applications inventory will be unobserved, we feel that understanding the features of the model that drive identification in this setting will help the reader to understand what drives identification in the setting where inventory is not observed. For convenience, we will maintain three assumptions: (i) prices do not vary; (ii) consumers can only purchase one package at a time; (iii) inventory is integral.

We outline the intuition behind the identification using numerical solution of the model at particular parameter values, and then provide some discussion of formal conditions that will guarantee identification. To illustrate this intuition we plot the purchase probabilities as a function of time for different discount factors in Figure 2, in the following situation: We suppose that in period 0 the consumer starts with an inventory level of $I = 16$, which is 2 full packages of the product. We maintain the assumption that the individual's consumption rate is 1, so she has no inventory at the beginning of period 17. Consider first the black line, which shows the probability of purchase for a completely myopic consumer. This consumer's purchase probability is flat except at 3 periods: period 8, where a package is used up and the storage cost drops, period 16 where the storage cost is

0 at the end of the period, and period 17 where the consumer runs out. In periods 1 through 8, an individual's purchase probability will be $Pr(-\alpha p - (\omega_3 - \omega_2) + \varepsilon_{i1t} - \varepsilon_{i0t} > 0)$. Given a particular value of the price coefficient α, the value of the purchase probability in this interval will identify the storage cost difference $\omega_3 - \omega_2$. Similarly, the level of the purchase probability in periods 9 through 15 will identify $\omega_2 - \omega_1$. In period 16, a consumer who runs out will not pay the stockout cost since she has a single unit of inventory, and will not pay the storage cost since she uses up her last package and storage costs are incurred at the end of the period, and so her purchase probability is $Pr(-\alpha p - \omega_1 + \varepsilon_{i1t} - \varepsilon_{i0t} > 0)$; this purchase probability will therefore identify ω_1. In period 17, the individual runs out and her purchase probability will be $Pr(-\alpha p + \nu + \varepsilon_{i1t} - \varepsilon_{i0t} > 0)$, so the purchase probability at zero inventory would identify the stockout cost ν.

The discount factor, β, will be identified from the slope of the purchase probability in areas where the myopic consumer's flow utility is flat as inventory drops because of the exclusion restrictions made in Assumption X1. In these areas of the state space, the purchase probability rises when inventory drops for a forward-looking individual is a result of Proposition 1. In addition, as β increases the probability of purchase rises for low values of inventory, which is a result of Proposition 2. Consumers who are more forward-looking will try harder to avoid stockouts, and this can be seen in larger slopes of the purchase probability for larger values of β. As a final note, we have not yet discussed the identification of α, the price coefficient. Technically, this coefficient is identified from purchase probabilities when the consumer has $M = 3$ packages in inventory. Because we assume that when an individual has M packages in inventory she throws away her current package and sets her inventory level to Mb, the purchase probability for a myopic consumer in such a situation would be $Pr(-\alpha p + \varepsilon_{i1t} - \varepsilon_{i0t} > 0)$. However, in general it is preferable to obtain identification of the price coefficient from price variation, rather than an assumption about how inventory is filled up when a consumer reaches her maximum storage capacity.

We now turn to a more formal discussion of identification. The researcher observes consumer choice probabilities at different values of the inventory state. These choice probabilities are functions of choice-specific values, which we define as the present discounted value of purchasing or not purchasing. To simplify the analysis we follow Fang and Wang (2015) and Abbring and Daljord (2016) and assume that the choice-specific error term follows a Type 1 Extreme Value distribution. Denote $\hat{P}(I)$ as the probability of purchase at inventory level I. Define v_j to be the choice-specific value of buying j packages at inventory level I and parameter vector $\boldsymbol{\theta} = (\alpha, \beta, \nu, \omega_1, ..., \omega_M)$:

$$v_j(I; \boldsymbol{\theta}) = -\alpha p \mathbf{1}\{j = 1\} - \omega_{B(j,I,1)} - \nu \mathbf{1}\{I = 0\} + \beta V(\max\{I + bj - 1, 0\}). \tag{15}$$

Under the logit error assumption we can write the choice probabilities in terms of choice-specific

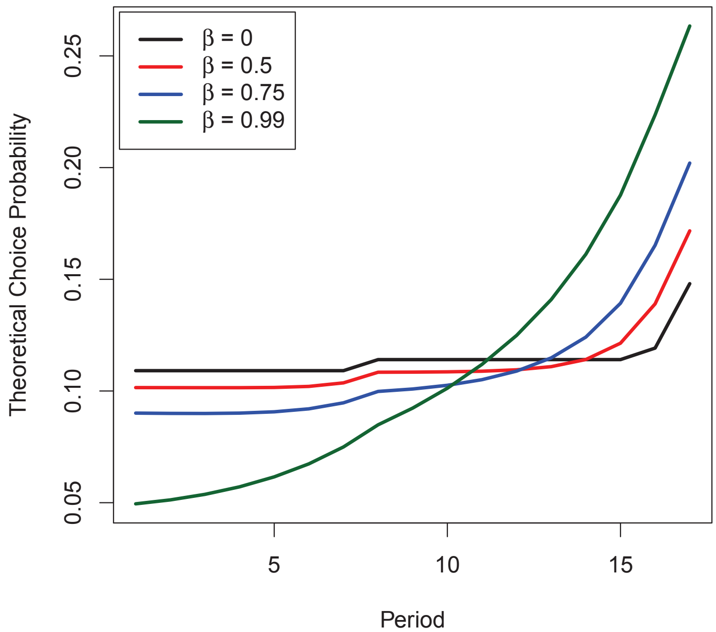

Figure 2: Probability of purchase in period t given inventory of 16 (2 packages) in period 0, where $c = 1$ for all periods. Parameter values $\nu = 0.25, \omega_1 = 0.1, \omega_2 = 0.25, \omega_3 = 0.75, \eta = 1$, $M = 3$, $p = 2$, and logit error term.

values as follows

$$\log(\hat{P}(I)) - \log(1 - \hat{P}(I)) = v_1(I; \boldsymbol{\theta}) - v_0(I; \boldsymbol{\theta}). \tag{16}$$

If a consumer can hold up to M packages, then the number of parameters we need to identify is $M + 3$: there are M different values of ω_B, there is the stockout cost ν, the discount factor β, and the price coefficient α. As a result, we need to be able to compute the moments in equation (16) for at least $M+3$ different values of inventory. In particular, the preceding informal discussion suggests that in order to identify all the model parameters, one must at least observe choice probabilities when inventory is 0 (to pin down the stockout cost), 1 (to pin down ω_1), and at least once for every possible number of packages a consumer can hold: In other words, if a consumer can hold up to M packages we need to observe the choice probability at a value of I in the interval $[2, b+1]$, the interval $[b + 2, 2b + 1]$, and so on up to $[(M - 1)b, Mb]$ (to pin down ω_2 up to ω_M and α). Additionally, for at least one package size we should observe two inventory levels so we can compute the slope of the purchase probability with respect to inventory, which will map into the discount factor as a result of our exclusion restriction. Denote the values of inventory defined thus as $\boldsymbol{I} = (I_1, I_2, I_3, ..., I_{M+3})$, where $I_1 = 0$, $I_2 = 1$, and I_3 through I_{M+2} are in the intervals $[2, b]$, $[b+1, 2b]$, etc, and the final value I_{M+3} is the value that lies in one of the intervals $[2, b]$, $[b+1, 2b]$, etc, but is different from I_3 through I_{M+2}.

Define the parameter vector we want to identify as $\boldsymbol{\theta} = (\alpha, \beta, \nu, \omega_1, ..., \omega_M)$, and denote the difference in choice-specific values on the right hand side of equation (16) as $\Delta v(I; \boldsymbol{\theta}) = v_1(I; \boldsymbol{\theta}) - v_0(I; \boldsymbol{\theta})$. Define the vector of choice-specific value differences $\Delta \boldsymbol{v}(\boldsymbol{I}; \boldsymbol{\theta})$ to be the vector of $\Delta v(I; \boldsymbol{\theta})$'s evaluated at the inventory levels $\boldsymbol{I}$: $\Delta \boldsymbol{v}(\boldsymbol{I}; \boldsymbol{\theta}) = (\Delta v(I_1; \boldsymbol{\theta}), \Delta v(I_2; \boldsymbol{\theta}), ..., v(I_{M+3}; \boldsymbol{\theta}))$. Suppose the researcher observes choices generated by the model outlined above, at a parameter value $\boldsymbol{\theta}_0$. Given a dataset with N observations, we define the vector of such choice probabilities for the state vector $\boldsymbol{I}$ as $\hat{\boldsymbol{P}}_0(\boldsymbol{I}; N)$. The parameters of the dynamic discrete choice model proposed in our paper are identified if the solution to the system of equations

$$\Delta \boldsymbol{v}(\boldsymbol{I}; \boldsymbol{\theta}) = \lim_{N \to \infty} \log(\hat{\boldsymbol{P}}_0(\boldsymbol{I}; N)) - \log(\mathbf{1} - \hat{\boldsymbol{P}}_0(\boldsymbol{I}; N)) \tag{17}$$

in terms of $\boldsymbol{\theta}$ is $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Our assumption that the dimension of $\Delta \boldsymbol{v}(\boldsymbol{I}; \boldsymbol{\theta})$ is at least $M+3$, which equals the number of parameters, is necessary to allow for identification. If the system of equations defined by (17) is full rank at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, then the parameter $\boldsymbol{\theta}_0$ is locally identified as a consequence of the Implicit Function Theorem: in some neighborhood of $\boldsymbol{\theta}_0$ there is a unique solution to equation (17) and the solution is $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. We formalize the assumptions necessary for local identification in Online Appendix 12.1. The rank condition can be verified by the researcher for a given set of choice probabilities.

