

The Impact of Soda Taxes on Consumer Welfare: Implications of Storability and Taste Heterogeneity

Emily Yucai Wang *

November 26, 2014

Abstract

The typical analysis on the effectiveness of soda taxes relies on price elasticity estimates from static demand models, which ignores consumers' inventory behaviors and their persistent tastes. This article provides estimates of the relevant price elasticities based on a dynamic demand model that better addresses potential inter-temporal substitution and unobservable persistent heterogeneous tastes. It finds that static analyses overestimate the long-run own-price elasticity of regular soda by 60.8%, leading to overestimated consumption reduction of sugar-sweetened soft drinks by up to 57.9% in some cases. Results indicate that soda taxes will raise revenue but are unlikely to substantially impact soda consumption.

*University of Massachusetts, Amherst; emilywang@resecon.umass.edu

This paper is a revised version of the second chapter of my Ph.D. thesis. I deeply appreciate the guidance and advice provided to me by Andrew Sweeting as well as Arie Beresteanu, James Roberts, and Chris Timmins. I sincerely thank the editor, Aviv Nevo, and two anonymous reviewers for their constructive comments and suggestions. In addition, this paper has benefitted from fruitful discussions with Pat Bayer, Federico Bugni, Paul Ellickson, Gautam Gowrisankaran, Jun Ishii, Carl Mela, Angela de Oliveira, Christian Rojas, Steven Tadelis, and Ken Wilbur. I am grateful for the opportunity to work with the IRI data, provided by Paul Ellickson, Carl Mela, and Andrew Sweeting. Furthermore, I would like to thank Christoph Bauner for the support and help he provided throughout this project. Any errors are my own.

1 Introduction

Obesity has become an alarming concern among health professionals and policy makers alike, with one in four American adults believed to be obese and estimated medical costs now exceeding 147 billion dollars a year (Finkelstein et al. 2009). The American Heart Association implicates the over-consumption of added sugars - largely from sodas and fruit drinks - as a major contributing factor to the high US obesity rate. Since 2009, the Centers for Disease Control and Prevention has listed reducing the intake of sugar-sweetened beverages as one of its top obesity prevention strategies (Keener et al 2009). The public health concern has prompted public calls for a tax on sugar-sweetened beverages. Proponents of the tax hope that consumers, faced with higher prices for sugar-sweetened beverages, will reduce their consumption of sugar-sweetened beverages by substituting to non-taxed, low-sugar alternatives. Currently, thirty-four states apply sales tax to soft drinks (Jeffords 2010). As of May 2011, fifteen states have discussed imposing specific taxes on sugar-sweetened beverages during their legislative session. In the November 4, 2014 election, Berkeley, California enacted the first soda tax in the U.S., establishing a penny-per-ounce tax on sugary drinks. This article estimates the effectiveness of potential soda taxes and shows that such policies may not be as effective at curbing consumption as previously predicted. The taxes are effective, however, at raising revenue.

A large health literature debates the effectiveness of soda taxes by predicting the reduction in the consumption of sugar-sweetened beverages induced by such a tax. Following Andreyeva, Long and Brownell (2010), most of these studies use price elasticity estimates from traditional static demand models. These static models face two key shortcomings. First, they ignore the fact that most beverage products are storable and experience frequent price reductions, suggesting potential for sizable inter-temporal substitution. Second, the traditional models ignore the fact that consumers tend to have strong preferences over product choices for reasons not entirely observable, suggesting unobservable persistent heterogeneous tastes. These two factors imply that existing studies over-predict price elasticity and exaggerate the consumption response from a given tax. This article provides estimates of the relevant price elasticities based on a dynamic demand model that better addresses potential inter-temporal substitution and unobservable persistent heterogeneous tastes. When applied to weekly scanner data from 2002 to 2004, the model finds price elasticities to be

substantially lower and the resulting public health gains from various proposed taxes significantly smaller than what has been claimed in the literature to date.

The chosen dynamic demand model is in the style of Hendel and Nevo (2006b). Households in the model are forward-looking and may choose to stockpile for future consumption. To accommodate the current application, I replace the traditional household-brand-size fixed effects with household-specific random coefficients. This allows consumers' taste heterogeneity to be modeled as both a function of observable attributes as well as unobservable persistent preferences. The resulting specification allows for more flexible substitution patterns, which provides the necessary foundation for analyzing the distributional impact of relevant taxes. I find this expanded scope of substitution important in my estimation and subsequent welfare analysis.

I use the estimated distribution of household preferences to perform policy analysis. Specifically, I study two of the more prominent tax proposals: a 10% sales tax and a penny-per-ounce tax. For each income bracket, I calculate the effects of the sugar taxes at four pass-through levels: 25%, 50%, 75%, and 100%. I simulate the post-tax soda consumption pattern, calculate the compensating variation, and estimate the consumer welfare loss. The results show that ignoring both inventory and unobservable persistent tastes leads to an overestimated long-run price elasticity of regular soda¹ of roughly 60.8%, further leading to overestimated reduction in sugar-sweetened soft drinks by as much as 57.9% in some cases. Moreover, I find that although the policies generate a small deadweight loss, they are regressive in nature and tax poor households more than their rich counterparts. This is because poor households not only consume more regular soda than rich households, they also have a less elastic demand for regular soda than rich households.

The article is organized as follows. The remainder of the introduction reviews the related literature. Section 2 presents industry details and the data. Sections 3 and 4 present the model and the estimation procedure. Section 5 presents the empirical results. Section 6 provides a discussion on welfare implications of the sugar taxes, and section 7 concludes the article.

¹The term "price elasticity of regular sodas" will be used to refer to the aggregate category-level price elasticity throughout the text. Price elasticities for individual products are specified directly (e.g., the price elasticity of regular Coke).

Related Literature

Various taxes have been proposed as a means of controlling the consumption of sugar-sweetened soft drinks. Proponents of such taxes draw on existing health literature that establishes a link between soft drinks and health problems. For instance, Schulze et al. (2004) provide evidence for a correlation between soda consumption and diabetes. Similarly, Bray et al. (2004) find a correlation between obesity and high fructose corn syrup, a main ingredient in regular soda. One possible explanation for the correlations is the increase in soda consumption. Nielsen and Popkin (2004) show that average daily caloric intake from soft drinks increased from 2.8%—roughly 50 calories—in 1977 to 7.0%—roughly 144 calories—in 2001. The large increase in soda consumption in conjunction with evidence showing that small dietary changes can effectively combat obesity (Hall and Jordan, 2008) suggests that soda taxes may help reduce obesity.

Several articles argue that sugar-sweetened soft drinks are the largest cause for obesity and suggest that they should be taxed in order to improve public health (Jacobson and Brownell, 2000; Nielsen and Popkin, 2004; Mello, Studdert, and Brennan, 2006; Brownell and Frieden, 2009; Brownell et al, 2009; and USDA, 2010). The most prominent of these are Brownell and Frieden (2009) and the USDA (2010) report.

Brownell and Frieden (2009) state that we are experiencing an obesity epidemic and hence should tax soda heavily, similar to other sin goods. The authors cite a study conducted by Yale University’s Rudd Center for Food Policy and Obesity suggesting that every 10% increase in price will lead to a 7.8% decrease in soda consumption.² Furthermore, they state that the penny-per-ounce excise tax proposed in New York is expected to reduce consumption by 13%. The authors do not document how these estimates are obtained, but the estimates are consistent with highly elastic demand for soda. This contrasts my estimates which suggest that demand for sugar-sweetened soft drinks is inelastic.

Following a prompt from the Institute of Medicine and the National Academies of Science, the Economic Research Service division of the USDA (2010) examined the health effects of taxing

²This study refers to Andreyeva, Long, and Brownell (2010), who review 160 studies conducted in the US since 1970, 14 of which are relevant for the soft drinks category. The authors report that according to these previous studies, soda drinks are among the products most responsive to price changes with an average price elasticity estimate of 0.79 (in absolute value). They state that a 10% increase in soft drink prices should reduce consumption by 8% to 10%. Following this paper, a large share of the literature in public health uses this estimate to determine the effectiveness of a soda tax.

sugar-sweetened beverages, reporting that a 20% soda tax would be expected to reduce overweight prevalence from the current 66.9% to 62.4% and obesity from the current 33.4% to 30.4%. The report draws on the 1998-2007 Nielsen Homescan and the National Health and Nutrition Examination Survey (NHANES). Applying Deaton and Muellbauer's (1980) Almost Ideal Demand System (AIDS) to the Nielsen weekly household purchase panel, the authors find that the category own-price elasticity of caloric sweetened beverages is -1.264, suggesting an elastic demand. The price elasticity estimate is then applied to individual beverage intake data reported in the NHANES to estimate changes in caloric intake in response to a tax-induced 20% increase in the price of caloric sweetened beverages. In comparison to this report, I find a much lower price elasticity of demand for regular sodas, at -0.5744.³

There has also been a growing body of literature opposing soda taxes. Kaplan (2010) points out that the medical trials used in Brownell et al. (2009) do not provide enough evidence for the claim that sugar taxes will decrease obesity in the population. Hall and Jordan (2008), Finkelstein et al. (2010), Katan and Ludwig (2010), and Patel (2012) suggest that small dietary changes would not cause much change in weight. The most salient to my research is Patel (2012).

Patel (2012) analyzes the impact of hypothetical soda taxes using a static Berry, Levinsohn, and Pakes (1995) (BLP) framework applied to a five-year panel of Nielsen Scantrack data from April 2002 to April 2006. The article finds that the median body mass index (BMI) for an obese individual is 33.39 and the median expected reduction in BMI for these individuals is 0.287. BMI changes of this magnitude are not likely to result in meaningful reductions in illnesses or medical costs. Although Patel (2012) finds little evidence that soda taxes will be as effective as claimed by their proponents, that research differs significantly from the current one in the models implemented, leading to significantly different policy implications. Similar to the USDA (2010) report, Patel (2012) uses weekly market sale volume changes from temporary price reductions. As consumers' stockpiling behavior is not modeled or incorporated in the analysis, it similarly overestimates the long-run own-price elasticities. As a result, the article finds demand for soda, both regular and diet, to be elastic, with an own-price elasticity of -3.097 for regular Coke and -2.183 for diet Coke. These estimates

³Since both papers use weekly scanner data, the differences in elasticity estimates come predominantly from the models employed. The AIDS model does not account for intertemporal substitution. Hence when applied to weekly data any differences in quantity changes are interpreted as changes in immediate consumption when in fact households may be stockpiling in anticipation of higher prices in future weeks.

are close to those previously found in the public health literature (generally estimated around -5). Despite these elastic demand estimates, Patel (2012) finds little change in consumers' BMIs. This result is driven by the conversion of calories in soda to consumers' BMIs in the steady state. It takes a large decrease in caloric intake to decrease an individual's BMI. This is an important but different point than the one illustrated in this article. When inventories are accounted for explicitly in the model, demand elasticity estimates are much smaller in absolute value (for instance, -1.20 for regular Coke). Therefore, the resulting decreases in consumers' weight become negligible.

Some proponents of soda taxes support these policies from a revenue-generation point of view. For example, Jacobson and Brownell (2000) acknowledge that sugar taxes may not improve public health but claim that they will generate sizable revenues for health programs. On the opposite side, Gostin (2007), Byrd (2004), and Powell et al. (2007) argue against these taxes on the grounds that they are regressive and target the poor and minorities. In this article, I find that the imposed taxes do not generate large deadweight losses: A large portion of the population has strong preferences for regular soda and do not switch out of their preferred drinks post tax. However, I find that soda taxes are regressive in nature. As poor households have heavier consumptions than their richer counterparts, they are also impacted more by the taxes.

On the methodological front, this article builds on two strands of literature:⁴ static demand models with consumer heterogeneity and dynamic demand models of storable goods. In terms of static demand models, the literature started by Bresnahan (1981) and Berry, Levinsohn, and Pakes (1995) and continued by Nevo (2000) models static consumer decisions for differentiated products. These articles and those that have followed show that it is important to incorporate consumer heterogeneity in demand systems in order to obtain realistic predictions for differentiated products. Although this article also incorporates consumer heterogeneity into the demand system, it differs from these articles in two ways. First, households in the model are forward-looking. And second, the model is better suited for disaggregated household panel data, unlike BLP-style models, which are adapted for aggregated market-level data.

