A Simple Model of Demand Anticipation

Igal Hendel   Aviv Nevo

January 26, 2010

Abstract

In the presence of intertemporal substitution, static demand estimation yields biased estimates and fails to recover long run price responses. Our goal is to present a computationally simple way to estimate dynamic demand using aggregate data. Previous work on demand dynamics is computationally intensive and relies on (hard to obtain) household level data. We estimate the model using store level data on soft drinks and find: (i) a disparity between static and long run estimates of price responses, and (ii) heterogeneity consistent with sales being driven by discrimination motives. The model's simplicity allows us to compute mark-ups implied by dynamic pricing.

1 Introduction

Demand estimation plays a key role in many applied fields. A typical exercise is to estimate a demand system and use it to infer conduct, simulate the effects of a merger, evaluate a trade policy or compute cost pass-through.[1] While for the most part the demand models used are static, there is evidence that product durability or storability may generate dynamics, which could contaminate estimates. Focusing on storable products, a number of papers (Erdem, Imai and Keane, 2003, and Hendel and Nevo, 2006b) use household level data to structurally estimate consumer inventory models and simulate long run price responses. The computational burden and (household level) data requirement have limited the use of these dynamic demand models.

We propose an alternative model to incorporate demand dynamics. Our goal is to present a computationally simple way to estimate dynamic demand for storable products, or test for its presence, using aggregate, rather than household level, data. In many studies dynamics are not the essence. A test for the presence of dynamics may help rule them out. If dynamics are present their impact can be quantified by comparing static estimates to estimates from our model.

The model allows us to separate purchases for current consumption from purchases for future consumption. That way we can relate consumption and prices, to recover preferences (clean of storage decisions); and translate short run responses to prices, observed in the data, into long run reactions. The latter are the object of interest in most applications. The way we impute purchases for storage is quite simple but intuitive. Its advantage is that it does not require solving the value function of the consumer and the estimation is straightforward.

A key to the simplicity of the model is in the storage technology: consumers are assumed to be able to store for a pre-specified number of periods. This assumption simplifies the solution to the consumer's problem. The intuition of the model can best be demonstrated by a simple example. Suppose there is a single variety of a product with (1) prices that take on two values: a sale and a non-sale price; and (2) some consumers can store the product for one period (while others cannot store). Given these assumptions the model defines four states depending on the current and previous period price. The states determine whether there are purchases for storage or not, and whether consumption comes out of storage. Thus,

[1] See, for example, Berry, Levinsohn, and Pakes (1995, 1999), Goldberg (1995), Hausman, Leonard and Zona (1994).

than the static estimates. The order of magnitude of the bias is comparable to what Hendel and Nevo (2006b) …nd when they estimate a dynamic inventory model for laundry detergents.

We discuss alternative approaches in Section 7. Alternatives to dealing with dynamics include aggregating the data from weekly to monthly and quarterly frequency, or approximating the missing inventory by including lagged prices/quantities (and computing long run e¤ects using impulse response). We show these alternatives perform poorly, yielding negative cross price e¤ects. We argue that the alternative methods also require a model to translate the estimated coe¢ cients into preferences.

Another advantage of the simplicity of the model is to make the supply side tractable. In principle, the presence of demand dynamics makes the pricing problem quite di¢ cult to solve. Especially so when there are multiple products sold by di¤erent sellers. In contrast, the demand framework we propose leads to a simple solution to the sellers'pricing problem.

Studying the supply side is interesting in its own right, but it is particularly important in many applications. Demand elasticities are typically used in conjunction with static …rst order conditions to infer market power. Demand dynamics render static …rst order conditions irrelevant. A supply framework consistent with demand dynamics is needed. We show that sellers'optimal behavior can still be characterized by …rst order conditions. Interestingly, the demand estimates show that consumers who store are signi…cantly more price sensitive than non-storers, which is consistent with price discrimination being the motive behind sales. We use the estimated demand elasticities and the dynamic …rst order conditions to infer markups.

Section 2 presents motivating facts and reviews the literature. The model is presented in Section 3 and the estimation in Section 4. Section 5 presents an application to soft drinks. Extensions of the model are presented in Section 6.

2 Evidence of Demand Accumulation

2.1 Motivating Facts

Several papers (discussed in the next sub-section) have documented demand dynamics. We …rst look at typical scanner data for direct evidence on the relevance of intertemporal demand e¤ects.

Figure 1 shows the price of a 2-liter bottle of Coke in a store over a year. The pattern is typical of pricing observed in scanner data: regular prices and occasional sales, with return to the regular price. Since soft-drinks are storable, pricing like this creates an incentive for consumers to anticipate purchases: buy during a sale for future consumption.