Abbring and Daljord (2016) provide examples of dynamic discrete choice model specifications where local identification holds but global identification fails: there may be two or more distinct values of the discount factor, and distinct flow utility parameters, that can rationalize the same observed choice probabilities.[10] Abbring and Daljord (2016) argue that if there are two states, I_A and I_B, where the difference $v_1(I_A; \boldsymbol{\theta}) - v_0(I_A; \boldsymbol{\theta}) - (v_1(I_B; \boldsymbol{\theta}) - v_1(I_B; \boldsymbol{\theta}))$ is monotonic in β, then the model parameters are globally identified. In our case, this monotonicity is difficult to verify when there is an error term since the value functions are complicated functions of all the model parameters. However, we can gain some insight from analyzing the explicit formulas for the choice-specific values derived in the proof of Proposition 2. Suppose that η is small, and that I_A and I_B are chosen such that storage costs do not change (i.e., the exclusion restriction holds), and suppose WLOG that $I_A > I_B$ and $\alpha = 1$. Then the difference in choice-specific values will be

[10] Here the parameters are distinct in the sense that there is an open neigborhood around each one that does not contain the others.

approximately $\beta^{I_A}p - \beta^{I_B}p$. Generally the equation $\beta^{I_A}p - \beta^{I_B}p = \log(\hat{P}(I_A)) - \log(1 - \hat{P}(I_A)) - (\log(\hat{P}(I_B)) - \log(1 - \hat{P}(I_B)))$ will have two solutions strictly between 0 and 1.[11] We can guarantee a unique solution to β however if we observe choice probabilities at inventory values of I_A, I_B, $I_A + 1$, and $I_B + 1$, and if the storage costs do not change in this interval. In this case there will be two equations to solve for β:

$$\beta^{I_B+1}(\beta^{I_A-I_B} - 1) = (\log(\hat{P}(I_A + 1)) - \log(1 - \hat{P}(I_A + 1)) - \log(\hat{P}(I_B + 1)) - \log(1 - \hat{P}(I_B + 1)))/p$$
$$\beta^{I_B}(\beta^{I_A-I_B} - 1) = (\log(\hat{P}(I_A)) - \log(1 - \hat{P}(I_A)) - \log(\hat{P}(I_B)) - \log(1 - \hat{P}(I_B)))/p$$

Note that if we divide the first equation by the second we can solve uniquely for β, which will be the ratio of log choice probability differences. Although the above example required observing choice probabilities at four values of inventory in some interval where storage costs do not change, it would be sufficient to observe choice probabilities at three values of inventory that are one unit apart, i.e., I_B, $I_A = I_B + 1$, and $I_A + 1$.

We make two comments on three simplifying assumptions stated at the beginning of the section. First, the assumption that prices do not vary, and that consumers can only purchase one package at a time, can be relaxed. If we relax these assumptions, identification in fact becomes easier, because there are more moments to help pin down the parameters. The tradeoff is that the model becomes more complicated to analyze, and we must completely rely on numerical solution of the model, which we leave to Section 6.2. The second comment relates to the assumption of inventory that is integral and consumption rates that are fixed over time. Under stochastic consumption rates, we can prove analogs to Propositions 1 and 2 (see Online Appendix 11). As a result, the identification arguments will not be substantially different. Continuous inventory should also not be problematic: even if inventory is continuous, as long as the exclusion restrictions hold there will be some areas of the state space where current utility does not vary with inventory, but the value function will if $\beta > 0$. We explore identification with continuous inventory (and a particular function form assumption on the storage cost function) using artificial data experiments in Online Appendix 13.

[11] To see this, note we can write $\beta^{I_A} - \beta^{I_B} = \left(\log(\hat{P}(I_A)) - \log(1 - \hat{P}(I_A)) - (\log(\hat{P}(I_B)) - \log(1 - \hat{P}(I_B)))\right)/p$ as $\beta^{I_B}(\beta^{I_A-I_B} - 1) = \left(\log(\hat{P}(I_A)) - \log(1 - \hat{P}(I_A)) - (\log(\hat{P}(I_B)) - \log(1 - \hat{P}(I_B)))\right)/p$. The difference $\beta^{I_A} - \beta^{I_B} \leq 0$ and equals zero at $\beta = 0$ and $\beta = 1$. The derivative of this equation will be $\beta^{I_B-1}(I_A\beta^{I_A-I_B} - (I_A - I_B))$. This derivative will be positive if $\beta > (I_B/I_A)^{1/(I_A-I_B)}$, and negative otherwise. The term $\beta^{I_A} - \beta^{I_B}$ thus has a unique minimum, and for any value of $\left(\log(\hat{P}(I_A)) - \log(1 - \hat{P}(I_A)) - (\log(\hat{P}(I_B)) - \log(1 - \hat{P}(I_B)))\right)/p$ between 0 and the minimum of the function there will be two solutions.

6 Identification with Unobserved Inventory

Scanner datasets that are typically used by researchers to estimate stockpiling models do not track consumer inventory, meaning that the identification strategy based on exclusion restrictions described in the previous section does not directly apply. In this section we discuss the type of variation in the data that can identify the model parameters when inventory is not observed. Although inventory is unobserved, the time between purchases, which is correlated with inventory, is observed. As a result, the discount factor may be identified from the impact of interpurchase time on a consumer's purchase probability, which is captured by the purchase hazard. The complication here is that inventory is unobserved, and it must be integrated out when forming the likelihood. Since this probability is analytically complicated, most of our analysis focuses on numerical simulations.

In this section we relax the assumption of a constant consumption rate, and allow the consumption rate c_{it} to be stochastic. In the numerical solutions below we will assume that the package size, b, is 8 units, and consumption shocks are in the set $\{1, 2\}$. We denote the probability that an individual receives a consumption draw of 1 as π_c. We will allow consumers to store up to $M = 3$ packages and assume the error term is standard logit. When we compute the purchase hazard, we will need to simulate out the steady state distribution of inventory in the population. To do this we will simulate purchases for 500 individuals for 600 periods. We will assume that in period 0 all individuals have 0 inventory. We find that aggregate inventory appears to reach the steady state at around 50 periods, so we will use periods 400 to 600 to compute steady state inventory. For much of the discussion (except the final discussion of Section 6.2) we will also hold prices fixed over time at a level of 2.

Remarks: It is worth comparing our approach here with that in Hendel and Nevo (2006b), which proposes a series of tests for the presence of forward-looking behavior in storable goods markets. The paper develops a stockpiling model with endogenous consumption from inventory, and where consumers are able to purchase quantities in continuous amounts. In their setting, the key difference between a myopic consumer and a forward-looking consumer is that a myopic consumer will alway purchase exactly the amount she will consume in the period where the purchase occurs, while a forward-looking consumer will purchase for future consumption. An implication of the model developed by Hendel and Nevo (2006b) is that the purchase hazard will be completely flat for myopic individuals, which allows a clean test for the presence of forward-looking behavior. This type of analysis will apply well to settings where consumers have the ability to purchase the product category in small increments: for example, canned tuna or soup. A key difference between Hendel and Nevo (2006b)'s setting and ours is that in our setting myopic consumers will purchase more than they can consume in a single period, since package sizes are large relative to consumption rates. As a result, in our setting the purchase hazard will not be completely flat for myopic individuals, violating the implications of Hendel and Nevo (2006b). Because of this complication, we rely on exclusion restrictions to separate out myopic consumers from forward-looking consumers, rather than relying on identification from quantity purchased. Because we rely on exclusion restrictions, we can identify the discount factor in situations where consumers are only able to purchase a single package in a purchase occasion. Another key difference is that in our setting we assume consumption rates are exogenous, in the sense that consumers use enough of a product to satisfy an exogenous consumption need (for example, one does not get extra utility from consuming more laundry detergent than is needed to do the weekly laundry, or one seldom gets extra utility from drinking more coffee than his/her consumption need). We note that the exclusion restriction may be violated in a setting where consumption is endogenous, since optimal consumption (and hence flow utility) can be a function of inventory.

6.1 Identification with Unobserved Inventory, and no Storage Costs

We begin by considering the case where $\omega_i = 0$ for $i = 1, ..., M$. In Figure 3 we plot the aggregate probability of purchase in period $t + \tau$ given a purchase in period t for different values of the discount factor. The discount factor primarily affects two features of the purchase hazard. The first feature is the slope of the purchase hazard in the periods immediately after a purchase occurs. For our particular parameterization, a purchase increases an individual's inventory by 8 units. Since consumption shocks are at most 2 units, it will take someone at least 4 periods to run out and incur a stockout cost. To see the implications of this, consider the purchase hazard for a myopic individual, shown by the black line in Figure 3. For the first 3 periods after a purchase, the purchase hazard is flat, since a myopic consumer's flow utility is fixed over this interval. In contrast, for a forward-looking consumer the purchase hazard has a positive slope over the first 3 periods, and this slope increases as the discount factor rises. This occurs because the expected future value of purchase rises as inventory drops, as we showed in Proposition 2. It is notable that without storage costs, there is a clean test for whether individuals are forward-looking or not - if individuals are myopic the purchase hazard should be flat for the initial few periods after a purchase, provided that it takes some time for individuals to run out of a package after a purchase (which we maintain in Assumption X3). If consumption rates are very high, then the purchase hazard will always be flat, and the discount factor will not be identified. Intuitively, if individuals always use up all their inventory right after a purchase, then it will be difficult or impossible to tell if individuals are myopic or not. We discuss this implication in a more formal context in Online Appendix 12.2.