⁴There is also a line of literature on dynamic demand models of durable goods. Examples include Gordon (2006), Conlon (2009), and Nair (2007). Some of these papers, like Gowrisankaran & Rysman (2009), incorporate unobservable heterogeneous tastes into dynamic durable goods demand models. However, in most cases, durable goods are purchased once instead of repeatedly. The decisions that consumers make in these models focus on the timing of product replacement instead of stockpiling. Hence these models and methods are not appropriate for studying consumption of storable goods.

This article also fits into the recent literature on dynamic demand models of storable goods. This literature originates from works both in industrial organization and marketing and has seen increased applications (Erdem, Imai, and Keane, 2003; Hendel and Nevo, 2006a,b; Hartmann and Nair, 2010; Hendel and Nevo, 2012; and Osborne, 2013). Within the field of industrial organization, Hendel and Nevo have produced a series of influential articles on storable products. Hendel and Nevo (2006a) find evidence for the presence of stockpiling behavior during periods of price reductions. Using scanner data on laundry detergent as an example, they find that as time-since-sale lengthens, total quantities purchased increase. This suggests that households intertemporally substitute purchases. The important implications here is that static demand estimates of long-run price elasticities may be misestimated for storable goods that experience frequent sales.

Having shown evidence for stockpiling behavior, Hendel and Nevo (2006b) build a dynamic demand framework that explicitly accounts for inventory. Again using laundry detergent as an example, the authors use this framework to estimate the magnitude of the misestimation that can result from using a static demand model in a storable goods market. Their results suggest that static demand estimates may 1) overestimate own-price elasticities by 30%, 2) underestimate cross-price elasticities by up to a factor of 5, and 3) overestimate the substitution to the no-purchase by over 200%. These results have significant ramifications for predicting the effect of soda taxes, suggesting that using a static demand model to estimate the elasticities of soda—a storable good—may lead to exaggerated reductions of regular soda consumption.

Hendel and Nevo (2012) propose a new dynamic demand model that accommodates how consumers respond to temporary price reductions. The model is simple in its design and implementation but still captures inventory. It provides a quick and appropriate way of computing price elasticities for storable products. As I show in the robustness check section in the technical appendix, the long-run own-price elasticities in the model proposed in Hendel and Nevo (2012) are comparable to those implemented in the article at hand, which provides assurances for the full model and its implementation. However, this framework does not allow for more complex welfare analyses that require estimating the distribution of tastes. To understand the impact of the proposed taxes on consumer welfare, we need to estimate the share of the population with strong tastes for the products taxed. These distributions are difficult to recover from aggregate data. Hence, implementing the richer model proposed in this article is necessary for predicting the policy implications of the

proposed taxes.

2 Data

I use weekly scanner data from 2002 to 2004 provided by Information Resources Inc. (IRI). The data comprises two components: a household panel of randomly selected households⁵ and a store panel. For each household in each week, I observe whether any soda was purchased; if so, I observe which products were purchased, where it was purchased, how much was bought, and the total dollar amount paid. From each store in each week I observe the price charged, the total quantity sold, and all promotional activities for each product sold in the store. Sales from these stores account for over 97% of all purchases of soft drinks observed from the households in the panel. For each household, I observe a few basic demographic variables such as race, income, and household size. For each product, I observe the product name, brand, packaging, volume, and whether it is regular or diet.

Each observation from both panels has unique identifiers for the product, store, and week. These identifiers allow me to link the two sets of data and track prices and promotions for all products available to households on any given shopping trip. That is, I observe not only information on the purchase itself but also information regarding each household's complete choice set.

Household Panel I separate households into three groups by per capita income: less than \$10K, between \$10K and \$20K, and more than \$20K per capita.⁶ Table 1 reports some statistics associated with each income bracket. The data reveals several interesting patterns. 1) On average, the household size is decreasing in income. 2) The average weekly household soda purchase is slightly increasing in income. 3) The number of trips where households bought soda is decreasing in income. High-income households make fewer purchases of soft drinks on average than lower income households. Because they also purchase more units, this implies that rich households tend to buy

⁵IRI randomly selects a sample of households in two suburban areas around the Midwest and New England and requests their participation. Participating households conform fairly well to local demographics. The data does not suggest a selection bias in the panel.

⁶The data is collected predominantly from suburban areas of New England and the Great Lakes region. As a result, there is little variation in ethnicity. The majority of households in the sample are white. However, there are substantial income differences. While in an ideal world we would like to observe racial diversity, previous literature indicates that obesity correlates more significantly with income than with race (FRAC, 2010). In addition, the main point made in this paper is that previous literature, which does not account for stockpiling or heterogeneous tastes, overestimates the effect of taxation on soft drink consumption. This can be generalized to areas with other demographic compositions.

more soda per shopping trip. There are two ways of reconciling this fact. First, if we assume soda is purchased only for immediate consumption, then rich households do not drink soda as frequently as poorer households do, but they consume more when they do drink it. The other explanation is inventory: rich households purchase more soda in fewer trips because they stockpile more than poor households. They are more likely to have larger houses and hence more storage area at home. The second explanation seems more plausible and I show in the last part of this section that the inventory explanation better fits the data.

Table 1 also presents a breakdown of the products that households purchased in terms of the shares of cola drinks (Coke, Pepsi, and store brand) as well as shares of diet versus regular drinks. It is clear from the table that high-income households purchase more branded products and poorer households purchase more regular soda. The same pattern holds for all brands of soda.

Store Panel From the store panel, I observe weekly prices and advertising information for all items sold.⁷ For each store in each week, there are over 250 different products offered on average, a combination of all brands, regular/diet types, packaging, volume, and flavors. In estimating the model, prices and promotions of all products become part of the state space. Carrying prices and promotions of all products is computationally infeasible in the dynamic programming problem.

Quantities sold of specialty items, such as IBC root beer, are very small. I therefore restrict the set of soda brands to those whose market shares exceed 1%. In addition, I also include the generic versions of these branded products where they are available, even if their market shares fall below 1%. Consumers' substitution behavior between regular and diet sodas is one of the crucial components of policy study here, so I allow for both varieties in the analysis. In table 2, I rank all products included in the analysis according to their market share by sales volume in 2002. The soda market is fairly mature and market shares therefore remain stable over time. Sales volumes in 2003 and 2004 are similar to those in 2002.

One pattern emerges clearly from table 2: the market for soda is fairly concentrated. A few large brands—namely Coca-Cola, Pepsi, Sprite, and Mtn Dew—capture its majority. At the top, regular Coca-Cola (Coke) achieves nearly 20% of the market. By the tenth-ranked product, diet

⁷I do not directly observe the price of each product. Instead, I observe weekly total revenue and total units sold for every UPC (product). I take the average and use it as a proxy for product price. In terms of promotional activities, for each product in every week, I observe whether it is on display, on feature, and/or has price discounts.

Mtn Dew, the market share has decreased to a little over 1%. This suggests that consumers have strong preferences for top brands and hence implies that capturing households' intrinsic product tastes in the model may be important for studying the impacts of policy changes. The combined market share of all products included in the analysis is close to 90%.

Seasonality One concern about soft drinks is that their consumption could be influenced by holidays. I observe a very small effect in the data. Figure 1 shows weekly cola product sales by volume as a percentage of annual total observed in stores. The gray bars indicate the presence of a holiday, where holidays are defined very liberally (Super Bowl Sunday is counted as a holiday). The figure below presents the panel for 2004; the graphs for the other two years are similar. We see that there are increases in the volumes sold around some holidays. The dominant ones are around July 4th and Thanksgiving. I do not model holiday effects here and use data from the 2nd to the 26th week of each year, excluding most major holidays such as July 4th, Thanksgiving, and Christmas. Households that are present in multiple panels are treated as separate observations. Robustness checks with each household counted once turned up no major differences in estimates.

Stockpiling Behavior The main competing model is one in which consumption is directly influenced by price and no storage occurs. If this were true we should expect to see sales volume increase when price reductions are available but decrease and remain largely constant when there is no sale. However, this is not the case. The following graph shows that sales volume drops dramatically immediately after each sale ends then slowly grows again. This can be explained by stockpiling: households fill their inventories during price reductions. Hence, they do not need to purchase much soda immediately afterwards. As time passes households deplete their inventories and more and more households have to restock.

As an example, I use the sale of regular Coke from a representative store in 2002. figure 2 shows how sales volume evolves over the duration of a year and how it is influenced by price reductions. Each bar shows the sale volume for one week. Black bars indicate price reductions and gray ones denote weeks without sales.

It is clear that the demand for regular Coke dramatically increases when there is a price reduction. Moreover, we see that sales volume decreases drastically immediately after a price reduction

and picks up again in the following weeks. (These occurrences are marked by the horizontal lines beneath the bars.) These dips in purchases are consistent with households stocking up on soda when there are price reductions, reducing the need to buy soda immediately afterward. However, as their stocks become low they make more purchases again. Static models do not account for this effect but assume that all purchased units are immediately consumed. Hence, these models are misspecified.

To further distinguish between static and dynamic behavior I analyze how past prices influence current purchase size choices. Using weekly store sales data, the following regression provides additional evidence for the presence of stockpiling behavior. To be more precise, I regress the weekly quantities sold for each product in each store on its current week's price, inflation adjusted, and the number of weeks (duration) since it last experienced a sale. Table 3 shows the results from this simple regression.

If a static model is correct, we should expect duration to have no impact. That is, purchase decisions only depend on current prices and marginal utilities of consumption. However, if an inventory model is correct, purchase quantities should be greater as more time passes since the last sale. If households are stockpiling items not needed for immediate consumption, they can delay their purchases when prices are unfavorable. But as they go through longer periods of high prices, their inventories become depleted and take larger purchases to restock. Indeed, as table 3 shows the coefficient on duration to last sale is positive and highly statistically significant. This indicates that purchases are correlated with past prices, justifying a dynamic specification.

Although stockpiling behavior is present across the entire population, this behavior differs significantly across income groups. Low-income households are less able to stockpile compared to their higher income counterparts. Recall from table 1, we see that higher income households, on average, purchase more per shopping trip than the lowest income group. Several factors may contribute to this phenomenon. For instance, low-income households may have less capacity to stockpile at home, they are more likely to experience higher transportation cost (e.g., having to rely on buses), and may face more budget constraints. These factors all limit poor households' ability to take advantage of temporary price reductions. In figure ??, I plot the shopping behaviors of two representative households.

The left graph plots the shopping frequency and timing of a middle-income household and the right one plots the same for a low-income household. Both households purchased around 30 two-liter bottles of Coca-Cola in 2003. In both graphs, I indicate the price movements of a two-liter bottle of Coke by grey lines and I denote the purchases by red dots.

We see from the above graph that richer households are more likely to make purchases only when prices are low, using their inventory during high-price periods and waiting for the next sale. Poorer households, on the other hand, make purchases more frequently, even during non-sale periods. This speaks to the differing abilities to stockpile across income groups and implies that poorer households are more limited in their ability to take advantage of temporary price decreases. As they cannot stockpile as much when a sale happens, they are also more affected by price fluctuations. In the estimation, this fact will be reflected as higher estimated storage costs.

Furthermore, poor households' limited storage ability coupled with higher persistent preferences negatively impacts their welfare. Households that are less able to stockpile but are also less willing to switch to cheaper alternatives will end up paying more for their preferred product. This will be further exaggerated by the proposed taxes. As we will see in the policy simulations, the combination of these factors implies that the tax is regressive in practice.

Persistent Heterogeneous Preferences Table 4 shows the probability of observing brand and diet purchases conditional on the observed last purchases. I display the transition of brand purchases in the first column and the transition of diet soda purchases in the second column.

If households do not have intrinsic brand or diet preferences, then the product choice in the second period should be independent from that of the first. That is, the probability of buying product x in period $t + 1$ conditional on buying the same product in period t should be roughly the same as the probability conditional on buying product y in period t . However, this is not the case, as seen in table 4. The probability of purchasing Coke in period $t + 1$ conditional on purchasing Coke in period t is 84.81%. The probability of purchasing Coke in period $t + 1$ conditional on purchasing Pepsi in period t , however, is only 26.66%. The same is true for diet preferences. The probability of observing a purchase of a regular soda in period $t + 1$ conditional on purchasing regular in period t is 63.83% and the probability of buying a regular soda in period $t + 1$ conditional on observing a diet purchase in period t is 22.04%. This suggests that households have persistent brand and diet

preferences.