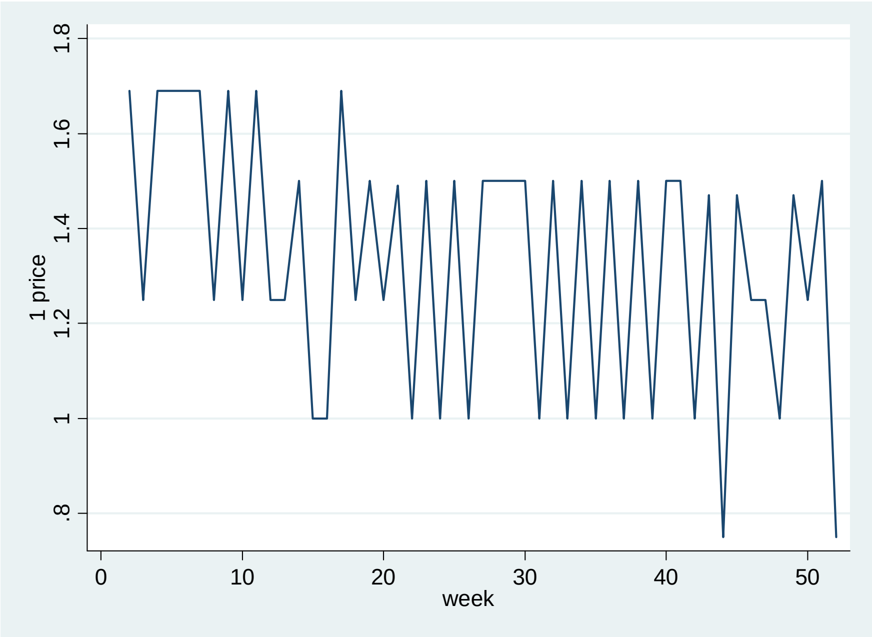

Figure 1: A typical pricing pattern

Quantity purchased shows evidence of demand accumulation. Table 1 displays the quantity of 2-liter bottles of Coke sold during sale and non-sale periods (we present the data in more detail below). During sales the quantity sold is signi…cantly higher (623 versus 227, or 2.75 times more). More importantly, the quantity sold is lower if a sale was held in the previous week (399 versus 465, or 15 percent lower).

The impact of previous sales is even larger if we condition on whether or not there is a sale in the current period (532 versus 763, or 30 percent lower, if there is a sale and 199 versus 248, or 20 percent lower in non sale periods).

We interpret the simple patterns present in Table 1 as evidence that demand dynamics are important and that consumers' ability to store detaches consumption from purchases. Table 1 shows that purchases are linked to previous purchases, or at least, to previous prices.

Table 1: Quantity of 2-Liter Bottles of Coke Sold

	$S_{t-} = 0$	$S_{t-} = 1$	
$S_t = 0$	247.8	199.4	227.0
$S_t = 1$	763.4	531.9	622.6
	465.0	398.9	

Note: The table presents the average across 52 weeks and 729 stores of the number of 2-litter bottles of Coke sold during each period. As motivated below, a sale is defined as any price below 1 dollar.

2.2 Related Literature

Numerous papers in Economics and Marketing document demand dynamics, specifically, demand accumulation (see Blattberg and Neslin (1990) for a survey of the Marketing literature). Boizot et al. (2001) and Pesendorfer (2002) show that demand increases in the duration from previous sales. Hendel and Nevo (2006a) document demand accumulation and demand anticipation effects, namely, duration from previous purchase is shorter during sales, while duration to following purchase is longer for sale periods. Erdem, Imai and Keane (2003), and Hendel and Nevo (2006b) estimate structural models of consumer inventory behavior.

Several explanations have been proposed in the literature to why sellers offer temporary discounts. Varian (1980) and Salop and Stiglitz (1982) propose search based explanations which deliver mixed strategy equilibria, interpreted as sales. Sobel (1984), Conlisk Gerstner and Sobel (1984), Pesendorfer (2002), Narasimhan and Jeuland (1985) and Hong, McAfee and Nayyar (2002) present different models of intertemporal price discrimination. Our estimates show that sellers have incentives to intertemporally price discriminate, suggesting that sales are probably driven by discrimination motives.

3 The Model

In order to convey the main ideas we start with the simplest model of product differentiation with storage. We later show the model can be generalized in several dimensions. For example, the proposed estimation can be applied to more flexible demand systems, e.g., Berry, Levinsohn and Pakes (1995).