The second, and more subtle difference between the purchase hazards, is that the purchase hazard becomes smoother as β rises (note that this feature of the purchase hazard also arises in the purchase probabilities with observed inventory in Figure 2). The intuition here is that a myopic consumer is not willing to trade off future utility for current utility, so her purchase hazard will start to rise sharply at $\tau = 4$, when people in the population start to run out. In contrast, a forward-looking consumer will be more willing to purchase early, and so the purchase hazard will be smoother for such a consumer.

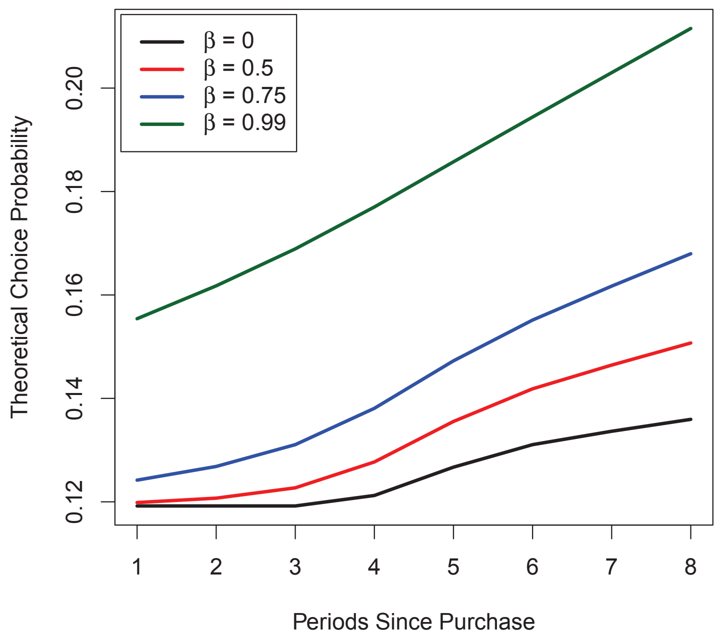

Figure 3: Probability of purchase in period $t+\tau$ given purchase of 1 package in period t. Parameter values $\nu = 0.25, \pi_c = 0.5, \omega_1 = 0, \omega_2 = 0, \omega_3 = 0, \eta = 1$, $M = 3$, $p = 2$, and logit error term.

The identification problem with unobserved inventory becomes more complicated than with observed inventory since we have to consider separate identification of the discount factor β, the stockout cost ν, and the probability of a low consumption shock π_c. A feature of the model that aids identification is that ν and π_c have very different effects on the purchase hazard than β. Figure 4 shows how ν affects the purchase hazard, for low and high values of the discount factor. Most of the impact of a change in ν on the purchase hazard occurs during later rather than earlier periods. This is sensible since ν should have more impact on purchase decisions when consumers begin to run out. Importantly, the shape of the purchase hazard is preserved as ν changes - for the low value of β, the purchase hazard displays a lot of curvature around period 4 for different values of

ν. Similarly, the purchase hazard is very smooth for high values of β for different values of ν. The impact of π_c on the purchase hazard for different values of β is shown in Figure 5. For low values of β, the impact of changing the probability of a low shock is similar to that of ν. For high values of β, changing π_c shifts the purchase hazard up and down. Our analysis suggests that ν and π_c could be difficult to separately identify if β is low. Indeed, we encounter this problem in our empirical application in section 8.2, and need to calibrate the consumption rate prior to estimating other structural parameters.[12] The price coefficient, α, will shift the overall purchase probability and will simply shift the purchase hazard up or down, and so (in the absence of price variation) it will be identified by the average purchase probability.

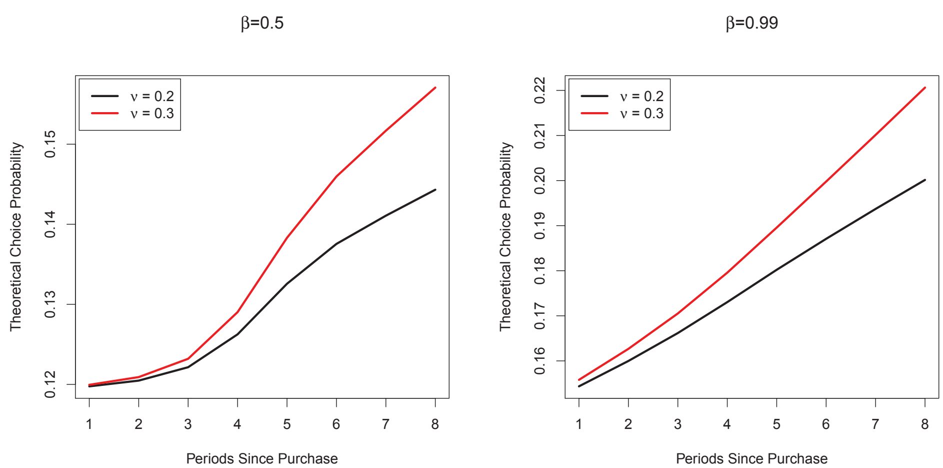

Figure 4: Probability of purchase in period $t+\tau$ given purchase of 1 package in period t, for different values of the stockout cost. Parameter values $\pi_c = 0.5, \omega_1 = 0, \omega_2 = 0, \omega_3 = 0, \eta = 1$, $M = 3$, $p = 2$, $\alpha = 1$ and logit error term.

6.2 Nonzero Storage Costs

If individuals have storage costs, the identification argument becomes somewhat more complicated because increases in storage costs can also increase the slope and decrease the curvature of the purchase hazard. To see why, note that when an individual makes a purchase, there is some chance that she has a small amount of a package left over. An individual in this situation will use up the package within a few periods after the purchase, and will observe a decrease in their storage costs. That decrease in storage costs will lead to an increase in the probability of a purchase. To

[12] In the empirical model consumption rates are constant, but π_c essentially controls the average consumption rate in our simulation, which is $\pi_c + 2(1-\pi_c)$. A similar identification issue should arise even with a constant consumption rate.

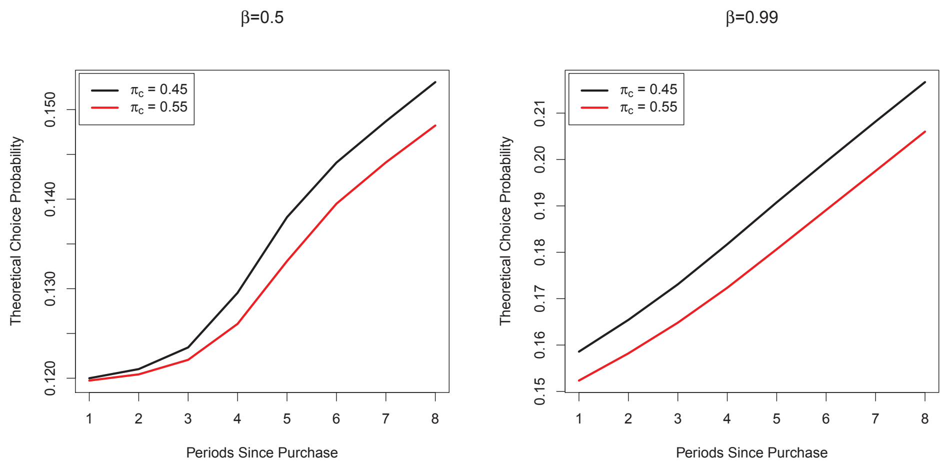

Figure 5: Probability of purchase in period $t+\tau$ given purchase of 1 package in period t, for different values of π_c. Parameter values $\nu = 0.25, \pi_c = 0.5, \omega_1 = 0, \omega_2 = 0, \omega_3 = 0, \eta = 1$, $M = 3$, $p = 2$, $\alpha = 1$ and logit error term.

show that this is the case, in Figure 6 we compute the purchase hazards for different discount factors, with $\omega_3 = 0.5$. The purchase hazard for a myopic consumer is positively sloped in the first 3 periods after a purchase - recall in Figure 3 a myopic consumer's purchase hazard was flat in this region, suggesting that the positive storage cost parameter is responsible for the increase in slope. Additionally, the black line in in Figure 6 is also smoother than the one in Figure 3, suggesting that storage costs can smooth out the purchase hazard.