3 Model

The model built here is household specific. Each household has its own intrinsic tastes for products, its own capacity for storage, and its own willingness to pay for goods. Households derive utilities from product specific consumptions and from any available advertising or promotions. They face prices for the goods and inventory costs for storing unused quantities of purchased goods. Together, these factors influence how each household responds to possible soda taxes.

Households in the model are forward looking and maximize their present value of expected future utilities. As such, they use current market information to form expectations over future prices and may choose to build inventories to guard against future price increases. Soft drinks are storable: Quantities not currently consumed are stored for future consumption. In each period, households make three decisions: how much soda to consume, how much soda to purchase, and—if they chose to buy anything at all—which soda products to buy. These decisions are governed by household persistent preferences, available information (current prices, advertising, consumption shock and inventories) and expectations concerning future prices and consumption shocks.

This model builds on the dynamic demand framework of Hendel and Nevo (2006b). Similar to Hendel and Nevo (2006b), the model flexibly incorporates observable consumer taste heterogeneity. Households differ according to their observable attributes, e.g. income and race. In addition, the model allows for a greater degree of persistent unobservable consumer taste heterogeneity. Households differ in their tastes for product characteristics, such as whether the product is regular or diet as well as their sensitivities for prices/advertising and their abilities to stockpile. These persistent preferences are modeled as random coefficients and pick up systematic differences in purchase patterns when observable characteristics are identical across households.

Model Setup

Households maximize the sum of discounted utilities derived from consuming soft drinks and an outside good. Each household h in each period t obtains a per period utility from consuming soft

drinks and an outside good:

$$U(c_t^h, v_t^h | \Theta^h) + \eta^h m_t^h$$

where c_t^h is the amount of soda consumed; v_t^h is a stochastic shock that changes the marginal utility from consumption; Θ^h is a vector of household-specific preferences; η^h is the marginal utility from consuming the outside good; and m_t^h is the consumption of the outside good. There are a total of J different products offered in the market, each is denoted by j . The total consumption of all soda products in period t by household h is $c_t^h = \sum_j c_{jt}^h$.

The utility of consumption, $U(c_t^h, v_t^h | \Theta^h)$, is a function of household h 's consumption in period t , c_t^h , and a shock to the marginal utility of consumption, v_t^h . Recall that consumption is equal to the total volume of all products consumed in the period. Following Hendel and Nevo (2006b), households do not obtain product-specific utility at the time of consumption. Instead, utility associated with each product is derived at time of purchase. Hence, the types of products available in storage do not affect consumption. Households receive a shock, v_t^h , in each period, which introduces randomness in households' need to consume and allows their consumption to depend on factors that are unobserved by the researcher. The shock influences the utility in the following way. If household h draws a high realization of v_t^h , then its marginal utility of consumption would be smaller than if a low realization had been drawn. Therefore, a high realization decreases households' need to consume and hence makes demand more elastic. In implementation, I follow previous literature in assuming the following functional form for the utility of consumption:

$$U(c_t^h, v_t^h | \Theta^h) = \alpha^h \log(c_t^h + v_t^h) \tag{1}$$

where the consumption level and the stochastic shock enter additively into the utility. Unlike previous literature, the marginal utility of consumption, α^h , enters into the model as a random coefficient and is household specific.⁸ As households cannot consume a negative amount of soda, the total consumption, c_t^h , is restricted to be non-negative.

Households have the ability to build inventories of goods, allowing them to smooth consumption

⁸Technically, the marginal utility of consumption is $\alpha^h / (c_t^h + v_t^h)$, which depends on both the consumption level and the consumption shock. For simplicity, I call α^h the marginal utility of consumption.

in times of high prices. However, this strategy is costly; the cost of the inventory at time t (i_t) is given by

$$F\left(i_{t+1}^h|\Theta^h\right) = \beta_1^h i_{t+1}^h + \beta_2^h \left(i_{t+1}^h\right)^2 \quad (2)$$

where β_1^h and β_2^h are household-specific random parameters. These parameters represent household h 's ability to store leftover products. Households' ability to take advantage of sales will depend on their marginal cost of inventory and thus on parameters β_1^h and β_2^h . Households' inventory costs influence their abilities to adapt to changes post tax. Households with lower costs of inventory are able to stock up on larger quantities of soda during sales. Households with higher costs of inventory will be more affected by the taxes because they will find it more costly to stock up. Following simple accounting rules, the end-of-period inventory, i_{t+1}^h , is equal to the beginning-of-period inventory, i_t^h , plus the total current purchase, s_t^h , minus the total consumption, c_t^h . That is, $i_{t+1}^h = i_t^h + s_t^h - c_t^h$. As households cannot consume more than what is available in their storage or make negative purchases, the total inventory, i_t^h , and purchase, s_t^h , are restricted to be non-negative.

The last piece of the flow utility, $G\left(p_{jst}, a_{jst}, \varepsilon_{jst}^h|\Theta^h\right)$, is associated with household h 's purchase of product j at volume s in period t . It is a function of prices, p_{jst} , promotional activities, a_{jst} , product preferences, Θ^h , and a random shock, ε_{jst}^h . In application:

$$G\left(p_{jst}, a_{jst}, \varepsilon_{jst}^h|\Theta^h\right) = \gamma_1^h p_{jst} + \gamma_2^h a_{jst} + \sum_j \mathbb{1}_{jt}^h \left(\zeta_j^h + \psi_j^h\right) s_{jt}^h + \varepsilon_{jst}^h \quad (3)$$

where $\sum_j \mathbb{1}_{jt}^h = 1$

Parameters γ_1^h and γ_2^h measure household h 's marginal utility of income and utility from any available promotional activities respectively. Variable a_{jst} is an indicator that denotes the presence of any features or displays. $\sum_j \mathbb{1}_{jt}^h (\zeta_j^h + \psi_j^h) s_{jt}^h$ represents household h 's intrinsic preferences for product j weighed by total volume purchased. Decomposing a household's product tastes into two sets of mutually exclusive product characteristics (i.e., $\sum_j \mathbb{1}_{jt}^h = 1$): a preference for the brand and a preference for whether the product is diet. Parameter ζ_j^h denotes household h 's persistent taste for the brand of product j and parameter ψ_j^h denotes household h 's persistent preference for whether product j is diet. Both are household-specific and, time-invariant random coefficients.

As sodas can be generally classified as either regular or diet, for identification purposes, ψ_j^h

measures how much more household h prefers regular over diet. If a household prefers regular over diet versions of the same brand then ψ_j^h is positive, and ψ_j^h is negative if the household prefers diet instead. Households' persistent preferences for the brand, ζ_j^h , and the diet type are especially important when it comes to studying the impact of the taxes, directly influencing whether the household would switch away from regular soda. Consider a household with strong preferences for regular Coke. This household may not switch away from its favorite product at all. Rather it would simply pay into the tax revenue. Alternatively, the household could purchase larger quantities of the product during a steep sale. These intrinsic tastes help determine how many people will alter their consumption behavior post tax. The last component, ε_{tsi}^h , denotes an idiosyncratic shock to household h 's product choice. Its distribution is discussed below along with other assumptions.

Household h 's objective is to maximize its discounted value of expected future utility, $V^h(\phi_t)$, for any state ϕ_t in any period t with respect to the consumption level c^h , the purchase volume s^h , and the product choice j^h , where the state space ϕ_t consists of current prices and promotions, the beginning-of-period inventory, the shock to utility from consumption, and the shock to utility for each product. Mathematically, household h 's problem in period $t = 1$ can be represented by the following infinite horizon maximization problem:

$$\begin{aligned}
V^h(\phi_1) &= \max_{\{c^h(\phi_t), s^h(\phi_t), j^h(\phi_t)\}} \sum_{t=1}^{\infty} \delta^{t-1} E \left[U(c_t^h, v_t^h | \Theta^h) - F(i_{t+1}^h | \Theta^h) + G(p_{jst}, a_{jst}, \varepsilon_{jst}^h | \Theta^h) | \phi_1 \right] \quad (4)
\end{aligned}$$

$$\text{s.t. } i_t^h, c_t^h, s_t^h \geq 0, i_{t+1}^h = i_t^h + s_t^h - c_t^h$$

where δ is the standard notation for the discount factor. This can equivalently be written as the

following Bellman equation:

$$\begin{aligned}
& V(\phi_1^h | \Theta^h) \\
= & \max_{\{c, s, j\}} \left\{ \left[\alpha^h \log(c^h + v^h) - \left(\beta_1^h i^{h'} + \beta_2^h (i^{h'})^2 \right) + \gamma_1^h p_{js} + \gamma_2^h a_{js} + \sum_j \mathbb{1}_{jt}^h (\zeta_j^h + \psi_j^h) s_j^h + \varepsilon_{js}^h \right] \right. \\
& \left. + \delta E \left[V \left((\phi^h)' | \phi_1^h, c^h, s^h, i^h, \Theta^h \right) \right] \right\} \tag{5}
\end{aligned}$$

where households maximize the sum of their current period flow utilities plus the discounted expected future values with respect to consumption, product, and volume purchases, conditional on the current state and unobserved persistent preferences.

Following Hendel and Nevo (2006b), product differentiation in this model takes place at the moment of purchase. Although households have intrinsic preferences for different products, these preferences and the differences between products affect households' behaviors exclusively at the store but do not give different utilities at the moment of consumption. Because consumptions and inventories are not observed, this specification avoids putting extra structure on consumption rates of different items in inventory. Furthermore, this assumption significantly decreases the state space. Instead of carrying the whole vector of inventories for each brand, only the total quantity in stock is tracked. Differentiation at purchase, represented by $\sum_j \mathbb{1}_{jt}^h (\zeta_j^h + \psi_j^h) s_{jt}^h$, captures the expected value of the future differences in utility from consumption. More specifically, the term $\zeta_j^h + \psi_j^h$ captures the expected utility from future consumption of one unit of product j at the time of purchase. As Hendel and Nevo (2006) state, this assumption is appropriate as long as 1) brand utilities are linear and 2) discounting is low. Condition 1 holds in this model, as can be seen from equation (4). Condition 2 is likely to be satisfied as well, given that decisions are made weekly and most purchases are presumably meant to be consumed in the relatively near future.

Additionally, in order to simplify the computational burden associated with estimating the model, I make the following assumptions, which are common in the literature:

Assumption 1. *Consumption shocks, v_t^h , are distributed independently and identically across*

households and over time.

In principle, serial correlation in v_t^h can be accommodated, but only with a significant increase in computational burden.⁹

Assumption 2. *Prices and advertising activities follow an exogenous first order Markov process:*

$$\Pr(p_t|p_1, \dots, p_{t-1}, a_1, \dots, a_{t-1}) = \Pr(p_t|p_{t-1}) \text{ and } \Pr(a_t|p_1, \dots, p_{t-1}, a_1, \dots, a_{t-1}) = \Pr(a_t|a_{t-1}).$$

More specifically, prices follow the AR1-process, $p_t = \alpha_0 + \alpha_1 p_{t-1} + v_t$, where v_t is distributed normally

with mean zero and advertising follows the AR1-process $\Pr(a_t = 1) = \frac{1}{1 + \exp(b_0 + b_1 a_{t-1})}$.

This assumption is often employed in the empirical literature in order to reduce the state space in the dynamic programming problem.¹⁰

Assumption 3. *The product choice shock, ε_{tsi}^h , follows a type I extreme value distribution and is independent and identically distributed across each household, period, and purchase.*

This assumption is commonly made in the literature and can be relaxed at the cost of significantly increased computational burden.

Assumption 3 allows me to calculate the probability of observing any particular brand-diet purchase, conditional on volume and unobservable persistent preference, according to the standard logit formula:

$$\Pr(j_t^h | s_t^h, \phi_t^h, \Theta^h) = \frac{\exp \left[\gamma_1^h p_{jst} + \gamma_2^h a_{jst} + \sum_j \mathbb{1}_{jt}^h (\zeta_j^h + \psi_j^h) s_{jt}^h + M(\phi_t^h, j^h, s^h | \Theta^h) \right]}{\sum_{j', s'} \exp \left[\gamma_1^h p_{j's't} + \gamma_2^h a_{j's't} + \sum_j \mathbb{1}_{jt}^h (\zeta_j^h + \psi_j^h) s_{jt}^h + M(\phi_t^h, j^{h'}, s^{h'} | \Theta^h) \right]} \quad (6)$$

where $M(\phi_t, j, s | \Theta^h) = \max_c \{ U(c_t^h, v_t^h | \Theta^h) - F(i_{t+1}^h | \Theta^h) + \delta E[V(\phi_{t+1}^h) | j_t^h, c_t^h, \phi_t^h, \Theta^h] \}$.