3.1 The Main Assumptions

Assume quadratic preferences:

$$U(q; m) = Aq \quad q'Bq + m \tag{1}$$

where $q = [q\ ; q\ ; \ldots ; q_N]$ is the vector of quantities consumed of the different varieties of the product (colas in our application) and m is the outside good. Absent storage, quadratic preferences lead to a linear demand system:

$$q_i^t(p) = \quad {}_i p_i^t + \sum_j \quad {}_{ij} p_j^t \tag{2}$$

In a multi-period set up with storage, consumers can anticipate purchases for future consumption. We make the following assumptions:

A1:

If everyone stored in the previous period our model would predict no purchases. With two prices this assumption is not very restrictive, but as we add more prices it will have bite since it assumes that the fraction of non-storers does not change with price.

In Section 3.4 we discuss the assumptions, their limitations, and possible generalizations.

3.2 Purchasing Patterns

We now characterize consumer behavior. To ease exposition we ignore discounting. The application involves weekly data, and therefore discounting does not play a big role.

Consumers who store, purchase for storage at p^S, and never store at p^N: When they store, they do so for one period. Thus, to predict consumer behavior we only need to define 4 events (or types of periods): a sale preceded by a sale (SS), a sale preceded by a non-sale (NS), a non-sale preceded by a sale (SN), and two non-sale periods (NN). We assume for now perfect price foresight, and later discuss (in section 6) behavior under rational price expectations.

current consumption comes from stored units, so purchases are for future consumption only, and the contribution to aggregate demand is $(1 - \omega) q_i(p_i^t; p_{-i}^t)$:[2]

Notice the difference in the second argument of the anticipated purchases relative to purchases for current consumption (i.e., during NN): Purchases for future consumption take into account the expected consumption of products $-i$: Here, for simplicity, we assume perfect foresight of future prices and therefore future demand is a function of p_{-i}^t. Alternatively, under rational price expectations the consumer would purchase based on the expected future price (see Section 6).

The key observation, regardless of price expectations, is the following: if a product is currently on sale we know its effective next period price is p^S (since the product will be stored today for consumption tomorrow). In other words, the way to incorporate the dynamics dictated by storage is to consider the effective cost (or price) of consumption, which does not necessarily coincide with current price. In an inventory model, the effective or shadow price is a complicated creature that requires solving the value function. In our framework effective prices is just the minimum of current and previous prices.

When all products are storable, the case we consider from here onward, accounting for the storability is no more complicated. We just need to control for the effective cross price. For example, consider the event NN (product i is not on sale at t or at t

prices are constant within each regime, there is no reason to store and therefore the di¤erence in purchases (and consumption) across regimes helps recover preference parameters and :

Instead of observing long lasting price di¤erences we may observe high frequency price changes, like in the case of sales. Consider for simplicity just three periods, and suppose product 1's price decreases during the second period $p = p = p^N$ and $p^S = p < p$; while product 2's price remains constant at p. Denote by $p = p \quad p = p^S \quad p^N < 0$:

Since storing is free, consumers (who store) will purchase all of period 3 consumption, $q\ (p^S; p)$; in period 2: Notice the e¤ective price of product 1 in period 3 is actually the lowest of periods 2 and 3 prices, $\min\{p; p\} = p^S$: The consumer can time her purchases to minimize expenses. In this case, period 3 consumption is determined by p:

Quantities purchased by a storing consumer over the three periods (according to equation 3) are:

	$t = 1$	$t = 2$	$t = 3$
$p^t =$	p^N	p^S	p^N
$x =$	q	$2(q \quad p)$	0
$x =$	q	$q + \quad p$	$q + \quad p$

where $q = q\ (p^N; p)$ and $q = q\ (p^S; p)$. Should we estimate demand statically we would estimate the following price e¤ects:

Own price

effect on cross price responses, but did not show the expected bias theoretically. The model predicts cross price effects are understated. In period 3 the observed and effective prices differ. The effective price, which dictates consumption of good 1, is the period 2 purchase price. In the estimation we would instead interpret the price increase (observed in period 3), which is not accompanied by an increase in purchases of product 2, as lack of cross price reactions.

3.4 Discussion of the Main Assumptions

in storage in di¤erent states. Second, it helps detach the storage decision of di¤erent prod-