How do we approach the issue of identification in the presence of storage costs? In the following discussion we explore two different avenues. One avenue is to argue that, because of the exclusion restrictions, there will not be enough storage cost parameters to completely fit the purchase hazard. In our example, we compute the purchase hazard for 8 periods, which means we have at least 8 moments. The number of parameters we have to fit these moments is 7 - three storage cost parameters (ω_1, ω_2,and ω_3), the stockout cost parameter (ν), the discount factor (β), the price coefficient (α), and the probability of a low consumption shock (π_c). Focusing on the discount factor, it is the case that even in the presence of storage costs, an increase in the discount factor still increases the slope of the purchase hazard (at least in early periods) and decreases its curvature. As a result, unless a rank condition fails letting the discount factor be free will provide an improved fit to these features of the purchase hazard. We state this rank condition in Online Appendix 12.2 - local identification can be obtained if the rank of the Jacobian of the theoretical purchase hazard is at least as large as the number of model parameters one needs to estimate. The exclusion restrictions

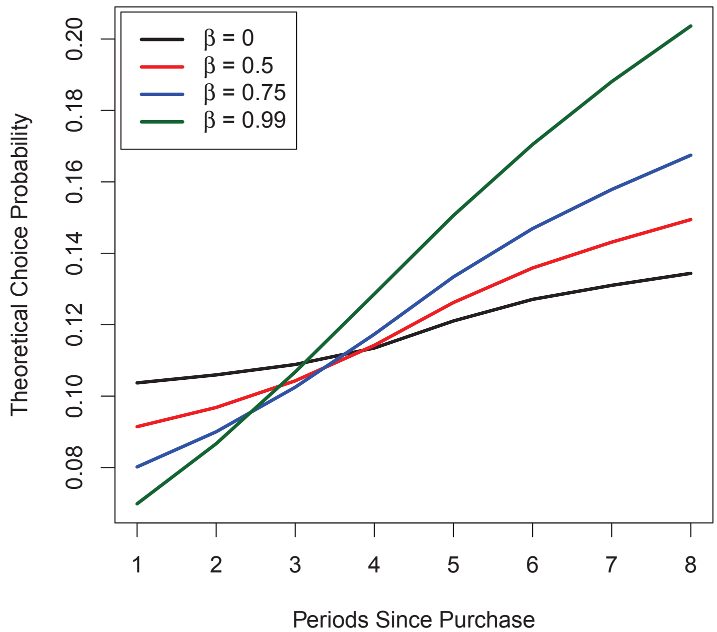

Figure 6: Probability of purchase in period $t+\tau$ given purchase of 1 package in period t. Parameter values $\nu = 0.25, \pi_c = 0.5, \omega_1 = 0, \omega_2 = 0, \omega_3 = 0.5, \eta = 1$, $M = 3$, $p = 2$, $\alpha = 1$ and logit error term.

X1 through X3 will help to guarantee that this rank condition holds. Exclusion restriction X1 reduces the number of model parameters that one needs to estimate to something manageable. Assumption X3 guarantees that individuals will not run out so quickly that the purchase hazard becomes degenerate. A caveat to the formal approach is that the rank condition which may be difficult to verify in practice. Moreover, the identification result will be local, rather than global, as global identification will be more difficult to verify.

A second solution is to focus on additional moments that may be generated by price variation, and stockpiling in response to price variation. A forward-looking consumer should become more sensitive to discounts as her inventory drops, since the value of avoiding future stockouts is higher the higher is the discount factor. To examine this, we analyze an extension to the model where we allow the price variable to take on two values, 1 and 2, where the value of the price follows a Markov transition process. The probability of the price 2 given 2 occurred the previous period is 0.8, and the probability of 2 given last period's price was 1 is 0.9. Thus most of the time prices are high, but periodically they drop to the low price for a short time, as is commonly observed in scanner data for storable goods. Additionally, we relax the restriction that individuals can only purchase a single package, and allow individuals to purchase up to 2 packages at once. Because we assume that the transition process for prices is known to consumers, individuals may stock up when prices are low.

One measure of an individual's propensity to stock up is the amount by which the probability a person buys 2 units relative to 1 unit increases when the price drops. Intuitively, if a consumer is myopic she has no need to purchase more than a single package - as a result, the propensity to purchase more than one package should be driven entirely by the error distribution. However, a forward-looking consumer should become relatively more likely to purchase multiple units at low prices, and this likelihood should increase as inventory drops. In Figure 7, we plot the ratio of probability of buying 2 units to 1 unit at the low price minus the same ratio at the high price, given t periods have elapsed since the last purchase occured. The left panel shows how this probability difference changes if there are no storage costs. It is notable that for a myopic individual, the propensity to stockpile in response to deals is completely unaffected by inventory, since the moment we show in the graph is totally flat. However, this moment increases as inventory drops for forward-looking individuals, and the slope of the curve rises with larger values of β. The right panel shows the same moment with positive storage costs. Here, the propensity to stockpile in response to low prices still rises if inventory is sufficiently low. An important point to note is that if the discount factor is low, the slope of the lines are relatively unaffected by the storage cost (compare the black and red lines on the left panel to the right panel). However, if the discount factor is high, the storage costs decrease the slope of the line, which is intuitive - if storage costs are high and individuals are forward-looking they should have less incentive to stockpile at low prices. The fact that the storage costs and discount factor have the opposite effect on the propensity to stockpile, while they both increase the slope of the purchase hazard, will help us to separately identify them. We note that price variation will also identify the price coefficient, but the price coefficient can be identified from the average change in the purchase probability for the low versus the high price; the preceding argument relies on how stockpiling in response to deal sensitivity changes as inventory changes.

6.3 Identification with Consumer Unobserved Heterogeneity

All of the analysis above has assumed that there is no persistent unobserved heterogeneity across consumers. A formal argument for identification with persistent unobserved heterogeneity would rely on the time dimension of our data going to infinity at a rate that is fast enough relative to the cross-sectional dimension that one could estimate the purchase hazards and average purchase quantities described above on an individual basis. Using this type of argument one could, in principle, identify individual-specific discount factors. In field settings infinite amounts of data are

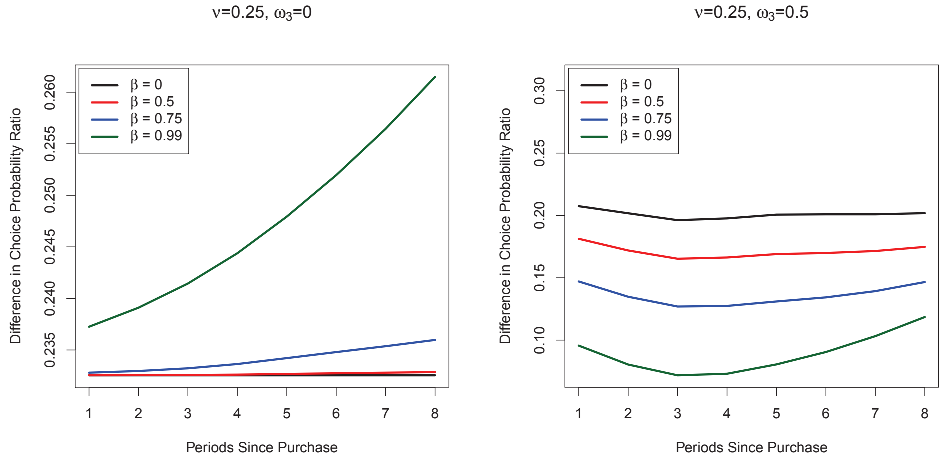

Figure 7: Ratio of probability of buying 2 units to 1 unit at $p = 1$ minus the same ratio at $p = 2$ in period t, given purchase of 1 package in period 0. Parameter values $\nu = 0.25, \pi_c = 0.5, \omega_1 = 0, \omega_2 = 0, \omega_3 = 0$ or $0.5, \eta = 1$, $M = 3$, and logit error term.

not available, and so rather than allowing parameters to be individual specific, the researcher would have to rely on distributional assumptions about the unobserved heterogeneity to aid identification.

7 Artificial Data Experiments

To provide further evidence that the model above can be identified in realistic settings, and to better understand when identification may become more difficult, we perform a series of artificial data experiments. As in the previous section, we perform our analysis on a dataset of 500 households who make purchases over 600 periods. We assume in period 1 everyone starts with 0 inventory; since in real data consumers will likely have been making purchases prior to the beginning of the data collection, we assume that the researcher only observes periods 201 to 600. The estimation method we use is in this section is simulated maximum likelihood. Since initial inventories are unobserved to the researcher, they must be simulated out. The approach we take is to use periods 201 to 400 to simulate initial inventories, and periods 401 to 601 to estimate parameters. In period 201 we assume all consumers begin with zero inventory, and draw a series of consumption shocks for each consumer. With simulated consumption shocks and observed purchase quantities one can construct an estimate of inventory in period 401. Our procedure of using the first part of a sample to construct inventories is standard in the literature (Erdem, Imai, and Keane (2003), Hendel and Nevo (2006a)). We use 100 simulated paths of consumption shocks for each household in constructing the likelihood.

Table 1: Price Transitions Used in Artificial Data Experiment

	$p_t = 0.5$	$p_t = 1$	$p_t = 2$
$p_{t-1} = 0.5$	0.1	0	0.9
$p_{t-1} = 1$	0	0.1	0.9
$p_{t-1} = 2$	0.1	0.1	0.8

To estimate the price coefficient we need sufficient price variation. We allow for 3 prices and use the price transition matrix shown in Table 1 to generate price processes. For the rest of the structure of the model, we allow consumers to purchase 2 packages at most, the package size $b = 8$, consumption shocks are in the set $\{1, 2\}$, and consumers can hold at most 3 packages. The error term is assumed to be logit and the weight on it is set to $\eta = 1$.

The results of the artificial data experiment are shown in Table 2. The top panel shows how the parameter identification is affected by including storage costs and by letting the storage cost function be more flexible. In the first 3 columns of this panel, we estimate the model in a situation where storage costs are zero. The first column shows the estimated parameters, the second the standard errors, and the third is the true values of the parameters. All the parameters are well identified. The next three columns show how the results change if we allow ω_3 to be free, while holding ω_1 and ω_2 fixed at 0. The parameter estimates are still close to the truth, although the standard errors are quite a bit larger. If we allow all 3 storage cost parameters to be positive, and estimate all of them, the standard errors rise significantly, although the parameter estimates are relatively close to the truth.[13] Note that the precision on the discount factor drops as the number of storage cost parameters increases. This highlights the importance of exclusion restrictions. They allow for more precise identification of the discount factor when it is applied to more of the state space: for instance, in cases where it might be reasonable to assume that the cost of storing the first 1 or 2 packages is 0. Turning to the other parameters, the storage cost coefficients are also imprecisely estimated, especially ω_3 - its standard error is three times higher than the estimate from the situation where ω_1 and ω_2 are fixed. This also highlights the fact that assuming zero storage costs for the first few packages may aid identification. It is notable that all the other model parameters, such as α, ν and π_c, are well-identified even if storage costs are flexible.