Similar to Hendel and Nevo (2006b), given the preceding model and assumptions, the optimal consumption, conditional on the volume purchased and household h 's persistent preferences, is not

⁹Under unobserved persistent preferences, households' consumption patterns could be correlated with their systematic preferences. It is therefore important that households' marginal utility of consumption, α^h , be allowed to differ across households. Once this has been controlled for, any remaining source of dependence would be present if households with different preferences have consumption shocks with different variances. However, it is unclear what would cause such a correlation, and any such effects are likely to have minor influences on the estimation results.

¹⁰This assumption implies that consumers use only the current period prices and promotions to predict future prices and promotions, which seems to be a reasonable approximation of the formation of consumers' price expectations. The main concern might be seasonal price fluctuations, when the probability of advertising increases. This concern should be ameliorated since I only use data from the 2nd to the 26th week of each year.

brand-specific. That is, $M(\phi_t, j, s | \Theta^h)$ is independent of brand choice. This implies that

$$\Pr(j_t^h | s_t^h, \phi_t^h, \Theta^h) = \frac{\exp[\gamma_1^h p_{jst} + \gamma_2^h a_{jst} + \sum_j \mathbb{1}_{jt}^h (\zeta_j^h + \psi_j^h) s_{jt}^h]}{\sum_{j', s'} \exp[\gamma_1^h p_{j's't} + \gamma_2^h a_{j's't} + \sum_j \mathbb{1}_{jt}^h (\zeta_j^h + \psi_j^h) s_{jt}^h]}. \quad (7)$$

From this equation, it is straightforward to see that a household's two purchase decisions, product j_t^h and volume s_t^h , are treated differently. Conditional on both the volume, s_t^h , and household h 's persistent preference, Θ^h , the product choices, j_t^h , are uncorrelated across periods. This split in the likelihood simplifies the state space as it allows me to keep track of only a single inventory variable instead of one inventory per product.

However, the same split in estimation as seen in Hendel and Nevo (2006b) is no longer feasible due to the persistent random tastes. As discussed in Aguirregabiria and Nevo (2013), the split of the likelihood in estimation relies on the assumption that taste parameters are conditionally independent of heterogeneity. Allowing for persistent random components, Θ^h , violates this assumption. The computation of the probability of product purchase j_t^h , conditional on volume s_t^h but unconditional on persistent preferences Θ^h requires the integration of the probability of Θ^h conditional on volume s_t^h , which is dynamic. As a result, this split cannot no longer be used in estimation.

The volume purchase decisions, s_t^h , depend on current consumption as well as past consumption levels, purchase decisions, initial inventories, and persistent preferences. Consumption and volume purchase decisions, in turn, are determined by households' optimal policies and are correlated over time. Hence, the probability of observing any sequence of volume purchased, $\Pr(s_1^h \cdots s_T^h | \phi_t^h, \Theta^h)$, depends on that household's optimal policy rule, which is the solution to the dynamic programming problem.

4 Estimation & Identification

The computational cost to estimating a dynamic demand model such as specified above is high. Some of the cost can be attributed to the large state space associated with the model: not only are all current prices and promotional activities of all products and sizes part of the state space, the current period inventories and consumption shocks are also state variables. This causes the calculation of the value functions underlying the dynamic model to be very time consuming. The computational

burden is exacerbated by the introduction of random coefficients. As these random coefficients need to be "integrated out," the value function needs to be calculated many more times than for the model without random coefficients. Combined, the computational burden makes the traditional fixed point approach to estimation infeasible. Consequently, I adopt a simulated maximum likelihood approach that exploits Importance Sampling techniques to mitigate the computational burden. I further reduce the computational burden by leveraging the assumption that household decision processes are mutually independent, allowing me to pre-calculate choice probabilities in a parallel fashion. Although the estimation follows roughly these methods, it has to deal with issues specific to this article. I start the discussion of the estimation by providing an overview of the estimation procedure; a more detailed description of the estimation details can be found in the technical appendix.

Estimation

The objective of the estimation is to maximize the likelihood of the observed sequence of actions with respect to the parameters that govern the distribution of households' tastes. Embedded in the likelihood are the infinite-horizon utility-maximization problems faced by every household in every period. Several key factors of the model complicate the dynamic programming problem. First, current period inventories are unknown as neither current consumption nor initial inventories are observed. The process of estimating current inventories becomes an initial conditions problem that is solved in the estimation procedure. Second, the distribution of households' unobservable persistent preferences is modeled as random coefficients; hence the likelihood has to be integrated over the joint distribution of these parameters. The necessity of having to solve households' dynamic programming problems at every guess of the parameters for every guess of the distribution of the random coefficients presents a significant computational burden.

Recall from the data section that only purchases made are observed. Therefore, the likelihood is a function of these purchasing decisions, d_t^h , which are decomposed into product purchases, j_t^h , and volume purchases, s_t^h (i.e., $d_t^h = \{j_t^h, s_t^h\}$). Assume for now that the econometrician observes not only households' purchases but also their initial inventories and unobserved persistent tastes. Furthermore, assume that households are independent from one another. Therefore, the probability

of observing the entire purchase sequence across all households can be written as:

$$L = \prod_{h=1}^H \int_{\mathbf{v}} \Pr \left(d_1^h \cdots d_T^h | \phi_1^h \cdots \phi_T^h, \Theta^h \right) dF \left(v_1^h \cdots v_T^h \right) \quad (8)$$

Decomposing household h 's purchase in period t into product purchases, j_t^h , and volume purchases, s_t^h , the likelihood can be rewritten as:

$$L = \prod_{h=1}^H \int_{\mathbf{v}} \left[\prod_{t=1}^T \Pr \left(j_t^h | s_t^h, \phi_t^h, \Theta^h \right) \right] \Pr \left(s_1^h \cdots s_T^h | \phi_t^h, \Theta^h \right) dF \left(v_1^h \cdots v_T^h \right) \quad (9)$$

Household h 's persistent preferences, Θ^h , and initial inventories, i_0^h , are not actually observed. To estimate the model we need to integrate equation (9) over the joint distribution of the random parameters and the initial inventories. Mathematically, this can be expressed as:

$$L = \prod_{h=1}^H \left\{ \int_{\Theta^h} \int_{i_0^h} \int_{\mathbf{v}} \left(\prod_{t=1}^T \Pr \left(j_t^h | s_t^h, \phi_t^h, \Theta^h \right) \Pr \left(s_1^h \cdots s_T^h | \phi_t^h, \Theta^h \right) \right) dF \left(v_1^h \cdots v_T^h \right) dF \left(i_0^h \right) dF \left(\Theta^h | \mu_{\Theta}, \Sigma_{\Theta} \right) \right\} \quad (10)$$

where μ_{Θ} and Σ_{Θ} denote the mean and variance-covariance matrix of the joint distribution of households' unobserved persistent preferences.

Households' initial inventories are unknown to the econometrician and hence need to be integrated out. Assume that we observe households' inventories at the beginning of period one. Once the optimal consumption and purchase choices are estimated, the end-of-period inventories for the first period can be calculated following simple accounting rules. By rolling this process forward we can obtain all inventories across all periods, implying that the inventory process boils down to an initial conditions problem.

There are multiple ways to estimate initial inventory. The standard approach is to use the estimated inventory distribution to generate the initial distribution. In practice, this procedure is implemented by starting at an arbitrary inventory level and using the first few weeks of observations in generating the distribution of inventories; Hendel and Nevo (2006b) take this approach. An alternative method is to use the model and observed households' purchase behaviors to explicitly estimate a distribution over the initial inventory. This approach requires an assumption on the

distribution of the initial condition. I assume initial inventories follow the distribution specified below.¹¹

Assumption 4. *The initial inventories, i_0^h , are distributed log-normal with mean μ_{inv} and standard deviation σ_{inv} .*

Households' persistent tastes are also unobserved. They are modeled as random coefficients that follow distributions whose parameters are estimated in the model. I assume the distribution follows the assumption below:

Assumption 5. *The random coefficients, $\Theta^h = \{\alpha^h, \beta_1^h, \beta_2^h, \gamma_1^h, \gamma_2^h, \zeta_j^h\}$, are distributed jointly log-normal with mean μ_Θ and variance-covariance matrix Σ_Θ . Parameter ψ_j^h , which measures the difference between tastes for regular over diet soft drinks, is distributed according to a normal distribution with mean μ_ψ and variance σ_ψ .*

The econometrician does not observe households' tastes, so the joint distribution of the random parameters is estimated. Many of the parameters are naturally bounded (e.g., the marginal utilities of consumption are non-negative) and hence are assumed to follow log-normal distributions. I further assume that the off-diagonal elements of the variance-covariance matrix are zero. This restriction can be lifted at the cost of an increased computational burden.

Calculating the likelihood, equation (10), is a two-step process. The first step solves the dynamic programming problem and finds the optimal policy rule that governs each household's decision process. The second step computes each household's choice probabilities. The econometrician is faced with significant computational difficulties in both steps, especially in the presence of unobserved persistent preferences. Three factors help overcome the technical challenges involved in the estimation routine. 1) Repeated observations on households with similar demographics are essential in pinning down the persistent tastes. Relying on aggregated store purchase data would imply employing methods similar to Berry, Levinsohn, and Pakes (1995), which would severely complicate the estimation process. 2) Prices are assumed to be independent of the error term conditional on the variables and controls included in the model. As discussed in section 2, this assumption

¹¹In implementation, I experiment with both methods and find that the results are robust to these specifications. For a more detailed comparison and discussion, please see the robustness checks documented in the Results section.

is reasonable given that my sample is relatively short and does not include times of particularly high demand. 3) Persistent preferences are treated as independent of the consumption and product purchase shocks in the model.

Even given these simplifying assumptions, common methods for addressing the dynamic programming problem and estimation routine are still infeasible. The large state space, encompassing not only current inventories and shocks but also the prices and promotions of all products, makes solving of the dynamic programming problem difficult. I overcome this problem by employing the Sieve Value Function Iteration process proposed by Arcidiacono et al. (2013). This method approximates the dynamic programming problem for combinations of observable states and interpolates the value function for unobserved combinations. It makes accurate approximations possible when the Bellman operator is evaluated at a subset of state space, thus not only speeding up the approximation routine but also allowing for the calculation of the value function even for large state spaces. Implementation details can be found in Technical Appendix A.1.

Calculating the choice probabilities for a given combination of household, household type, initial inventory, and parameter try is not problematic per se. The problem arises from repeating this process for all households, initial inventory guesses, and parameter tries in maximizing the likelihood function. Solving the maximization problem using the standard maximum likelihood estimator requires a prohibitive amount of time. I adopt a simulated maximum likelihood approach and use sampling techniques to cut down on computation burden. This has the advantage that for any parameter draw we can calculate the value function, solve the model, and compute the choice probability outside the maximization routine, which allows for easy parallel computing and hence speeds up the calculation. The specific sampling procedure I use is the Importance Sampling approach discussed in Akerberg (2009). This method reduces the number of times the dynamic programming problem needs to be solved in estimation, which also helps to decrease computational burden. Please see the Technical Appendix A.2 for a more detailed description of the estimation procedure as well as implementation details.

Identification

As discussed earlier, the identification of most of the parameters is straightforward. The only one that may cause concern is the persistent product preference for diet versus regular soda. The

identification of this parameter comes from variations in the relative prices across products and over time; here I show weekly price variations both across brands and within brands. As an example, table 5 presents the percentage of times I observe different prices offered in the same week between Coke and Pepsi conditioned on the type of product as well as different prices between regular and diet sodas of the same brand. Other brands follow a similar pattern.

From table 5, most of the time Coke and Pepsi offer different prices. In fact, the same price is offered at less than 20% of the time. Most of the price differences observed are between 10 to 25 cents. There are fewer price variations for products within the same brand. Regular and diet drinks of a given brand are priced differently about 30% of the time. Most of these price differences are small, mostly below 10 cents. However, large differences do occur. We see that about 8% of the time regular and diet sodas of the same brands have price differences of 10 to 25 cents. These price variations are used to identify households' taste parameters. Please see the technical appendix for a discussion of the other parameters.