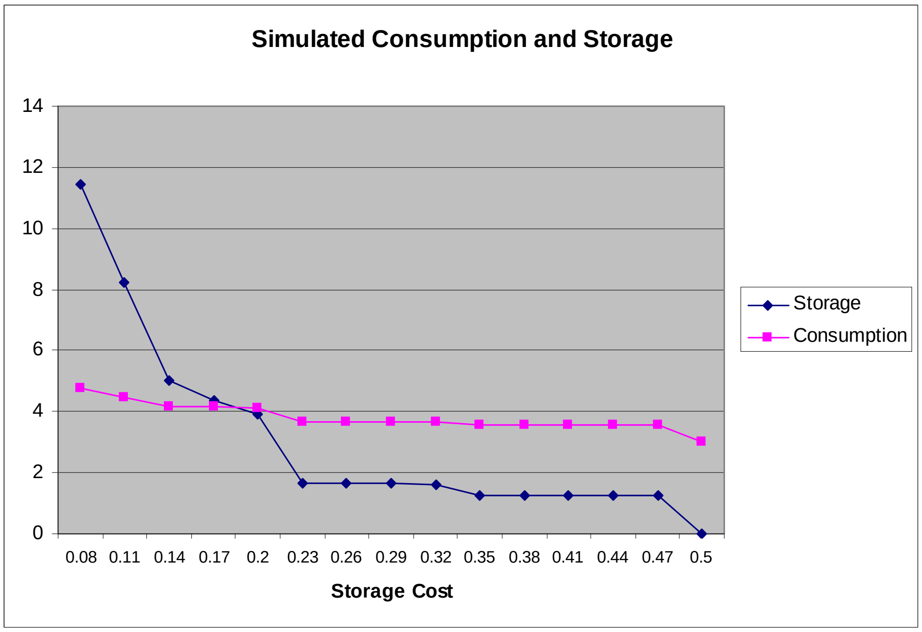

Figure 2: Optimal Dynamic Behavior as a Function of Storage Costs

Figure 3 displays the percent bias in the price coe¢ cient for OLS estimates and our …x assuming $T = 1$ and $T = 2$. For moderate levels of anticipated purchases the proposed …x does well. On the other hand, OLS shows substantial bias, about 60%, even for modest levels of storage. For very low storage costs all estimates overstate price responses. However, while the $T = 1$ …x is o¤ the mark by 40% the OLS estimate is over 160% o¤. As expected the $T = 2$ …x does better than the $T = 1$ …x for very low storage costs.

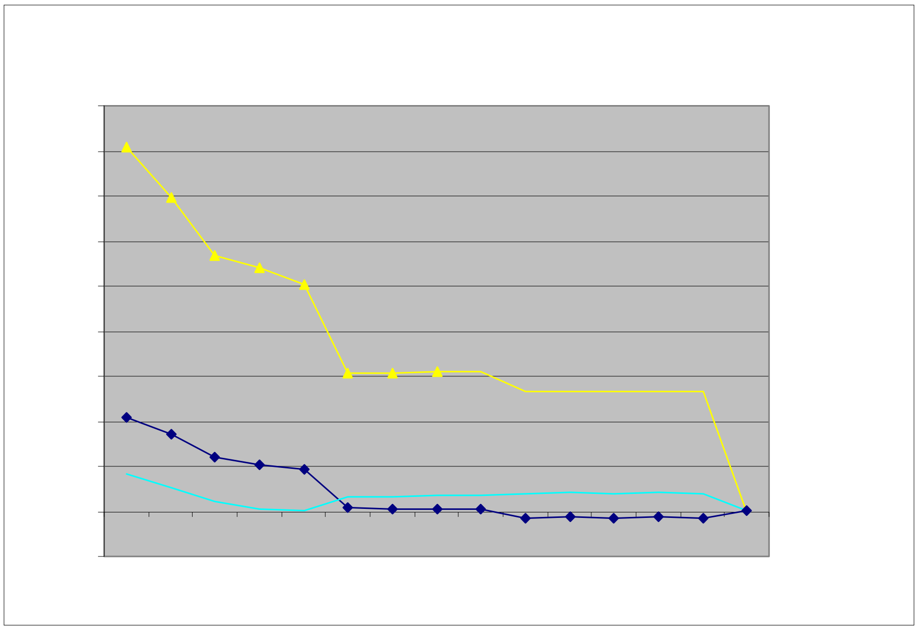

Figure 3: Percent Bias in Estimated Slope Parameter

Table 2 presents mean estimates and mean squared error of the different estimates by storage cost. It shows that the $T = 1$ ...x does best when the average storage (conditional on holding storage) is in the ballpark of one period of consumption (i.e., $c = 0.29$ and 0.38), while the $T = 2$ ...x is closest to target for $c = 0.17$ when the average storage is about twice the flow consumption. Both uniformly dominate OLS, unless storage is absent.