The bottom panel of the table shows how identification of the discount factor varies as consumers

[13] We do not show the results when we allow for 2 storage costs to be free to save space; in that case the standard errors are a little higher than when we have only a single storage cost free.

Table 2: Artificial Data Experiment: Results

Parameter	No Storage Costs Est	S.E.	Truth	ω_2 Free Est	S.E.	Truth	ω_1, ω_2 Free Est	S.E.	Truth
Price Coeff (α)	1.004	0.007	1	1.002	0.014	1	1.001	0.015	1
Stockout Cost (ν)	0.098	0.01	0.1	0.103	0.006	0.1	0.101	0.036	0.1
Discount Factor (β)	0.957	0.016	0.95	0.957	0.032	0.95	0.954	0.067	0.95
ω_1	-	-	0	-	-	0	0.105	0.059	0.1
ω_2	-	-	0	-	-	0	0.243	0.051	0.25
ω_3	-	-	0	0.499	0.055	0.5	0.508	0.142	0.5
π_c	0.489	0.007	0.5	0.49	4.72e-04	0.5	0.5	0.002	0.5

Parameter	$\beta = 0.001$ Est	S.E.	Truth	$\beta = 0.6$ Est	S.E.	Truth	$\beta = 0.99$ Est	S.E.	Truth
Price Coeff (α)	1.002	0.007	1	1.002	0.009	1	1	0.014	1
Stockout Cost (ν)	0.096	0.022	0.1	0.1	0.011	0.1	0.102	0.006	0.1
Discount Factor (β)	0.001	0.149	0.001	0.619	0.052	0.6	0.994	0.034	0.99
ω_1	-	-	0	-	-	0	-	-	0
ω_2	-	-	0	-	-	0	-	-	0
ω_3	0.479	0.061	0.5	0.47	0.044	0.5	0.488	0.059	0.5
π_c	0.492	0.001	0.5	0.496	2.80e-04	0.5	0.494	4.03e-04	0.5

get more forward-looking. The first column shows the case where consumers are essentially myopic. In this case, the discount factor is not precisely identified. The reason for this is that a consumer with a positive, but low discount factor such behaves very similarly to a myopic consumer. As the discount factor rises, the precision with which we can estimate it also rises.

Before turning to the empirical application, we note that we have also performed our artificial data experiments under the assumption that inventory is discrete rather than continuous. This exercise is presented in Online Appendix **??**. Our findings in that section are similar: we can identify the discount factor well in general.

8 Empirical Application

8.1 Data

To demonstrate how to apply our technique in practice we estimate a stockpiling model using individual level IRI data in the laundry detergent category (Bronnenberg, Kruger, and Mela 2008). An observation in our data is a household-week pair. The data we are currently using covers the years 2001 through 2007. Estimation uses the final 3 years of the data while the first 4 are used to construct initial inventories. In our sample we include households who only purchase the 5 most popular sizes of detergent: the 50 oz, 80 oz size, 100 oz size, the 128 oz size, and the 200 oz size. We restrict the sample to include households who purchase from the top 25 brands by overall purchase share. We also allow consumers to purchase up to 5 bottles units of a size - for instance people will sometimes purchase 2 or 3 bottles of the 100 oz bottle. We remove households who ever purchase different products within the same week (this is very infrequent), or who purchase more than 5 bottles of a product in a week. We only include households who make at least 5 purchases between 2005 and 2007, and for whom the maximum number of weeks between purchases is smaller than 40 weeks. This will cut out households who disappear from the sample for long periods of time, and who may be making laundry detergent purchases that aren't recorded in the data. Additionally, for all households in the data we compute an estimate of the weekly consumption rate and drop individuals for whom the estimated rate is extremely high or low. Extremely low consumption rates likely indicate missing data or purchases made outisde the store sample. Our final sample contains 540 households.

Some statistics on our sample are shown in Table 3. An average household makes a purchase every 10 weeks, and in most weeks no purchase occurs. In our sample, consumers mostly purchase the smallest size bottle containing 100 ounces. Table 4 shows the purchase shares (the number bottles purchased of a particular brand divided by the total number of bottles purchased in the sample) as well as average prices (in cents per ounce) for each brand. (When constructing the sample we initially include the top 25 brands by purchase share. After reducing the sample to 540 households by removing those who purchase too infrequently or who purchase too much, only 18 brands have positive purchases) We are also interested in understanding the relationship between discount factors and demographic variables. Sample averages of these demographic variables are shown in Table 5. We include four demographic variables, all of which are coded as dummy variables. The income variable codes whether the household's income is above $35,000 (the median in our estimation sample), whether the household head's age is about 55 years (also the median household age in the sample), whether the household head has a college degree, and whether the

Table 3: Characteristics of Household Data

Number of households	540
Avg interpurchase time (weeks)	9.9
Fraction of weeks with 0 bottles bought	0.902
Fraction of weeks with 1 bottles bought	0.072
Fraction of weeks with 2 bottles bought	0.018
Fraction of weeks with 3+ bottles bought	0.026
Fraction of purchases where 100 oz size chosen	0.709
Fraction of purchases where 128 oz size chosen	0.146
Fraction of purchases where 200 oz size chosen	0.096
Fraction of purchases where 50 oz size chosen	0.029
Fraction of purchases where 80 oz size chosen	0.02

household has 3 or more individuals in it. The estimation sample somewhat oversamples elderly households and households with 2 individuals, relative to the U.S. population.

8.2 Estimation Details

This section outlines the estimation procedure used to recover the model parameters. Although we argue above that stockpiling models can in principle be identified under relatively flexible specifications of the storage cost function and distribution of consumption shocks, in practice such flexibility can greatly increase the computational burden of estimation. With respect to consumption shocks, we have found that modeling stochastic consumption shocks greatly increases the computational burden of estimation since the shocks need to be integrated out while estimating the other model parameters. Therefore, instead of using simulated maximum likelihood to estimate our model, we use the modified Bayesian MCMC algorithm proposed by Imai, Jain, and Ching (2009).[14] As we use MCMC to estimate our model, we would need to add another Gibbs step to our estimation routine where we draw the consumption shocks for every individual and every period in the data. Moreover, for standard distributions of consumption shocks (such as a normal distribution) the posterior density of the shocks given the data will not have a form that is easy to draw from, necessitating the use of a Metropolis-Hastings step. Adding this step to the algorithm substantially slows

[14] Ching, Imai, Ishihara, and Jain (2012) provides a practitioner's guide to this approach.

Table 4: Brand Level Purchase Shares and Prices

Brand	Purchase Share	Price (Cents Per Ounce)
TIDE	22.6	8.63
XTRA	9.8	2.49
PUREX	10	4.84
ALL	7.1	5.65
ARM & HAMMER	10.4	4.71
ERA	6.1	5.26
DYNAMO	11.5	4.6
WISK	8.9	6.3
PRIVATE LABEL	4	3.53
CHEER	1.5	7.13
FAB	1	6.13
YES	2.4	4.51
AJAX FRESH	0.4	3.2
GAIN	0.5	6.18
AJAX	0.5	3.14
TREND	0.4	2.22
SUN	1	4.44
SOLO	1.7	3.88
IVORY SNOW	0.2	10.46

Table 5: Averages of Demographic Dummy Variables

Variable	Average
HH Income $\geq$ \$35,000	0.6
HH Head Age $\geq$ 55	0.59
HH Head has college degree	0.22
HH Size 3+	0.36

down convergence.[15] As a result, we assume that consumption rates may vary across individuals, but are fixed over time.

The other issue we introduced in the previous paragraph related to the specification of the storage cost function and the way inventory is modeled. Significant computational complications arise in situations where individuals can choose among more than a single size of bottle or brand. In our dataset individuals choose among 5 different package sizes and 19 brands. Adding multiple brands and package sizes to the stylized model will increase in the size of the state space. This is because one would need to: (i) track inventory for each brand, and a price for each brand separately; (ii) track the number of bottles of each size held in inventory, and model the order in which different sizes of bottles are consumed. The inventory composition would matter to the consumer since her storage cost will decrease as she uses up a bottle. A consumer who has two small bottles in her inventory will lower her storage cost more quickly than someone who has two large bottles. An additional complication is that multiple package sizes would require us to model the order in which packages are consumed. For instance, if a consumer has a large bottle and a small bottle in her inventory we would have to decide whether she would use the small bottle before the large one, or vice versa. Below we will first describe how we handle the issues arising from including different bottle sizes, and then from including multiple brands.

First, to deal with the issues arising from multiple bottle sizes, we make a simplification in how we model storage costs. Specifically, we assume that each consumer has an upper bound on the amount that they can store, which we call $\overline{\omega}_i$, and that storage costs are zero before the amount stored hits $\overline{\omega}_i$. An intuitive explanation for this assumption is that a consumer has some storage space dedicated to bottles of laundry detergent, and that she does not purchase more bottles than what she can put in that space. Formally, our formulation of the storage cost is

$$s(I; \omega_i) = \begin{cases} 0 & \text{if } I \leq \omega_i \\ \infty & \text{otherwise} \end{cases}$$

With this formulation, we do not have to model the composition of bottles in inventory or the order in which bottles are used. Even though we have put inventory directly into the storage cost function for convenience, the exclusion restriction still holds since storage costs are zero until the bound is reached, and a consumer will never purchase more than she can store. The storage cost bound $\overline{\omega}_i$ is an individual specific parameter we estimate.