5 Results

The estimation is carried out on a sample of 781 households from two suburban areas in the Midwest and New England. The households are divided into three different per capita income brackets: less than \$10K (low income), between \$10K and \$20K (middle income), and more than \$20K per capita (high income). The model and method can accommodate large variations in observable household characteristics. Here I account for income and household size only because the sample is largely homogeneous in ethnicity. With regard to products, I include all products listed in table 2. Furthermore, packaging of all types (i.e., bottles or cans) is included in the analysis.

Parameter Estimates

I estimate the demand model as stated by equation (4). Table 6 presents the mean and standard deviation governing the distribution of each random parameter. Estimates for different income brackets are listed in separate columns.

The differences between household types are clear. With regard to the marginal utility of income, there is a large difference between poor households and rich households. Poor households

value money significantly more than rich households. This result is consistent with decreasing marginal utility of income: as households' per capita income increases, money becomes less important for them. For the marginal utility of consumption, middle-income and high-income households are roughly the same, with values around 3; low-income households, on the other hand, value consumption of soft drinks much more, with an average marginal utility of around 4.4. This implies that, holding inventory constant across all income brackets, poor households would consume more soda than either middle-income or rich households. This corroborates findings reported in much of the research in public health regarding poor households consuming more high-calorie, low-cost foods (Drewnowski and Specter, 2004; Kumanyika and Grier, 2006). The cost of inventory has an opposite story from the previous two effects. As income increases, households' cost of inventory decreases. It is more costly for poor households to store products at home than for middle-income households and similarly more costly for middle-income households to store products than for rich households.

Households from different income brackets also differ by their persistent product tastes. Although estimates for different brands exhibit different patterns, brand loyalty for more prominent brands is generally increasing in income. For instances, brand preferences for Coca-Cola and Pepsi, the two dominant brands, are strictly increasing in income. High-income households like Coca-Cola more than poor households by over 40%. This pattern is less clear for other brands. For example, low-income and high-income households both have higher preferences for Sprite than middle-income households. Preferences for 7 Up, on the other hand, are strictly decreasing in income. Poor households prefer this brand much more than higher income households. The preference pattern for Mtn Dew is the opposite of that for Sprite, and the pattern for Dr. Pepper follows that of 7 Up.

The last parameter measures the difference between households' preferences for regular and diet soft drinks. If households prefer regular soda to diet soda, then the mean of this random parameter is positive. As table 6 shows, the means across all income groups are positive. Therefore, all households prefer regular soda to the diet versions. However, this effect is different across different groups and is strictly decreasing in income. Low-income households' preferences for regular soda are much stronger than those for middle- and high-income households. This result doesn't come as a surprise. Rich households tend to be more health conscious and tend to drink diet soda more than regular ones. As seen in table ??, diet soda captures a larger market share among high-income

households than among low-income households. The parameter estimate reflects this fact. This also has significant policy implications. As some tax proposals target only regular soda, poor households will be impacted more. Those households with strong preferences for regular soda may choose to pay more for it instead of switching away.

Implications

Long-run price elasticities measure consumers' responsiveness to permanent price changes. Their estimates determine the predictions of post-tax consumption patterns in sugar-sweetened soda and influence the resulting welfare loss. As the proposed policies predominantly target sugar-sweetened soft drinks, I discuss here the long-run elasticities of products affected, both for the aggregate category and for individual products. In table 7, I present the long-run own-price elasticity of regular sodas. I further decompose this aggregate category elasticity and present the long-run own-price elasticity estimates of regular soda by brand. Then following a similar fashion I present and discuss the long-run cross-price elasticities in table 8. For the full matrix of own-price and cross-price elasticities, see table 12 and 13 in appendix A.6.

The estimated long-run own-price elasticity of regular sodas is -0.5744 , meaning if the price of regular soda increases by 1% its demand will fall by about 0.57%. This suggests that the demand for regular soda is inelastic and households are not likely to switch out of them after small permanent price increases. Post tax, households will likely pay into the tax but will not alter their behavior. Hence, the proposed policies may not achieve their intended consequences.

In the above table, I also report the long-run own-price elasticity of regular soda for each individual brand considered in the analysis. Demand for each brand is somewhat elastic. For example, the long-run own-price elasticity of regular Coke is -1.20 . That is, if the price of regular Coke increases by 1%, holding all other prices constant, the demand for regular Coke will fall by 1.20%. This indicates that, on average, consumers are somewhat likely to substitute out of regular Coke and into other products if the price of regular Coke goes up. As we can see, several other brands are more elastic than Coke. These include Pepsi, 7 Up, and Dr. Pepper. Several brands are less elastic than Coke. For instance, demand for regular Mtn Dew seems very inelastic. For a 1% price increase, demand will decrease by 0.83%. One possible explanation is that Mtn Dew is a specialized soda. It is advertised as a sport/energy drink. Additionally, having the largest quantity

of caffeine, many consumers might rely on Mtn Dew to increase their work efficiency.

To understand the consequences of the proposed taxes is it important to know not only how elastic demand for regular sodas is but also how sensitive demand for diet sodas is to price changes in regular sodas. These sensitivities govern the substitution patterns post tax. Table 8 reports the long-run cross-price elasticities of demand for diet sodas at the category level and for each individual brand. As a result of permanent price hikes on all regular sodas, the long-run cross-price elasticity estimates here measure the responsiveness of demand for diet Pepsi, for instance, to a change in the price of, say, regular Coke. If the long-run cross-price elasticity for diet Pepsi with respect to regular Coke is low, a permanent price hike on regular Coke would not induce households to switch to low-sugar alternatives such as diet Pepsi¹². As documented in table 8, the overall demand for diet soda goes up by roughly 0.6% in reaction to a 1% price increase in all regular soda. This implies that households are not likely to substitute to diet soft drinks. This should not come as a surprise. As seen in table 7, the demand for the regular soda category is inelastic.

We see that the demand for diet Coke increases by 0.87% if the price of regular Coke increases by 1%. Other brands follow a similar pattern, with the exception of Dr. Pepper. The long-run cross-price elasticity of diet Dr. Pepper is 1.29, which implies that its demand increases by over 1%. The important number to take away here is the long-run own-price elasticity of demand for all regular sodas. With an estimate of -0.57, we should not expect to see a large decrease in the consumption of regular soda post tax.

To analyze the effects of the tax proposals, I use the distributions estimated from the model to simulate households of each type and income bracket. With the simulated households, I predict the consumption reductions and welfare losses for all income groups and at different pass-through levels.

¹²Accounting for storability has been shown to lead to larger and more positive long-run cross-price elasticities in some products (Hendel and Nevo 2006b). If the same holds for the soft drinks category, it would imply that a higher percentage of households will be predicted to switch to diet sodas after the taxes have been implemented compared to cross-price elasticity estimates obtained without accounting for storage. Table 8 shows that while this may be true, the cross-price elasticity of diet soft drinks, as a category, with respect to regular soft drinks is still relatively low.

Robustness Checks

In this section, I document two robustness checks. In the first one, I employ a simpler method of estimating the long-run price elasticities. In the second, I estimate the model using two different methods of controlling for the initial conditions problem. I discuss each in the paragraphs below.

As a robustness check and a clarification of my identification strategy, I implement a model in the spirit of Hendel and Nevo (2012). Hendel and Nevo (2012) build a dynamic demand model that is simple in design and implementation but still capturing inventory. This model provides a straightforward way of checking elasticity estimates. I implement two versions of this model. In the first one, I replicate Hendel and Nevo (2012) using a similar set of data. In the second, I expand the model to allow for regular and diet versions of soda.

In both versions, I use weekly scanner data at the store level from IRI, which covers 1503 stores across most major metropolitan areas. The data used spans the same 25 weeks in 2003 as that examined by the main model. To make the analysis comparable, I follow Hendel and Nevo (2012) in using 2-liter bottles of Coke, Pepsi, and generic brands. I implement the log-linear demand model with store fixed effects assuming that households have perfect foresight and that $T = 1$.

In the first version, my replication results confirm those documented in Hendel and Nevo (2012). I then extend the analysis to allow for diet and regular drinks. Table 9 below documents the results. I present the long-run own-price elasticity estimates for both regular and diet products of Coca-Cola and Pepsi. The left column presents elasticity estimates obtained using the full model discussed in the main text and the right column presents ones from implementing the simple model. As seen in table 9, the elasticity estimates from the simple model conform to those from the full model. This provides assurances for the full model and its implementation.

As briefly discussed in the estimation section, the initial inventory poses an initial conditions problem. A standard method of solving the initial conditions problem uses the estimated inventory distribution to generate the initial distribution. One possible implementation is to start the process at a randomly chosen starting point and use the first few weeks of observations to obtain the distribution inventories. Another approach to implement this is to estimate the distribution of the initial inventory explicitly. This approach requires the specification of a distributional assumption.

To obtain a good estimate using the first method, the panel would ideally need to be fairly

long, so many purchases per household are observed. Hendel and Nevo (2006b) take this approach. Their panel spans two years and they use the first eleven weeks to recover estimates for the initial inventories. The length of panel in this article is relatively short, twenty-five weeks. However, because soft drinks are a frequently purchased item, there are an adequate number of observations, even after using the first five weeks, to approximate the initial inventory distribution. I implement both methods and report the estimated parameters in table 10. The parameter estimates are comparable regardless of the approach adopted. Results are not sensitive to the approach used in estimating the initial inventory distribution. Parameter estimates for the other income groups show the same pattern.

6 Policy Study

There are a large variety of taxes currently being debated, some as ad valorem taxes and others as excise taxes. Here, I follow a previous study conducted by Brownell and Frieden (2009) from the Yale Rudd Center for Food Policy and Obesity. They analyze two different taxes: a 10% sales tax and a penny-per-ounce excise tax. They estimate the total decrease in sugary soda consumption assuming a 100% pass-through. That is, they assume the entire burden of the tax will fall on consumers. I analyze both taxes for four levels of pass-through: 100%, 75%, 50%, and 25%. For each tax, income bracket, and pass-through level, I present and discuss the predicted total decrease in consumption and the estimated total welfare loss.

Reduction in Consumption

Table 11 shows the estimated average reduction in households' consumption of regular soft drinks for both the 10% sales tax and the penny-per-ounce tax. In monetary terms, the 10% sales tax ranges from around 9 to 17 cents per 2 liters. The penny-per-ounce tax, on the other hand, is independent of the price and is just under 68 cents for every 2 liters of regular soda. For each tax and pass-through combination, I report the average estimated total reduction in consumption for each income bracket. I decompose the total reduction in consumption into two parts: 1) substitution from regular to diet sodas and 2) substitution from regular sodas to no purchase (or outside option).

For all taxes at all pass-through levels, poor households experience the largest reduction in

consumption and rich households experience the least. Recall from table 6 that poor households have the highest preferences for regular drinks and rich households have the lowest. This may seem like a contradiction. Because low-income households have stronger preferences for regular soda, one might expect to see a smaller decrease in their consumption of regular soda. However, recall that households' taste preferences are not the only factors that influence their consumption behavior. Poor households have the highest marginal utility of income. That is, they react most strongly to price increases. Without their tastes, we would expect them to decrease their consumption by the highest margin among all household income brackets. Their taste for regular drinks decreases that margin but is not enough to offset the entire burden of the price increase. Rich households, on the other hand, have the lowest marginal utility of income and the highest share of diet-loving households (63.56%). Not only do they have lower responsiveness to price changes, most of these households already consume a large quantity of diet drinks. Therefore, the tax does not impact them as much. We see the reduction in sugar-sweetened soft-drink consumption is highest for low-income households and lowest for high-income households.

For the 10% sales tax at 100% pass-through level, the largest predicted decrease in consumption of regular soda is 3.77%. According to Brownell and Frieden (2009), the expected decrease in the consumption is 7.9%. This more than doubles my findings. For the penny-per-ounce tax at 100% pass-through, the largest predicted decrease in consumption of regular soda is 9.69%. This tax imposes about 68 cents for every two-liter bottle of regular soda. This is an aggressive tax, considering a bottle of regular soda costs around \$1.50. Thus, it is not surprising that its impact is larger than the 10% sales tax. For this tax proposal, Brownell and Frieden (2009) provide a somewhat vague estimate. They predict a reduction in consumption above 10%. My estimates are somewhat lower and for some pass-through levels significantly lower than their estimate.

For both taxes at all pass-through levels, I decompose the reduction in regular soda consumption into two parts. The reduction comes from substitutions to diet drinks or substitution to no purchase (or outside good). The share of both of these factors across all household income brackets, taxes, and pass-through levels is roughly the same. The largest share of the reduction, about 76% to 97%, is replaced by diet drinks. The rest is replaced by no purchases (or outside options).