Table 2: Monte Carlo Simulations

	Simulated Data		Mean			MSE		
			OLS	T = 1	T = 2	OLS	T = 1	T = 2
c	Consumption	Storage	N=100					
0.08	4.80	16.20	10.57	5.82	4.73	43.98	3.59	0.66
0.17	4.19	9.31	8.36	4.89	4.07	19.26	0.92	0.13
0.29	3.64	5.80	6.50	4.08	4.30	6.39	0.14	0.35
0.38	3.54	5.05	6.16	3.91	4.37	4.75	0.15	0.38
0.50	2.97	0	3.99	4.00	3.99	0.05	0.20	0.19
			N=200					
0.08	4.80	16.20	10.50	5.71	4.68	42.59	3.02	0.52
0.17	4.19	9.31	8.35	4.87	4.07	19.06	0.83	0.07
0.29	3.64	5.80	6.50	4.07	4.29	6.32	0.07	0.21
0.38	3.54	5.05	6.14	3.90	4.35	4.63	0.08	0.23
0.50	2.97	0	4.00	4.01	3.99	0.03	0.09	0.09
			N=500					
0.08	4.80	16.20	10.48	5.66	4.66	42.08	2.79	0.46
0.17	4.19	9.31	8.32	4.84	4.05	18.69	0.73	0.03
0.29	3.64	5.80	6.47	4.05	4.28	6.13	0.03	0.12
0.38	3.54	5.05	6.13	3.89	4.33	4.52	0.04	0.15
0.50	2.97	0	4.00	4.00	4.00	0.01	0.04	0.03

Note: Means and mean squared error of estimates of the slope coe¢ cient, beta, computed based on 1,000 repetitions of each estimation. The data was generated using with a slope parameter of 4. The storage level is the average storage conditional on being positive, as oppose to Figure 2 that shows the unconditional average storage.

4 Identi…cation and Estimation

4.1 How Do We Recover Preferences?

Before presenting the estimation we discuss intuitively how the model helps recover preferences. We o¤er two approaches, both are part of the full estimation, but discussing them separately helps clarify what variation in the data identifies the parameters. The first approach is based on events without storage, while the second approach imputes storage and purges it from purchases.

For simplicity, assume a single product in which equation 3 suggests that during NN and SS demand is given by $q(p^t)$, while during SN demand is scaled down by ω and during NS it is scaled up by $2 - \omega$. This suggests two different ways to recover the model's parameters from the data. We will refer to the first as "timing" restrictions. According to the model during sale periods that follow a sale (event SS) purchases equal consumption $x(p) = q(p)$. Basically, after purchasing for storage, the pantry is filled, consumers (whether they are a storer or not) purchase for a single consumption event. Since both NN and SS events involve purchases dictated by $q(p)$ we can rely on them to estimate preferences. Price variation across efficient estimators, coming in the next section, will combine all this information and further control for differences across stores, prices of other products, and promotional activities.

Notice that both restrictions render a lower price sensitivity than the one implied by the static estimates.

4.2 Estimation

We follow the two strategies described above to estimate preferences. The first strategy uses data only from the NN and SS periods, which involve no storage. The second approach uses data from all periods, and is therefore more efficient, but it requires non-linear estimation. Linear estimation allows us to recover all the parameters of the model, except the fraction of consumers who store. To obtain the exact estimating equations we combine equations 2 and 3, and allow for a panel structure (that exists in the data we use below). To account for the store level fixed effects we de-mean the data. For prices this is straightforward. For quantities we have to account for the re-scaling in different regimes. We show in the Appendix how to modify the estimating equation to account for this re-scaling.

We estimate all the parameters by least squares, linear or non-linear depending on the equation. In principle, we could use instrumental variables to allow for correlation between prices and the econometric error term. However, we do not think correlation between prices and the error term is a major concern in the example below.

5 An Empirical Application: Demand for Colas

The average numbers (from Table 1) used in the previous section do not exploit price variation across stores, or within a regime (for a given store). They also neglect to properly control for the prices of substitute products.[4] We now estimate the model using all the events adding these additional controls.

[4] This is a serious concern since promotions of Coke and Pepsi are probably correlated, thus, a low Coke price may be also reflecting a high price of the closest substitute, thus contaminating the price reactions we infer.

5.1 Data

The data we use was collected by Nielsen and it includes store-level weekly observations of prices and quantity sold. The data set includes information at 729 stores that belong to 8 different chains throughout the Northeast, for the 52 weeks of 2004. We focus on 2-liter bottles of Coke, Pepsi and store brands, which have a combined market share of over 95 percent of the market.

There is substantial variation in prices over time and across chains. A full set of week dummy variables explains approximately 20 percent of the variation in the price in either Coke or Pepsi, while a full set of chain dummy variables explains less than 12 percent of the variation.[5] On the other hand, a set of chain-week dummy variables explains roughly 80 percent of the variation in price. Suggesting similarity in pricing across stores of the same chain (in a given week), but prices across chains look quite different. As a first approximation it seems that all chains charge a single price each week. However, three of the chains appear to define the week differently than Nielsen. This results in a change in price mid week, which implies that in many weeks we do not observe the actual price charged just a quantity weighted average. In principle we could try to impute the missing prices. Since this is orthogonal to our main point we drop these chains.