Second, to deal with issues arising from including brand differentiation, we follow Hendel and

[15] Some earlier empirical work on stockpiling such as Hendel and Nevo (2006a) and Sun (2005) has included stochastic consumption shocks, but those papers did not model unobserved preference heterogeneity as we do.

Nevo (2006a) and make two simplifying assumptions: (i) consumers only care about brand differentiation at the time of purchase, and (ii) a form of inclusive value sufficiency modified from what Hendel and Nevo (2006a) proposed (the modifications we use were introduced in Osborne (2017)). Assumption (i) means that all utility from consuming a particular brand arises when a consumer makes a purchase, and at the time of consumption only the overall level of inventory matters.[16] This implies that the composition of the inventory does not matter, and it drastically reduces the size of the state space. We assume that the flow utility received from a particular brand scales linearly with the number of packages purchased: the flow utility from purchasing j packages of brand k is equal to $\frac{j}{J}\xi_{ik}$, where one of the ξ_{ik} coefficients is normalized to zero.[17] The assumption that brand utility scales with the number of packages purchased relates the inclusive value sufficiency assumption (ii), and we will show below that it will help reduce the size of the model's state space. The consumer's flow utility function from buying $j > 0$ units of size x of brand k can be written down as:

$$u_{it}(k, x, j, I_{it}, \varepsilon_{ijt}, p_{it}, c_i; \boldsymbol{\theta}_i) = \begin{cases} \frac{j}{J}\xi_{ik} - s(B_{i,t+1}(j, I_{it}, c_i); \boldsymbol{\omega}_i) - \alpha_i p_{ixkt} j + \varepsilon_{ijxkt} & \text{if } I_{it} + b(x)j \geq c_i \\ \frac{j}{J}\xi_{ik} - \nu_i \frac{c_{it} - (I_{it} + b(x)j)}{c_i} - \alpha_i p_{ixkt} j + \varepsilon_{ijxkt} & \text{otherwise} \end{cases}, \tag{18}$$

where $b(x)$ is the number of ounces in a bottle of size x. Before we write down the Bellman equation we need to clarify the elements of the consumer's choice set. Consumers can either purchase nothing ($j = 0$), or purchase $j > 0$ units of a single brand-size combination (k, x). Denoting the feasible set of (j, k, x) combinations as C, the consumer value function can be written as:

$$V_{it}(I_{it}, \boldsymbol{p}_{it}) = E_{\varepsilon_{it}} \max_{(j,k,x) \in C} \{u_{it}(k, x, j, I_{it}, \varepsilon_{ijkt}, \boldsymbol{p}_{it}, c_i; \boldsymbol{\theta}_i) + \beta_i E_{p_{i,t+1}|p_{it}} V(I_{i,t+1}, p_{i,t+1})\}, \tag{19}$$

where p_{it} is a vector of brand-size level prices. Our second assumption of inclusive value sufficiency (IVS) simplifies the state space by assuming that rather than tracking individual prices, consumers track the expected flow utility arising from each available package size, which are the inclusive values. The standard formulation of IVS used in Hendel and Nevo (2006a) relies on the assumption

[16]Formally, assumption (i) means that the consumption utility, γ_i, does not depend on the brand purchased (as we argued earlier, the parameter γ_i is not identified so we normalize $\gamma_i = 0$).

[17]We have made a restriction that the flow utility for the brand does not depend on the particular package size chosen - ξ is not indexed by x. We experimented with allowing the ξ_{ik} coefficients to vary across sizes, x, but found it difficult to identify brand-package size interactions.

of logit errors, and under their formulation the number of inclusive values equals the number of packages multiplied by the number of packages an individual can purchase. Osborne (2017) shows that the number of inclusive values one needs to track can be further reduced to only the number of packages, under the assumption that flow utility scales with the number of packages purchased, coupled with an assumption that the choice specific error can be written in the form of a nested logit:

$$\varepsilon_{ijxkt} = e_{ijt} + \frac{j}{J} v_{ixkt},$$

where the distribution of v_{ixkt} is Type 1 extreme value and the distribution of e_{ijt} has a distribution of the form denoted as $C(\lambda)$ from Cardell (1997), where $\lambda = j/J$ (note the $C(\lambda)$ notation does not refer to choice sets, but to a particular distribution derived in Cardell (1997)). Denoting C_1 as the set of feasible (j, x) combinations and $C_2(x)$ as the set of brands which are available in size x, the two aforementioned assumptions entail that an individual's expected utility over brands for choosing j packages of size x can be written as

$$\frac{j}{J}\Omega_{it}(x) = \frac{j}{J} \ln \left(\sum_{k \in C_2(x)} \exp\left(\xi_{ixk} - J\alpha_i p_{ixkt}\right) \right).$$

Details on the above derivation are shown in Online Appendix 14. To summarize, our implementation of IVS assumes that consumers track $\Omega_{it}(x)$, rather than each individual price p_{ixkt}. As a result, the Bellman equation in equation (19) can be written as

$$V(I_{it}, \mathbf{\Omega}_{it}) = \ln \left(\sum_{(j,x) \in C_1} \exp\{ \frac{j}{J}\Omega_{it}(x) - \nu_i \frac{c_i - (I_{it} + bj)}{c_{it}} \mathbf{1}\{I_{it} < c_i\} + \beta_i E_{\mathbf{\Omega}_{it}|\mathbf{\Omega}_{i,t-1}} V(I_{i,t+1}, \mathbf{\Omega}_{it})\} \right), \quad (20)$$

where $\mathbf{\Omega}_{it}$ is an X-dimensional vector of inclusive values for all package sizes.

To estimate the model we use the Bayesian estimation method of Imai, Jain, and Ching (2009) (henceforth abbreviated IJC). Hendel and Nevo (2006a) propose a three step estimation method that uses maximum likelihood, but their approah cannot allow for unobserved heterogeneity across individuals. The IJC method can more easily handle unobserved heterogeneity than the standard approach since one does not have to solve the value function repeatedly - rather one iterates on the value function over the course of the MCMC chain making solution much faster.

In addition to the different specification used for storage cost, we make three other more minor changes to the model specification from the specification used for artificial data experiments. First, we incorporate a fixed cost of purchase, FC_i, which is the disutility a consumer receives from making a purchase. We found it necessary to include this parameter in order to properly fit the low frequency of purchase we observe in the data.

Second, rather than estimating consumption rates we calibrate them from the data. Consistent with what we note earlier, we found it difficult to identify both the consumption rate and the stockout cost together, and this problem seemed especially pronounced when the discount factor was low. Thus, we set each individual's consumption rate to the total quantity purchased over the estimation period, divided by the total number of weeks where the individual is observed. To ensure our results are not materially affected by this assumption, we perform a robustness exercise where we increase every individual's consumption rate by 25% and re-estimate the model.[18] We find our parameter estimates are relatively insensitive to the consumption rate.

The third change relates to store visits. In the data, there are some weeks where consumers do not visit any store. To capture this, the third change we have made is that we assume there is an exogenous probability a consumer goes to the store, which we estimate prior to estimating the other model parameters. This probability is incorporated into consumers' expectations when they update their value functions in equation (20).

For simplicity of our exposition below, we outline how the solution of the model works when it is assumed that consumers always visit a store. We found it difficult to identify unobserved heterogeneity in brand coefficients, so we assume that all those parameters are fixed across the population. However, we do allow for unobserved heterogeneity as well as demographic interactions in the price coefficient, the cost of stocking out, the discount factor, the fixed cost of purchase, and the upper bound on storage.

We allow for unobserved heterogeneity in all other model parameters except for two of the brand coefficients - we found we could not identify all the variances of all the brand coefficients. The basic steps of the algorithm are as follows:

1. Draw the population-varying parameters using Metropolis-Hastings,

2. draw the means and of population-varying parameters,

3. draw the variance of population-varying parameters,

4. draw the population-fixed parameters using Metropolis-Hastings, and

[18] If an individual always purchases more of a product at the time she runs out, which we might expect with necessities such as laundry detergent, the calibrated consumption rate will equal the underlying consumption rate. We have found in simulations that if stockout costs are low enough that individuals sometimes wait a few periods after running out to make a purchase, the calibrated rate somewhate understates the actual consumption rate.

5. update the value function.

We describe how we implement steps 1 to 4 in Online Appendix 15.1, and step 5 in Online Appendix 15.2. Some other details related to the construction of the inclusive value transition process and setup of the MCMC chain are described in Online Appendices 15.3 and 15.4.

8.3 Estimation Results

This section presents our estimation results. Table 6 shows the estimates of the brand parameters. Since none of these parameters vary across the population, we present the posterior mean of the 6,000 saved draws as well as the 95% confidence interval around the estimated mean.[19] Table 7 shows the estimates of the non-brand model parameters: the price coefficient, stockout cost, discount factor, fixed cost of purchase, and inventory bound. Since all of these parameters vary across the population, we show the 25th, 50th, and 75th percentiles as well as the population mean. To compute an estimated moment, say the 25th percentile, first for each Gibbs draw we compute the 25th percentile of the population distribution of taste draws for the price coefficient. The estimated 25th percentile is the average of the 25th percentiles over all 6,000 saved Gibbs draws. The second row shows the 95% confidence bounds on each of the estimated moments. There is a significant amount of heterogeneity in all of these model parameters. Stockout costs, as well as the fixed cost of purchase, are all large in utility terms.[20] Turning to the discount factor, the population average of the weekly discount factor is about 0.73, which is much lower than the value of 0.995 that would be consistent with rational expecations assumption. There is also some heterogeneity across individuals in discount factors. This can be seen in Figure 8, where we plot a kernel density of the average estimated discount factor for the population (for each individual, we compute the average of the discount factor estimate for all saved draws). Most individuals' discount factors lie between about 0.6 and 0.85. Although our estimated discount factors are less than the rational expectations benchmark assumed in past work, low estimated discount factors are consistent with some other field studies that allow the parameter to be free (for example, Yao, Mela, Chiang, and Chen (2012) estimate in data on cellular phone usage that consumer discount factors are around 0.91).