Welfare Loss

I use compensating variation in measuring households' welfare loss. Compensating variation can be interpreted as the dollar amount the government needs to pay to compensate a household in order for the household to reach its pre-tax utility level. In table 12, I show the compensating variation for each income bracket for both taxes and all pass-through levels. I further decompose the compensating variation into the dollar amount paid in taxes, compensation for utility loss in consumption decreases, and compensation for product switching. Note that the sum of the decomposition does not equal the total compensating variation. I do not present the compensating variation associated with different advertising schemes or the cost associated with inventory costs.

For the 10% sales tax, the overwhelming portion of compensating variation comes from the tax paid. For example, for poor households, the tax makes 89.85% of the total compensating variation at 100% pass-through. This result is intuitive. Households' demand for regular soda is not very elastic, they do not switch to diet drinks very much. Hence, they do not incur large deadweight losses. Indeed, we see that the non-tax payment portion of the compensating variation is only about twelve cents for poor households at 100% pass-through. The highest compensating variation comes from the high-income households, which switch to other products the least. The reason that they have to be compensated the most is that they are the least price sensitive. They do not tend to wait for sales and purchase products from more expensive stores. Therefore, they end up paying more into the system in an ad valorem tax.

In contrast, from the penny-per-ounce tax, we see that rich households have to be compensated the least among all households. The difference between these two taxes is that the penny-per-ounce tax is independent of price. Hence, because rich households switch to diet drinks less than their poor counterparts, they are compensated less. Poor households change their behavior more and hence have to be compensated for the deadweight loss created. Even so, the largest share of the compensating variation still comes from the taxes paid. Overall, this implies that the taxes are largely efficient in that they do not create large deadweight losses in the process.

An additional issue that policy makers should consider is how regressive the proposed taxes are. A tax is regressive if the poor pay more into it as a percentage of their income than the rich. Such a tax would shift the incidence disproportionately away from those with higher incomes to those

with lower income. In the present context, the proposed taxes are regressive if poor households pay more into the tax than their richer counterparts as a percentage of their income. This could happen if households continue to consume more regular soda post tax. Assume for a moment that the tax burden is entirely shifted to the consumers (100% pass-through). For the 10% sales tax, the lowest income group is predicted to pay \$1.24 in tax and the highest income group is predicted to pay \$1.39. In absolute terms poor households pay less into the tax than the rich households, but they also earn less than half the income of those in the highest income bracket. The lowest income group therefore pays proportionally more than the highest income group, making the tax regressive. For the penny-per-ounce tax, the lowest income group is predicted to pay \$4.48 in tax and the highest income group is predicted to pay \$3.83. Hence, this tax is also regressive. Results for all other pass-through combinations are qualitatively identical.

In combination, we see that the taxes proposed do not create large deadweight losses but are regressive. Most of the compensating variation comes from tax payment. However, poor households pay more in tax than rich households in proportion to their respective incomes.

7 Concluding Remarks

To accurately analyze the effects of public policies imposed on differentiated storable goods, I account for storability and unobservable persistent heterogeneity. I use the estimated distribution of household preferences to perform policy analyses. Following a study conducted by Brownell and Frieden (2009), I analyze two specific tax proposals: a 10% sales tax and a penny-per-ounce tax. I divide households into three income brackets and calculate the welfare effects of the sugar taxes at four pass-through levels: 25%, 50%, 75%, and 100%. The reduction in sugar-sweetened soda consumption is at most half of what was previously predicted. Moreover, I find that the policies generate a small deadweight loss but tax poor households more than their rich counterparts.

Tables and Figures

Table 1: Household Demographics

	Bracket 1	Bracket 2	Bracket 3
Demographics			
bracket definition (K)	<10K	[10K 20K)	≥ 20K
avg Income (K)	7.526	15.952	36.106
avg household size	2.571	2.426	1.865
number of observations	334	272	175
Soda Purchase			
avg weekly vol purchased (liters)	3.42	3.48	3.54
avg weekly dollars spend	3.11	3.14	3.24
avg annual total number of cola trips	12.12	11.46	10.44
Brand Purchase Shares			
Coke market share	50.13%	46.69%	54.14%
Pepsi market share	40.13%	40.80%	41.72%
Store Brand market share	9.74%	12.51%	4.14%
Diet vs. Regular Shares			
diet market share	39.49%	36.32%	51.55%
regular market share	60.51%	63.68%	48.45%

Table 2: Market Share in 2002 by Volume

	Product Name	Market Share	Cumulative Share
1	Coca Cola (Regular)	19.39%	19.39%
2	Pepsi Cola (Regular)	18.34%	37.73%
3	Coca Cola (Diet)	13.47%	51.20%
4	Pepsi Cola (Diet)	10.23%	61.43%
5	Mt. Dew (Regular)	5.83%	67.26%
6	Sprite (Regular)	5.14%	72.40%
7	Dr. Pepper (Regular)	4.23%	76.63%
8	7 UP (Regular)	2.58%	79.21%
9	Dr. Pepper (Diet)	1.80%	81.01%
10	Mt. Dew (Diet)	1.47%	82.48%
11	Generic Cola (Regular)	1.44%	83.92%
12	Sprite (Diet)	1.27%	85.19%
13	7 UP (Diet)	1.20%	86.39%
14	Generic Cola (Diet)	0.54%	86.93%
15	Generic Lemonlime (Regular)	0.48%	87.41%
16	Generic Lemonlime (Diet)	0.12%	87.53%
	Other	12.47%	100%

Table 3: Regression of Quantity Purchased

	Coefficient
duration since last sale	0.6901* (0.1044)
current price of product	-41.6410* (1.3694)
constant term	244.7918* (5.0503)

Note: *statistically significant at 99% level.

Table 4: Transitional Probability of Purchase

Prob(brand in per. 2 brand in per.1)	Prob(diet in per. 2 diet in per.1)
Prob(coke coke)	84.81%
Prob(pepsi coke)	15.19%
Prob(coke pepsi)	26.66%
Prob(pepsi pepsi)	73.34%
Prob(regular regular)	63.83%
Prob(diet regular)	36.17%
Prob(regular diet)	22.04%
Prob(diet diet)	77.96%

Table 5: Price Variation

Price Difference	Coke vs Pepsi	Diet vs Regular
	Share	Share
no difference	19.63%	72.39%
(\$0 \$0.1)	9.26%	13.27%
[\$0.10 \$0.25)	36.63%	7.56%
[\$0.25 \$0.50)	24.95%	3.96%
[\$0.50 \$0.75)	8.83%	0.94%
[\$0.75 \$1.00)	0.62%	0.94%
[\$1.00 \$2.00)	0.08%	0.94%

note: price differences are in units of 2 liters.

Table 6: Parameter Estimates

	Low Income		Middle Income		High Income	
	mean	std. dev.	mean	std. dev.	mean	std. dev.
Marginal Utility of Consumption	4.4393 (0.1489)	1.8642 (0.0918)	3.3301 (0.1977)	1.5400 (0.0431)	3.4376 (0.1875)	1.5489 (0.0861)
Cost of Inventory, Linear	1.4197 (0.2037)	0.8610 (0.0411)	1.3186 (0.1706)	1.2721 (0.0780)	1.0307 (0.2158)	1.3361 (0.0840)
Cost of Inventory, Quadratic	1.5990 (0.3255)	1.2820 (0.0658)	1.3311 (0.1305)	1.3847 (0.0557)	0.9052 (0.1697)	1.0375 (0.0567)
Marginal Utility of Income	1.1692 (0.2003)	1.0268 (0.1348)	1.1229 (0.1925)	1.7290 (0.1379)	0.8141 (0.2037)	0.8060 (0.1474)
Brand Preference - Coke	2.1679 (0.1787)	1.8355 (0.0844)	2.2653 (0.1449)	1.5267 (0.0557)	3.7865 (0.2107)	1.6845 (0.0785)
Brand Preference - Pepsi	2.5203 (0.1768)	1.6986 (0.0580)	2.4073 (0.1453)	1.4585 (0.0777)	3.2057 (0.1340)	2.1734 (0.1354)
Brand Preference - Sprite	1.8300 (0.1154)	0.5008 (0.0769)	1.2833 (0.2155)	0.5807 (0.0907)	1.9010 (0.1670)	0.5399 (0.0798)
Brand Preference - 7 UP	1.2048 (0.1519)	1.6907 (0.1634)	1.0289 (0.2200)	1.0305 (0.0993)	0.8415 (0.1078)	0.6455 (0.0663)
Brand Preference - Mtn Dew	1.0190 (0.3165)	1.0510 (0.0616)	1.1887 (0.2210)	1.6091 (0.0581)	0.8913 (0.2065)	0.8253 (0.0626)
Brand Preference - Dr. Pepper	1.0528 (0.2761)	1.1624 (0.0376)	1.0189 (0.1837)	1.0624 (0.0942)	0.7873 (0.1047)	1.4637 (0.0767)
Diet Preference - Regular	1.1107 (0.1701)	0.9753 (0.0832)	0.7605 (0.1897)	0.7991 (0.0686)	0.3025 (0.1087)	0.6193 (0.0756)

Note: The estimates are presented on top and the standard errors are in the parenthesis.

Table 7: Long-Run Own-Price Elasticities

Own-Price Elasticities	
All Regular Soda	-0.5744
By Brand:	
Regular Coke	-1.2016
Regular Pepsi	-1.2481
Regular Sprite	-1.1085
Regular 7 Up	-1.7685
Regular Mt. Dew	-0.8284
Regular Dr. Pepper	-1.2503

Table 8: Long-Run Cross-Price Elasticities

Cross-Price Elasticity	
All Diet Soda	0.6302
By Brand:	
Diet Coke	0.8660
Diet Pepsi	0.7779
Diet Sprite	0.6284
Diet 7 Up	0.5591
Diet Mt. Dew	0.7825
Diet Dr. Pepper	1.2923

Table 9: Elasticity Comparison

	Full Model	Simple Model
E_d Regular Coke	-1.20	-1.25
E_d Regular Pepsi	-1.25	-1.28
E_d Diet Coke	-1.00	-1.04
E_d Diet Pepsi	-1.22	-1.26

Table 10: Parameter Estimates of Low Income Bracket

	Estimation with Random Starting Values		Estimation with Distributional Assumption on Inventories	
	mean	std. dev.	mean	std. dev.
Marginal Utility of Consumption	4.4393	1.8642	4.3551	1.3652
Cost of Inventory, Linear	1.4197	0.8610	1.4132	0.8447
Cost of Inventory, Quadratic	1.5990	1.2820	1.5806	1.7604
Marginal Utility of Income	1.1692	1.0268	1.3431	1.5622
Brand Preference - Coke	2.1679	1.8355	2.2710	1.6382
Brand Preference - Pepsi	2.5203	1.6986	2.5352	1.6788
Brand Preference - Sprite	1.8300	0.5008	1.7346	1.0534
Brand Preference - 7 UP	1.2048	1.6907	1.1408	1.2365
Brand Preference - Mt. Dew	1.0190	1.0510	1.0734	1.0578
Brand Preference - Dr. Pepper	1.0528	1.1624	0.9110	0.7267
Diet Preference - Regular	1.1107	0.9753	1.0695	0.7992
Initial Inventory			2.6240	1.4680

Table 11: Post-Tax Reduction in Regular Soda Consumption

10% Sales Tax			
	Low Income	Middle Income	High Income
100% Pass-Through			
reduction in consumption	-3.77%	-3.40%	-3.26%
- decomposition: substitute to diet	75.82%	87.45%	95.86%
- decomposition: substitute to no purchase	24.17%	12.55%	4.14%
75% Pass-Through			
reduction in consumption	-2.69%	-2.62%	-2.37%
- decomposition: substitute to diet	75.99%	88.25%	96.23%
- decomposition: substitute to no purchase	24.01%	11.75%	3.77%
50% Pass-Through			
reduction in consumption	-2.03%	-1.59%	-1.54%
- decomposition: substitute to diet	76.12%	88.93%	96.86%
- decomposition: substitute to no purchase	23.88%	11.07%	3.14%
25% Pass-Through			
reduction in consumption	-0.99%	-0.79%	-0.69%
- decomposition: substitute to diet	76.82%	89.53%	97.01%
- decomposition: substitute to no purchase	23.17%	10.47%	2.99%
Penny per Ounce Tax			
	Low Income	Middle Income	High Income
100% Pass-Through			
reduction in consumption	-9.69%	-8.15%	-7.59%
- decomposition: substitute to diet	82.47%	88.92%	95.97%
- decomposition: substitute to no purchase	17.53%	11.08%	4.02%
75% Pass-Through			
reduction in consumption	-7.63%	-6.88%	-5.94%
- decomposition: substitute to diet	83.24%	88.98%	96.14%
- decomposition: substitute to no purchase	16.76%	11.02%	3.86%
50% Pass-Through			
reduction in consumption	-5.61%	4.97%	-4.18%
- decomposition: substitute to diet	83.98%	89.14%	96.67%
- decomposition: substitute to no purchase	16.02%	10.86%	3.33%
25% Pass-Through			
reduction in consumption	-3.85%	-3.36%	-3.01%
- decomposition: substitute to diet	84.15%	89.82%	96.84%
- decomposition: substitute to no purchase	15.85%	10.18%	3.16%