We need a definition of a sale, or more precisely, we need to identify periods of advance purchases. Figure 4 displays the distribution of the price of Coke in the five chains we examine below. The distribution seems to have a break at a price of one dollar, which we use as the threshold to define a sale. Any price below a dollar is considered a sale, namely, a price at which storers purchase for future consumption. This is an arbitrary definition. A more flexible definition may allow for chain specific thresholds, or perhaps moving thresholds over time. For the moment we prefer to err on the side of simplicity. Using this definition we find that approximately 30 (36) percent of the observations are defined as a sale for Coke (Pepsi). Interestingly, sales are somewhat asynchronized with only 7 percent of the observations exhibiting both Pepsi and Coke on sale (compared to a 10.5 percent predicted if the sales were independent).

[5]These statistics are based on the whole sample, while the numbers in Table 2 below are based on only five chains as we explain next.

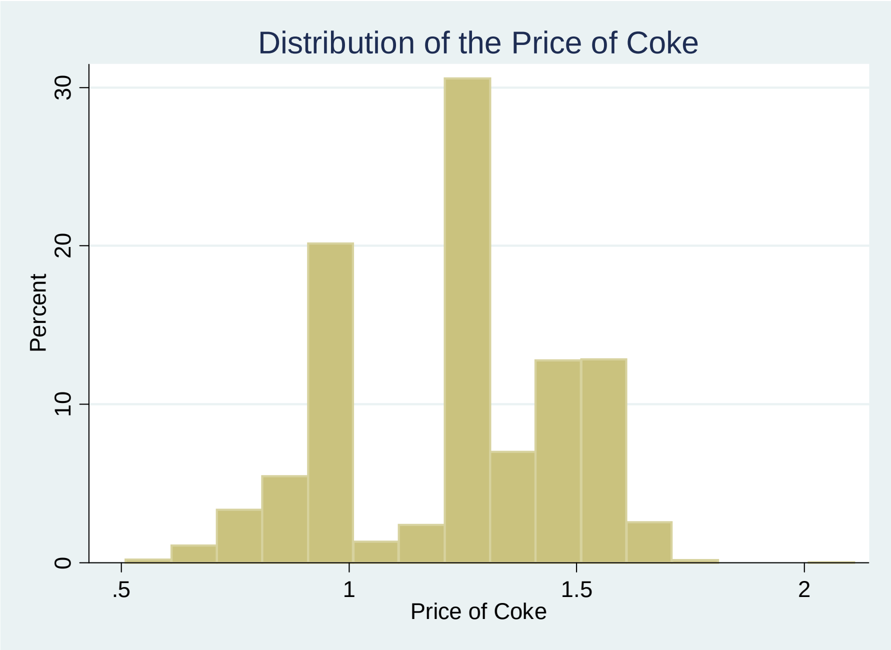

Figure 4: The Distribution of the Price of Coke

For the analysis below we use 24,674 observations from …ve chains. The descriptive statistics for the key variables are presented in Table 3.

Table 3: Descriptive Statistics

			% of variance explained by:		
Variable	Mean	Std	chain	week	chain-week
Q_{Coke}	446.2	553.2	5.6	20.4	52.5
Q_{Pepsi}					

5.2 Results

The estimation results are presented in Tables 4 and 5. All columns present least squares estimates of linear demand. The dependent variable is the number of 2-liter bottles of Coke or Pepsi sold in a week in a particular store. All the columns include the price of the store brand and store fixed-effects. The first column displays estimates from a static model, with store fixed effects. Column 2 presents our model estimated using the timing restriction only (namely, using the sub-sample with the events in which the model predicts no storage). Columns 3 and 4 present estimates of the full model.

The difference between columns 3 and 4 is that in column 3 we control for the current price of the competing products. Column 4 instead controls for the effective price. According to the model the effective price faced by a storer is the minimum of current and last period price.[6]

In column 5 we present the results from a model that allows different price sensitivity between storers and non-storers. Finally, in column 6 we present results where we replace the perfect foresight assumption with a rational expectation assumptions. We discuss this model and the results in the next section.

All the estimates from our model suggest lower (in absolute value) own price effects and higher cross price effects, for both Coke and Pepsi. The estimated proportion of consumers who do not stockpile is around half the population, and is slightly higher for Coke. Consistent with this estimate, the differences between the static and dynamic estimates are larger for Pepsi than for Coke.