Taking our results at face value it may be tempting to argue that our estimates suggest con-

[19] In Figure 9 of Online Appendix 10 plots the estimated mean parameter $\boldsymbol{b}$ at each of the 10,000 Gibbs steps for selected parameters. The parameters seem to stabilize at or before draw 4,000.

[20] The inventory bound is presented in hundreds of ounces, meaning that individuals have a large amount of free space for laundry detergent.

sumers are irrational, as weekly a discount factor of 0.73 would translate to a yearly discount factor of close to 0, implying consumers are essentially myopic when making financial decisions where the time horizon was on the order of a year. With a discount factor at this range, it essentially implies that consumers make their purchase decisions thinking a couple of weeks ahead. Since consumer package goods are small ticket items, such a short planning horizon may be reasonable and could be rationalized as rational behavior taking into account sarced mental resources. When making important financial decisions, consumers may behave in a more forward-looking way due to the fact that more money is on the line - there are more gains to plan for a longer horizon and hence it is worth exerting more mental resource to think further ahead.

Table 8 shows the estimated marginal impact of each demographic variable on all the parameters which are allowed to vary across the population. The first column of this table shows the estimated parameter at the modal value of the demographic variables. For example, if we denote the modal demographics as $\overline{\boldsymbol{Z}}$ then for the price coefficient we show in the first column $\exp(\boldsymbol{b}_k'\overline{\boldsymbol{Z}})$, where k is the row of $\boldsymbol{b}$ corresponding to the untransformed price coefficient. The table shows the value averaged across saved draws. Each column shows how the estimated parameter changes when the corresponding demographic variable is changed from zero to one. For example, if a household is high income its price coefficient is closer to zero by 0.009 (i.e., higher income households are more price sensitive). The results indicate that higher income households, older households, and larger households are more forward-looking, although the overall effects of the demographics are not large when compared to the amount of unobserved heterogeneity. This finding suggests that most of the heterogeneity in discount factors seems to be driven by unobserved factors is consistent with the results of Dubé, Hitsch, and Jindal (2014).

We also estimated two alternative model specifications as a robustness check. In one specification, we increase all consumption rates by 25% and estimate the full model. We find that our parameter estimates are similar to the basic specification. In particular the estimated discount factors are a little bit higher, as the average discount factor is about 0.79. We also estimate a version of the model where we fix the discount factor to be zero across the population, to verify we are indeed able to tell apart forward-looking from myopic behavior. The forward-looking model fits the data better, having a Deviance Information Criterion of 82967.69, while that of the myopic model is higher at 83132.69. The average marginal likelihood of the forward-looking model is -40988.16 while the myopic model is -41051.3. A complete set of results from these specifications are available from the authors upon request.

Table 6: Brand Parameter Estimates

Parameter	Average Posterior Estimates	95% CI
XTRA	-4.32	[-4.46, -4.17]
PUREX	-3.91	[-4.02, -3.79]
ALL	-2.92	[-3.01, -2.8]
ARM & HAMMER	-3.13	[-3.25, -3.01]
ERA	-3.27	[-3.38, -3.16]
DYNAMO	-3.52	[-3.67, -3.35]
WISK	-2.32	[-2.43, -2.21]
PRIVATE LABEL	-5.66	[-5.81, -5.47]
CHEER	-2.64	[-2.84, -2.43]
FAB	-4.4	[-4.72, -4.12]
YES	-4.94	[-5.21, -4.68]
AJAX FRESH	-6.24	[-6.65, -5.87]
GAIN	-4.64	[-4.98, -4.31]
AJAX	-6.23	[-6.6, -5.86]
TREND	-6.6	[-6.98, -6.24]
SUN	-5.04	[-5.46, -4.64]
SOLO	-6.11	[-6.41, -5.82]
IVORY SNOW	-3.8	[-4.32, -3.33]

Notes: The first column of the table show the average of the estimated posterior distribution of the brand parameters. The second shows the 95% confidence bound around the mean. Brand coefficients for Tide (the most popular product) are normalized to be zero across the population.

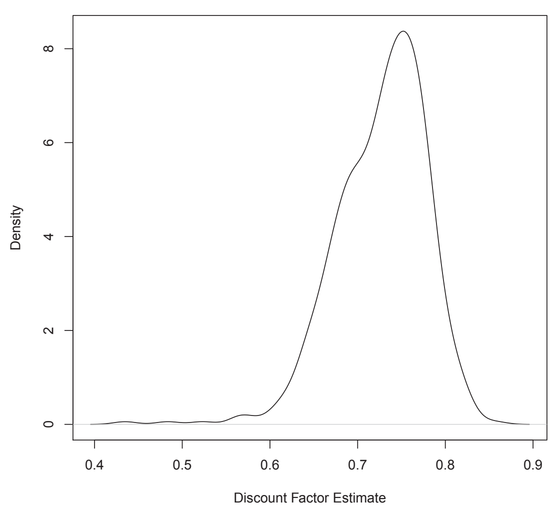

Figure 8: Kernel Density of Individual-Specific Discount Factor Estimates.

Table 7: Dynamic Parameter Estimates

Parameter	1st Quartile	Median	Mean	3rd Quartile
Price Coefficient	-0.23	-0.16	-0.19	-0.11
	[-0.25, -0.22]	[-0.17, -0.16]	[-0.2, -0.19]	[-0.11, -0.1]
Stockout Cost	0.28	0.39	0.44	0.53
	[0.24, 0.32]	[0.34, 0.43]	[0.39, 0.49]	[0.46, 0.61]
Discount Factor	0.67	0.74	0.73	0.8
	[0.59, 0.72]	[0.66, 0.78]	[0.65, 0.77]	[0.73, 0.83]
Fixed Cost of Purchase	-4.09	-3.38	-3.36	-2.68
	[-4.19, -4]	[-3.46, -3.3]	[-3.42, -3.3]	[-2.77, -2.6]
Inventory Bound	12.04	15.57	16.07	19.59
	[9.32, 16.46]	[13.04, 19.87]	[13.58, 19.98]	[16.07, 23.55]

Notes: This table shows average moments of the posterior distribution of the population distribution of the dynamic parameters. For example, the median columns shows the average of the population median of a given parameter, where the average is taken across MCMC draws. Square brackets show 95% confidence intervals.

9 Conclusion

We note that our strategy to identify the consumer discount factor will work well for many, but not all product categories. Product categories which fit our framework should have three key features. First, they should be product categories where a consumer does not gain from consuming beyond weekly needs. Products such as laundry detergent, ketchup, instant or ground coffee will fit this criterion well. One does not gain utility from consuming more laundry detergent than what is necessary to do laundry, or more ketchup than what is necessary to put on a hamburger. Products where temptation is a large part of purchase, such as ice cream or potato chips, may not "provide" exclusion restrictions. The reason for this is that the more of the product one has in inventory, the more one is tempted to consume the product, and the more one gains in current utility. The second key feature is that the cost of storing a product (in terms of space used) does not in general change as inventory drops. This feature will exist in product categories where a product's package size does not decrease with inventory - outside of rare instances where one has multiple packages and a package is used up, the space taken up by packages won't change as the amount in a package changes. For products such as laundry detergent, where the product is a liquid

Table 8: Marginal Effects of Demographic Variables

Parameter	Baseline	HH Income	HH Head Age	HH Head College	HH Size
Price Coefficient	-0.174	0.009	0.005	0.003	0.004
		[0.002, 0.015]	[-0.002, 0.011]	[-0.004, 0.009]	[-0.002, 0.011]
Stockout Cost	0.579	-0.087	-0.091	-0.027	-0.049
		[-0.117, -0.058]	[-0.121, -0.061]	[-0.055, 0.001]	[-0.075, -0.022]
Discount Factor	0.646	0.036	0.038	0.012	0.021
		[0.021, 0.051]	[0.022, 0.053]	[-0.001, 0.025]	[0.008, 0.034]
Fixed Cost of Purchase	-1.945	-0.502	-0.516	-0.172	-0.305
		[-0.651, -0.351]	[-0.668, -0.367]	[-0.331, -0.012]	[-0.457, -0.153]
Inventory Bound	18.98	-1.228	-1.331	-0.418	-0.765
		[-1.959, -0.431]	[-2.074, -0.505]	[-1, 0.12]	[-1.397, -0.159]

Notes: This table shows the estimated impact of changing one of the demographic dummy variables from zero to one on a particular parameter. The respective demographic dummy variables are defined to be 1 under the following conditions: Income above $35,000; age of household head above 55; household head has a college degree; size of household is more than 2 individuals. The baseline column shows the predicted value of a parameter at the mode of the demographic distribution. The modal values are high income, older household head, no college degree, and two individuals in the household.

stored in bottles, this assumption will hold: the cost of storing the product only depends on the number of bottles held, but not the amount of inventory within a bottle. If this were not the case, we would not have exclusion restrictions because inventory would affect storage costs continuously, which are part of a consumer's current period payoffs. Product categories we think would work well with our identification strategy include laundry detergent, ketchup, cereal, deodorant, facial tissue, household cleaners, mustard, mayonnaise, margarine, peanut butter, or shampoo.