Table 12: Compensating Variation

10% Sales Tax			
	Low Income	Middle Income	High Income
100% Pass-Through			
Average Compensating Variation	\$1.38	\$1.34	\$1.46
- tax paid	\$1.24	\$1.23	\$1.39
- change in utility of consumption	\$0.04	\$0.03	\$0.02
- change in brand-diet specific utility	\$0.08	\$0.06	\$0.04
75% Pass-Through			
Average Compensating Variation	\$1.06	\$1.05	\$1.12
- tax paid	\$0.96	\$0.96	\$1.06
- change in utility of consumption	\$0.03	\$0.02	\$0.01
- change in brand-diet specific utility	\$0.06	\$0.05	\$0.03
50% Pass-Through			
Average Compensating Variation	\$0.73	\$0.71	\$0.75
- tax paid	\$0.65	\$0.66	\$0.72
- change in utility of consumption	\$0.02	\$0.01	\$0.01
- change in brand-diet specific utility	\$0.04	\$0.04	\$0.01
25% Pass-Through			
Average Compensating Variation	\$0.38	\$0.37	\$0.39
- tax paid	\$0.34	\$0.33	\$0.37
- change in utility of consumption	\$0.01	\$0.01	\$0.01
- change in brand-diet specific utility	\$0.02	\$0.02	\$0.01
penny-per-ounce Tax			
	Low Income	Middle Income	High Income
100% Pass-Through			
Average Compensating Variation	\$7.98	\$7.16	\$5.76
- tax paid	\$4.48	\$4.51	\$3.83
- change in utility of consumption	\$1.04	\$0.87	\$0.65
- change in brand-diet specific utility	\$2.38	\$1.72	\$1.22
75% Pass-Through			
Average Compensating Variation	\$6.54	\$5.99	\$4.67
- tax paid	\$3.93	\$3.93	\$3.22
- change in utility of consumption	\$0.83	\$0.69	\$0.49
- change in brand-diet specific utility	\$1.76	\$1.34	\$0.93
50% Pass-Through			
Average Compensating Variation	\$4.76	\$4.41	\$3.40
- tax paid	\$3.06	\$3.03	\$2.38
- change in utility of consumption	\$0.56	\$0.44	\$0.33
- change in brand-diet specific utility	\$1.12	\$0.91	\$0.65
25% Pass-Through			
Average Compensating Variation	\$2.67	\$2.43	\$1.81
- tax paid	\$1.76	\$1.74	\$1.32
- change in utility of consumption	\$0.32	\$0.22	\$0.16
- change in brand-diet specific utility	\$0.56	\$0.44	\$0.31

Figure 1: Weekly Cola-Drinks Sold (Percentage of Annual Total)

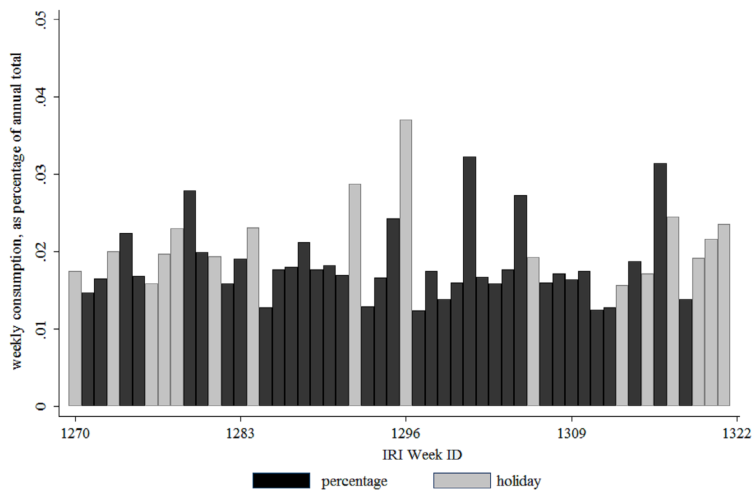


Figure 2: Sales Volume of Regular Coke (with and without Price Reduction)

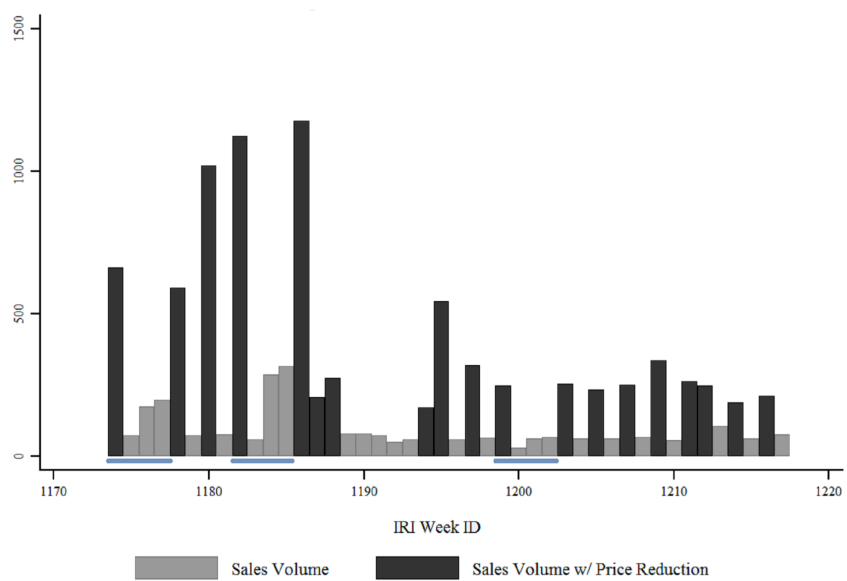
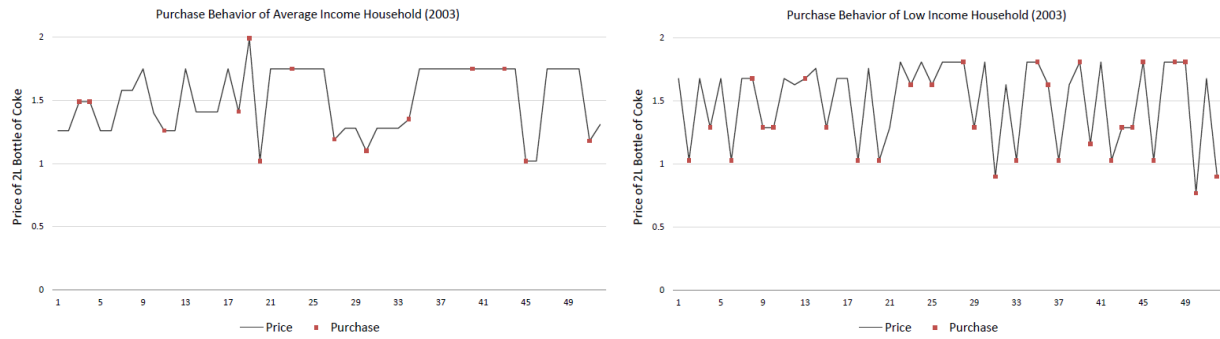


Figure 3: Purchase Behavior of Two Households (by income)



References

- Ackerberg, D. "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation." *Quantitative Marketing and Economics*, Vol. 7(2009), pp. 343-376.
- Aguirregabiria, V. & Nevo, A. "Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games." In D. Acemoglu, M. Arellano, and E. Dekel, *Proceedings of the Econometric Society World Congress*, 2013.
- Arcidiacono, P., Bayer, P., Bugni, F., & James, J. "Approximating High Dimensional Dynamic Models: Sieve Value Function Iteration." *Advances in Econometrics*, Vol. 31(2013), pp. 45-96.
- Andreyeva, T., Long, M., & Brownell, K. "The Impact of Food Prices on Consumption: a Systematic Review of Research on Price Elasticity of Demand for Food." *American Journal of Public Health*, Vol. 100(2010), pp. 216-222.
- Bajari, P., Fox, J., Kim K., & Ryan, S. "The Random Coefficients Logit Model is Identified." *The Journal of Econometrics*, Vol. 166(2012), pp. 204-212.
- Bajari, P. & Benkard L. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *The Journal of Political Economy*, Vol. 113(2005), pp. 1239-1276.
- Benitez-Silva, H., Hall, G., Hittsch, G., Pauletto, G. & Rust, J. "A Comparison of Discrete and Parametric Approximation Methods for Continuous-State Dynamic Programming Problems." Working Paper, 2000.
- Berry, S., Levinsohn, J., & Pakes, A. "Automobile Prices in Market Equilibrium." *Econometrica*, Vol. 63(1995), pp. 841-890.
- Bray, G., Nielsen, S., & Popkin, B. "Consumption of High-Fructose Corn Syrup in Beverages May Play a Role in the Epidemic of Obesity." *American Journal of Clinical Nutrition*, Vol. 79(2004), pp. 537-543.
- Bresnahan, T. "Departures from Marginal Cost Pricing in the American Automobile Industry: Estimates for 1997-1978." *The Journal of Industrial Economics*, Vol. 17(1981), pp. 201-277.

Brownell, K., Farley, T., Willett, W., Popkin, B., Chaloupka, F., Thompson, J., & Ludwig, D. "The Public Health and Economic Benefits of Taxing Sugar-Sweetened Beverages." *The New England Journal of Medicine*, Vol. 361(2009), pp. 1599-1605.

Brownell, K. & Frieden, T. "Ounces of Prevention - The Public Policy Case for Taxes on Sugared Beverages." *New England Journal of Medicine*, Vol. 360(2009), pp. 1805-1808.

Byrd, S. "Civil Rights and the "Twinkie" Tax: the 900-Pound Gorilla in the War on Obesity." *Louisiana Law Review*, Vol. 65(2004), pp. 303-305.

Conlon, C. "A Dynamic Model of Costs and Margins in the LCD TV Industry." Working Paper, Department of Economics, Columbia University, 2010.

Deaton, A. & Muellbauer, J. "An Almost Ideal Demand System." *The American Economics Review*, Vol. 70(1980), pp. 312-326.

Drewnowski, A. & Specter, S. "Poverty and Obesity: the Role of Energy Density and Energy Costs." *The American Journal of Clinical Nutrition*, Vol. 79(2004), pp. 6-16.

Erdem, T., Imai, S., & Keane, M. "Consumer Price and Promotion Expectations: Capturing Consumer Brand and Quality Choice Dynamics under Price Uncertainty." *Quantitative Marketing and Economics*, Vol. 1(2003), pp. 5-64.

Fletcher, J., Frisvold, D., & Tefft, N. "The Effects of Soft Drink Taxes on Child and Adolescent Consumption and Weight Outcomes." *Journal of Public Economics*, Vol. 94(2010), pp 967-974.

Food Research and Action Center. "Relationship Between Poverty and Overweight or Obesity." <http://frac.org/initiatives/hunger-and-obesity/are-low-income-people-at-greater-risk-for-overweight-or-obesity>.

Finkelstein, E., Trogon, J., Cohen, J., & Dieta, W. "Annual Medical Spending Attributable To Obesity: Payer And Service-Specific Estimates." *Health Affairs*, Vol. 28(2009), pp. w822-w831.

Goettler, R. & Gordon, B. "Does AMD Spur Intel to Innovate More?" *The Journal of Political Economy*, Vol. 119(2011), pp. 1141-1200.