The estimates in column 3 are of no interest on their own. According to the model, the cross prices controls are incorrect. The model prescribes the use of past prices during periods preceded by a sale (i.e., the effective price that dictates the consumption of storers is the lagged price). However, if the model is irrelevant, or demand dynamics absent, as we move from column 3 to column 4 we would be introducing noise in the price of the competing product. As such we would expect the coefficient of Pepsi in the Coke equation (and Coke's in the Pepsi equation) to be lower, due to measurement error (assuming the introduced noise will generate classical measurement error). Interestingly, both cross price

[6] For the non-storer the current price is the effective one. Thus, because of linerity of the demand curve, the aggregate effective price is the weighted average of the prices faced by storers and non-storer; weighted by the proportion of each type of buyer in the population. The price is recomputed as the estimation algorithm seaches for the optimal ω.

effects increase substantially as we replace current price by effective price. Suggesting the latter is the correct control, and that indeed dynamics are present.

Table 4: Demand for Coke

	FE	Timing Only	All Restrictions		Different slopes	Rational Exp
	(1)	(2)	(3)	(4)	(5)	(6)
P_{Coke}	-1428.2	-743.4	-967.4	-938.9	-522.4	-767.8
	(11.1)	(11.8)	(11.2)	(11.1)	(35.1)	(8.8)
P_{Pepsi}	66.5	191.6	82.8	150.1	-73.1	195.4
	(11.9)	(10.9)	(11.0)	(11.4)	(36.4)	(12.3)
P_{Coke} storers					-1273.1	
					(16.7)	
P_{Pepsi} storers					145.01	
					(21.3)	
ω (fraction non- storers)			0.53	0.53	0.35	0.57
			(0.01)	(0.01)	(0.01)	(0.01)
Cross price Corrections			No	Yes	Yes	Yes

Note: All estimates are from least squares regressions. The dependent variable is the quantity of Coke sold at a store in a week. The regression in column (1) includes store fixed effects. The regression in column (2) is the same as in column (1) but uses only the NN and SS periods. The regressions of columns (3)-(4) impose all the restrictions of the model using the actual and effective price. Column (5) allows for different slopes for consumers who store and those that do not. Column (6) assumes rational expectations rather than perfect foresight. Standard errors are reported in pharenthesis.

Table 5: Demand for Pepsi

	FE	Timing Only	All Restrictions		Di¤erent Slopes	Rational Exp
	(1)	(2)	(3)	(4)	(5)	(6)
P_{Coke}	-20.9	71.8	62.8	106.6	-246.2	140.0
	(12.2)	(11.0)	(10.5)	(10.5)	(26.9)	(12.1)
P_{Pepsi}	-1671.3	-994.0	-1016.5	-996.9	-341.3	-762.3
	(13.1)	(15.8)	(11.6)	(11.6)	(29.3)	(8.9)
P_{Coke} storers					216.7	
					(13.3)	
P_{Pepsi} storers					-1255.0	
					(14.8)	
! (fraction non-storers)			0.44	0.44	0.28	0.47
			(0.01)	(0.01)	(0.01)	(0.01)
Cross Price Corrections			No	Yes	Yes	Yes

All estimates are from least squares regressions. The dependent variable is the quantity of Pepsi sold at a store in a week. The regression in column (1) includes store …xed e¤ects. The regression in column (2) is the same as in column (1) but usese only the NN and SS periods. Column (5) allows for di¤erent slopes for consumers who store and those that do not. Column (6) assumes rational expectations rather than perfect foresight. Standard errors are reported in pharenthesis.

As long as $p^*_{NS} > p^*_S$ then $\overline{p} = p^*_{NS}$ while $\underline{p}$ is the price charged by a non-discriminating monopolist who faces demand $! Q_{NS}(\underline{p}) + 2(1 \quad !)Q_s(\underline{p})$; namely, demand with additional weight on the storing population. It is easy to see that $p^*_S < \underline{p} < p^*_{ND}$:

Optimal pricing involves high prices targeting non-storers who are less price sensitive and sales, targeting storers. Under constant prices the seller would set a price that targets

The quantities $\overline{q}^P$ and $\underline{q}^P$ are still linear in prices and have the same functional form, but they di¤er from the demand functions of the static problem (in equation 2). The reason is simple, in the static problem the consumer reacts to a Coke sale by adjusting both Coke and Pepsi quantities. Instead, $\overline{q}^P$ and $\underline{q}$

where $_{ij}$ is the utility from the attributes of the product both observed and unobserved[8], $_{i}$ is the marginal utility of income and $''_{ijt}$ is a transitory shock. For now, we assume perfect foresight of both prices and individual shocks. We can think of $''_{ijt}$ as capturing transitory needs known in advance, like having guests the following week. As in the standard discrete of the different products. In some applications an unknown ε_{ijt} might be appropriate. In two periods ($t = 0; 1$) and the last two periods ($t = 2; 3$) and define prices as revenue divided by quantity, then estimation based on the aggregate data would recover long run effects.