The third key feature is that the consumption rate is small relatively to package size, such that it takes consumers several periods to use up a package. As we have argued throughout the paper, both the slope and curvature of the purchase hazard will help identify the discount factor. If consumers use up a package of the product very quickly, then the purchase hazard will have little to no curvature, and it will provide less information about how forward-looking consumers are.

Consumer stockpiling behavior in consumer package goods is often cited as an example where consumers are forward-looking. However, previous research (most notably, Erdem, Imai, and Keane

(2003), Hendel and Nevo (2006a)) assumes (i) consumer are homogeneous in their discount factors, and (ii) consumers do not arbitrage and hence discount factor can be set according to the prevailing interest rate. By explicitly modeling storage costs to depend on number packages instead of inventory, our model generates exclusion restrictions that have not been previously studied. By exploring these exclusion restrictions and using recently developed estimation methods, we are able to relax these two assumptions. To classical economists, our findings may be surprising because consumers are not only heterogeneous in their discount factors, but their magnitudes are also significantly lower than what the interest rate predicts. Our estimated weekly discount factors average at around 0.69, lower than the value of 0.99 this is obtained if one uses a common interest rate to set it. The differences are large and they could lead to material impact on the results of counterfactual experiments conducted in prior research which fixes consumer discount factors.

For instance, Erdem, Imai, and Keane (2003) quantify the importance of consumer expectations in their response to promotions using a dynamic structural model of purchases in the ketchup category. The modeling approach in that paper has many similarities to ours, but the discount factor is fixed to the rational expectations benchmark. Their paper finds that long run cross-price elasticities are much greater than short run cross-price elasticities, and that temporary price reductions drive category expansion rather than switching between brands. If the actual discount factor is smaller than the calibrated value, then long run and short run elasticities cross-elasticities should be closer to each other. As a result, if a manager were to incorrectly assume a value of the discount factor that was too high, she would choose an optimal price that was too low and react too much to competitor price changes. The paper also quantifies the impact on firm profits of switching from a Hi-Lo pricing strategy to an Everyday Low Price Strategy, and finds for a leading ketchup brand that such a change in strategy can increase profits. The increase in profits may be driven by the fact that under Hi-Lo pricing, some of the current period increase in quantity arising from a temporary price discount steals from future demand. If individuals are myopic they will be less likely to stockpile at low prices, which will dampen this effect and may make Hi-Lo pricing more profitable.

An important factor that drives the profitability of Hi-Lo pricing is heterogeneity in the ability of consumers to stockpile (Hendel and Nevo (2013), Hong, McAfee, and Nayyar (2002)). Hi-Lo pricing is a form of intertemporal price discrimination: price sensitive individuals will also tend to wait for promotions, while price insensitive individuals will not stockpile and will usually be charged the high price. Hendel and Nevo (2013) empirically examine the implications of banning intertemporal price discrimination on welfare and profits, and find that intertemporal price discrimination is profitable. In their paper, the discount factor is also fixed and the ability of consumers to stockpile is driven by differences in storage cost. Our paper allows for another dimension of heterogeneity: differences across individuals with respect to the discount factor. Although we find evidence of heterogeneity in individual discount factors, we also find most individuals are relatively myopic. As a result, studies which fix the discount factor at a high value mightl overstate the profitability of a Hi-Lo pricing strategy.

Although our approach relies on a fixed consumption rate, in situations where consumption is endogenous a price promotion will lead to both increased current consumption, as well as stockpiling, and it is managerially relevant to be able to separate the two drivers of the increase in purchase. For example, Sun (2005) finds in two product categoryies that a temporary 25% price discount will substantially increase consumption for two or three weeks after prices return to baseline levels. If consumers are less myopic than the rational expectations baseline, then most of the consumption increase arising from a temporary price discount will occur during the period when the discount occurs.

Our results also have policy relevance, as Hendel and Nevo (2006a) argue that ignoring consumers' forward-looking incentives would lead to overestimation of price-cost margins, underestimation of cross-price elasticities, and overestimates of the amount of substitution to the outside alternative. The latter two findings imply that using estimates from a static demand model could lead to misleading policy decisions in approving mergers (an antitrust authority that relied on static demand estimates would be too lenient). However, our findings suggest that the standard practice of setting the discount factor using the prevailing interest rate could generate the opposite outcome, i.e., price-cost margins would be underestimated, and hence merger decisions would be made too conservatively. In particular, it is common for researchers to use price-cost margins to test whether firms collude. Using a discount factor that is too high would increase the incidence of type I errors, i.e., rejecting the collusion hypothesis when it is indeed happening.

References

Abbring, J. and O. Daljord (2016). Identifying the discount factor in dynamic discrete choice models. Working Paper.

Akça, S. and T. Otter (2015). Identifying the discount factor of forward looking consumers based on consumption from inventory. Working Paper.

Bronnenberg, B., M. Kruger, and C. Mela (2008). The IRI marketing data set. Marketing Science 27(4), 745–748.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. Econometric Theory 13(2), 185–213.

Chevalier, J. and A. Goolsbee (2009). Are durable-goods consumers forward-looking? evidence from college textbooks. Quarterly Journal of Economics 124(4), 1854–1884.

Ching, A., T. Erdem, and M. P. Keane (2014). A simple method to estimate the roles of learning, inventories, and category consideration in consumer choice. Journal of Choice Modelling 13, 60–72.

Ching, A., S. Imai, M. Ishihara, and N. Jain (2012). A practitioner's guide to bayesian estimation of discrete choice dynamic programming models. Quantitative Marketing and Economics 10(2), 151–196.

Chung, D., T. Steenburgh, and K. Sudhir (2013). Do bonuses enhance sales productivity? a dynamic structural analysis of bonus-based compensation plans. Marketing Science 33(2), 165–187.

Coase, R. H. (1972). Durability and monopoly. Journal of Law and Economics 15, 143–149.

Crawford, G. and M. Shum (2005, July). Uncertainty and learning in pharmaceutical demand. Econometrica 37, 1137–1174.

Dubé, J.-P., G. Hitsch, and P. Jindal (2014, December). The joint identification of utility and discount functions from stated choice data: An application to durable goods adoption. Quantitative Marketing and Economics 12(4), 331–377.

Erdem, T., S. Imai, and M. Keane (2003). A model of consumer brand and quantity choice dynamics under price uncertainty. Quantitative Marketing and Economics 1(1), 5–64.

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

Fang, H. and Y. Wang (2015). Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. International Economic Review.

Feenstra, R. and M. Shapiro (2003, August). High-Frequency Substitution and the Measurement of Price Indexes. Univ. of Chicago and NBER. in Robert Feenstra and Matthew Shapiro (eds.), *Scanner Data and Price Indexes.*

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002). Time discounting and time preference: A critical review. Journal of Eocnomic Literature 15, 351–401.

Geweke, J. and M. Keane (2000). Bayesian inference for dynamic choice models without the need for dynamic programming. Cambridge University Press. in Roberto Mariano, Til Schuermann, Melvyn J. Weeks (eds.), *Simulation-based Inference in Econometrics: Methods and Applications.*

Haviv, A. (2014). Does purchase without search explain counter cyclic pricing? working paper, Simon Business School, University of Rochester.

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

Hendel, I. and A. Nevo (2006b). Sales and consumer inventory. The RAND Journal of Economics 37, 543–561.

Hendel, I. and A. Nevo (2013). Intertemporal price discrimination in storable goods markets. American Economic Review 103(7), 2722–2751.

Hong, P., P. McAfee, and A. Nayyar (2002). Equilibrium price dispersion with consumer inventories. Journal of Economic Theory 105, 503–517.

Hotz, J. and R. Miller (1993, July). Conditional choice probabilities and the estimation of dynamic models. The Review of Economic Studies 60(3), 497–529.

Imai, S., N. Jain, and A. Ching (2009). Bayesian estimation of dynamic discrete choice models. Econometrica 77(6), 1865–1899.

Ishihara, M. and A. Ching (2012). Dynamic demand for new and used durable goods without physical depreciation: The case of japanese video games. Working Paper.

Lee, R. S. (2013). Vertical integration and exclusivity in platform and two-sided markets. American Economic Review 103(6), 2960–3000.

Liu, Y. and S. Balachander (2014). How long has it been since the last deal? consumer promotion timing expectations and promotional response. Quantitative Marketing and Economics 12(1), 85–128.

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete decision processes. Econometrica 20(2), 801–816.

McFadden, D. (1981). Econometric Models of Probabilistic Choice. MIT Press: Cambridge, MA. in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Application.*

Norets, A. (2009). Inference in dynamic discrete choice models with serially correlated unobserved state variables. Econometrica 77, 1665–1682.

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

Osborne, M. (2017). Approximating the cost-of-living index for a storable good. Working Paper.

Pesendorfer, M. (2002). Retail sales: A study of pricing behavior in supermarkets. Journal of Business 75(1), 33–66.

Reis, R. (2009). A dynamic measure of inflation. Working Paper.

Rossi, P., G. Allenby, and R. McCulloch (2005). Bayesian Statistics and Marketing. Wiley.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurchner. Econometrica 55, 993–1033.

Seiler, S. (2013). The impact of search costs on consumer behavior: A dynamic approach. Quantitative Marketing and Economics 11(2), 155–203.

Sobel, J. (1984). The timing of sales. Review of Economic Studies 51, 353–368.

Sun, B. (2005). Promotion effect on endogenous consumption. Marketing Science 24, 430–443.

Yang, B. and A. Ching (2014). Dynamics of consumer adoption of financial innovation: The case of atm cards. Management Science 60(4), 903–922.

Yao, S., C. Mela, J. Chiang, and Y. Chen (2012). Determining consumers' discount rates with field studies. Journal of Marketing Research 49(6), 822–841.