The success of this approach in recovering long run responses relies crucially on several assumptions, like lack of heterogeneity in storage. We provide, in the Appendix, an analytic example that shows this.

We now apply these alternative corrections. The results for Coke are presented in Table 6. The first two columns repeat the results from the store fixed-effects regression, and from our model. The next two columns present the long run effect from models that include 1 and 4 lags, respectively. The results are not very promising. Both lagged prices models impact the own price elasticity in the "right" direction but the magnitude is smaller than our correction. The results do not look good for the cross price effect. The first model does not change the cross-price effect by much. The second, with more lags, does but estimates a negative cross-price elasticity.

The last two columns present the results from aggregating over time: into bi-weekly of prices. The regressions in columns (5) and (6) aggregate the data to a bi-weekly and monthly level (and use unit prices). Standard errors are reported in pharenthesis.

8 Concluding Comments

We offer a simple model to account for demand dynamics due to consumer inventory behavior. The model can be estimated using store level data. An application to demand for Coke and Pepsi yields reasonable estimates. At the same time, corrections based on alternative methods, like aggregation or control for lagged variables, do not perform well.

The base results rely on many assumptions, most of which can be relaxed. As we showed we can allow for heterogeneity in preferences, more flexible demand systems, and rational expectations. We can also let the fraction of consumers who store vary with price. Of course, some of these extensions increase the complexity of the model and defeat our goal of delivering a simple model.

We use the simplicity of the model to derive markups implied by dynamic pricing, rather than plugging demand estimates into static first order conditions. The standard static approach underestimates market power for two reasons. First, demand elasticities biases (both own and cross) imply lower markups. Second, the static first order conditions imply lower mark-ups than the dynamic ones.

9 References

Berry, Steven, James Levinsohn, and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841-890.

Berry, Steven, James Levinsohn and Ariel Pakes (1999) "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy," *American Economic Review*, 89(3), 400–430.

Blattberg, R. and S. Neslin (1990), Sales Promotions, Prentice Hall.

Boizot, C., J.-M. Robin, M. Visser (2001), "The Demand for Food Products. An Analysis of Interpurchase Times and Purchased Quantities," *Economic Journal*, 111(470), April, 391-419.

10 Appendix

10.1 Purchases when $T = 2$

The predicted purchases when $T = 2$ (assuming a single product) are given by:

$$x(p^t) = \begin{cases} \omega q(p^t) & SNN \\ q(p^t) & NNN \\ \omega q(p^t) & NSN \text{ or } SSN \\ \omega + 3(1-\omega)q(p^t) & NNS \\ \omega + 2(1-\omega)q(p^t) & SNS \\ q(p^t) & NSS \text{ or } SSS \end{cases} \quad \text{if} \tag{5}$$

First, notice there are 8 states, some of them involve similar predicted purchases. In contrast to equation 3 where demand is affected by lagged prices, when $T = 2$ demand depends on whether there was a sale two periods ago. Second, notice how (some) events are split. Event NN needs to be split into SNN and NNN; because a storer who purchased two periods ago on sale does not buy today at a regular price, while she would buy if two periods earlier there was no sale, namely, in event NNN: Predicted purchases in events SS and NS are not affected by $t-2$ events, thus they require no modification from equation 3. Purchases differ between SNS and NNS because in SNS current consumption comes out of storage.

10.2 Estimating equations

We choose the parameters to minimize the sum of squares of the difference between observed purchase and those predicted by the model. The data consists of a panel of quantities and prices in different stores. Since purchases are scaled differently in different states in order to account for store fixed effects we need to transform the predicted purchases as follows. Let j denote the store.

$$x_{ijt} = f_t\left(\frac{1}{T}\sum\left(\frac{x_j}{f}\right) + (p_{ijt} - \bar{p}_{i:t}) + (pe_{-ijt} - \overline{pe}_{-i:t})\right)$$

where f_t is the factor by which demand is scaled up in period t, $\overline{p}_{i:t}$ is the within store average, and pe is the effective cross price (as defined in the text). Note, that the effective price is a function of !. For the base model

$$f_t = \begin{cases} 1 & \text{if } NN \text{ or } SS \\ ! & \phantom{\text{if }} SN \\ 2 \quad ! & \phantom{\text{if }} NS \end{cases}$$

10.3 Example where aggregation fails

Consider the following example where aggregation fails. Suppose there are two types of consumers. Type A consumers can store for one period, type B cannot store. Assume four time periods with $p \;<\; p \;=\; p \;=$