- Gostin, L. "Law as a Tool to Facilitate Healthier Lifestyles and Prevent Obesity." *The Journal of the American Medical Association*, Vol. 297(2007), pp. 87-90.
- Gowrisankaran, G. & Rysman, M. "Dynamics of Consumer Demand for New Durable Goods." *The Journal of Political Economy*, Vol. 120(2011), pp. 1173-1219.
- Hall, K., & Jordan, P. "Modeling Weight-loss Maintenance to Help Prevent Body Weight Regain." *American Journal of Clinical Nutrition*, Vol. 88(2008), pp. 1495-1503.
- Harding, M., Leibtag, E., & Lovenheim, M. "The Heterogeneous Geographic and Socioeconomic Incidence of Cigarette and Beer Taxes: Evidence from Nielsen HomeScan Data." Working Paper, Cornell University, 2010.
- Hartmann, W. "Intertemporal Effects of Consumption and Their Implications for Demand Elasticity Estimates." *Quantitative Marketing and Economics*, Vol. 4(2006), pp. 325-349.
- Hartmann, W. & Nair, H. "Retail Competition and the Dynamics of Consumer Demand for Tied Goods." *Marketing Science*, Vol. 29(2010), pp. 366-386.
- Hendel, I. & Nevo, A. "Sales and Consumer Inventory." *The RAND Journal of Economics*, Vol. 37(2006a), pp. 543-561.
- Hendel, I. & Nevo, A. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica*, Vol. 74(2006b), pp. 1637-1673.
- Hendel, I. & Nevo, A. "Intertemporal Price Discrimination in Storable Goods Markets." *The American Economics Review*, Vol. 103(2013), pp. 2722-2751.
- Jacobson, M. & Brownell, K. "Small Taxes on Soft Drinks and Snack Foods to Promote Health." *American Journal of Public Health*, Vol. 90(2000), pp. 854-857.
- Johnson, R., Appel, L., Brands, M., Howard, B., Lefevre, M., Lustig, R., Sacks, F., Steffen, L., & Wylie-Rosett, J. "Dietary Sugars Intake and Cardiovascular Health: A Scientific Statement From the American Heart Association." *Journal of the American Heart Association*, Vol. 120(2009), pp. 1011-1020.

- Kaplan, M. "Taxing Sugar-Sweetened Beverages." *The New England Journal of Medicine*, Vol. 362(2010), pp. 368-369.
- Katan, M., & Ludwig, B., "Extra Calories Cause Weight Gain But How Much?" *Journal of the American Medical Association*, Vol. 303(2010), pp. 65-66.
- Keener, D., Goodman, K., Lowry, A., Zaro, S., & Kettel, L. "Recommended Community Strategies and Measurements to Prevent Obesity in the United States: Implementation and Measurement Guide." *Center for Disease Control and Prevention, Recommendations and Reports* Vol.58 (2009), pp. 1-26.
- Kumanyika, S., & Grier, S. "Targeting Interventions for Ethnic Minority and Low-Income Populations." *Future Child*, Vol. 16(2006), pp. 187-207.
- Mello, M., Studdert, D., & Brennan, T. "Obesity - The New Frontier of Public Health Law." *The New England Journal of Medicine*, Vol. 354(2006), pp. 2601-2610.
- Nevo, A. "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry." *RAND Journal of Economics*, Vol. 31(2000), pp. 395-421.
- Nielsen, S., & Popkin, B. "Changes in Beverage Intake Between 1977 and 2001." *American Journal of Preventive Medicine*, Vol. 27(2004), pp. 205-210.
- Osborne, M. "Estimation of a Cost-of-Living Index for a Storable Good Using a Dynamic Structural Model." Working Paper, Rotman School of Management, University of Toronto, 2013.
- Patel, K. "Essays in Industrial Organization," (Doctoral Dissertation). Northwestern University, ProQuest, UMI Dissertations Publishing, 3508326, 2012.
- Politis, D. N. & Romano, J. P. "Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions." *The Annals of Statistics*, Vol. 22(1994), pp. 2031-2050.
- Powell, LM., Slater, S., Mirtcheva, D., Bao, Y., & Chaloupka FJ. "Food Store Availability and Neighborhood Characteristics in the United States." *Preventive Medicine*, Vol. 44(2007), pp. 189-195.

Rust, J. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurche." *Econometrica*, Vol. 55(1987), pp. 999-1033.

Schulze, M., Manson, J., Ludwig, D., Colditz, G., Stampfer, M., Willett, W., & Hu, F. "Sugar-Sweetened Beverages, Weight Gain, and Incidence of Type 2 Diabetes in Young and Middle-Aged Women." *Journal of the American Medical Association*, Vol. 292(2004), pp. 927-934.

Smith, T., Lin, B., & Lee, J. "Taxing Caloric Sweetened Beverages." *United States Department of Agriculture, Economic Research Report*, Report No. 100(2010), pp. 33.

Sturm, R., Powell, L., Chriqui, J., & Chaloupka, F. "Soda Taxes, Soft Drink Consumption, And Children's Body Mass Index." *Health Affairs*, Vol. 29(2010), pp. 1052-1058.

Wang, Y.C., Coxson, P., Shen, Y., Goldman, L., & Bibbins-Domingo, K. "A Penny-Per-Ounce Tax On Sugar-Sweetened Beverages Would Cut Health and Cost Burdens Of Diabetes." *Health Affairs*, Vol. 31(2012), pp. 199-207.

A Technical Appendix

A.1 Implementation of Sieve Value Function Iteration

Sieve Value Function Iteration (SVFI), as introduced in Arcidiacono et al (2012), is a value function approximation method designed to make high-dimensional dynamic models tractable. For any guess of the model’s structural parameters, SVFI approximates the integrated value function with a sieve function, which is constructed with observable states. The closest approximation is found when parameters governing the sieve function come as close as possible to making the Bellman equation hold exactly, i.e. when the contraction mapping converges. The sieve space is relatively simple, so the minimization process is fairly straightforward to solve.

The method can be used in the context of single agent dynamic programming problems with either a finite or an infinite horizon setting. In the current context, I focus on the infinite horizon variant. To implement this version, the process boils down to estimating the parameters of the sieve function. Because the state space for the calculation at hand is continuous, the parameters are computed based on a large set of observed states. The value function values for states unobserved are then interpolated using the previously estimated parameters.

In the estimation routine, for any particular guess of the model’s structural parameters, sieve value function iteration is used to approximate the solution to the dynamic problem and to compute the policy functions as well as conditional choice probabilities. I embed this method for approximating the solution to the dynamic problem within the Importance Sampling estimation routine. There are a few choices that need to be made along the approximation process. The first is the construction of the sieve approximation using polynomials of the states and their interactions. Let x_t denote the set of relevant state variables, including inventories, consumption shocks, and prices and promotions of 16 products. Let \vec{p} denote the vector of prices for all 16 products and \vec{a} denote the vector of advertising, then:

$$x = [i, v, \vec{p}, \vec{a}]$$

To form the terms of the sieve space, I take all the state variables themselves and all pair-wise interactions. This makes a total of 595 terms. Following notation introduced in Arcidiacono et al (2012), let $w_j(x)$ denote the j^{th} term in the polynomial. Because the goal of SVFI is to approximate

the true value function as closely as possibly with the sieve, the problem of finding the optimal sieve approximation becomes:

$$\theta_n(x) = \rho(1)w_1(x) + \dots + \rho(n)w_n(x) = \rho_n W_n(x)$$

Let $u(x, c, j, s) = U(c_t^h, v_t^h | \Theta^h) - F(i_{t+1}^h | \Theta^h) + G(p_{jst}, a_{jst}, \varepsilon_{jst}^h | \Theta^h)$ denote the flow utility. The objective function, thus, becomes:

$$\hat{\rho}_n = \arg \min_{\rho_n} \sum_{x \in X} \left[\rho_n W_n(x) - \ln \left(\sum_{\{c, j, s\}} \exp(u(x, c, j, s) + \beta E[\rho_n W_n(x')]) \right) + \gamma \right]^2$$

Because the sieve approximation uses only polynomials, the expectation over next period's state space passes through only to the elements of the polynomial. To solve the minimization problem, I adopt the iterative least squares method with a convergence criterion of 1E-6.

A.2 Estimation Procedure for Importance Sampling

The estimation procedure can be roughly broken into five steps.

Step 1 (Initial Parameter Distribution) Importance Sampling methods require specifications of initial distributions over parameters, which are updated during the convergence process. The initial distributions can be chosen randomly at the discretion of the econometrician as long as they have a fairly large range that well covers the true parameter value.¹³ To start the estimation process, I predefine an initial joint distribution of the random coefficients, $g(\tilde{\Theta}^h | \mu_{\Theta}^0, \Sigma_{\Theta}^0)$, and a distribution over the initial inventory, $g(\widetilde{inv}_0 | \mu_{inv_0}^0, \sigma_{inv_0}^0)$, where the parameters that govern these distributions, $(\mu_{\Theta}^0, \Sigma_{\Theta}^0)$ and $(\mu_{inv_0}^0, \sigma_{inv_0}^0)$, are chosen arbitrarily. Then I draw N sets of parameter values and initial inventories according to the predefined distributions. I denote the randomly drawn parameters by $\tilde{\Theta}^h$ and the randomly drawn initial inventory by \widetilde{inv}_0 .

Step 2 (Product Choice) With the randomly drawn parameter tries, we can calculate the dynamic programming problem and compute the choice probabilities associated with each household's observed sequence of purchases. Recall that a household's purchase decisions are decomposed into a

¹³It is important to choose distributions that have a sufficiently large variance. Otherwise, certain parameters may be excluded from being drawn. This would lead to misestimation if the true parameters lay outside the range of the initial parameter draws. For a discussion of the how to choose initial distributions, refer to Akerberg (2009).

product choice, a static component, and a volume choice, which is a serially correlated component. Because product choices are static, conditional on the quantity purchased, they follow the standard logit formula. For each household, household type, and parameter draw the probability of observing any product choice is:

$$\Pr\left(j_t^h | s_t^h, \phi_t^h, \tilde{\Theta}^h\right) = \frac{\exp\left[\tilde{\gamma}_1^h p_{jst} + \tilde{\gamma}_2^h a_{jst} + \sum_j \mathbb{1}_{jt}^h \left(\tilde{\zeta}_j^h + \tilde{\psi}_j^h\right) s_{jt}^h\right]}{\sum_{j',s'} \exp\left[\tilde{\gamma}_1^h p_{j's't} + \tilde{\gamma}_2^h a_{j's't} + \sum_j \mathbb{1}_{j't}^h \left(\tilde{\zeta}_j^h + \tilde{\psi}_j^h\right) s_{j't}^h\right]}$$

Step 3 (Dynamic Programming Problem) Households' decisions on how much to purchase depend on their current level of inventory, which in turn depends on past consumption and purchase decisions. Hence these volume choices are correlated over time. These serially correlated components of households' purchases along with their consumptions are solutions to dynamic programming problems. The Bellman equation equivalent to households' infinite horizon maximization problem is:

$$\begin{aligned} & V\left(\phi_1^h | \tilde{\Theta}^h\right) \\ = & \max_{\{c,s,i\}} \left\{ \left[\tilde{\alpha}^h \log(c^h + v^h) - \left(\tilde{\beta}_1^h i^{h'} + \tilde{\beta}_2^h (i^{h'})^2 \right) + \tilde{\gamma}_1^h p_{js} + \tilde{\gamma}_2^h a_{js} + \sum_j \mathbb{1}_{jt}^h \left(\tilde{\zeta}_j^h + \tilde{\psi}_j^h\right) s_j^h + \varepsilon_{js}^h \right] \right. \\ & \left. + \delta E \left[V\left(\left(\phi^h\right)'\right) | \phi_1^h, c^h, s^h, i^h, \tilde{\Theta}^h \right] \right\} \end{aligned}$$

where households maximize the sum of their current period flow utilities plus the discounted expected future values with respect to consumption, product and volume purchases, conditional on current inventories, prices and promotions, and shocks.

I employ the sieve value function iteration method in solving for the optimal decision rule that governs consumption and purchase decisions. The sieve value function iteration method provides an approximation to the dynamic programming problem. It solves the value function at a large set of randomly chosen states and then extrapolates the result to other states. For more details, please refer to Arcidiacono et al. (2012).

With the computed optimal consumption and purchase decisions, the probability of observing

any sequence of volume purchases for a given household type and parameter draw is:

$$\begin{aligned} & \Pr \left(s_1^h \cdots s_T^h | \phi_t^h, inv_0^h, \tilde{\Theta}^h \right) \\ &= \prod_t \Pr \left(s_t^h | p_t, v_t^h, inv_t^h \left(s_1^h \cdots s_{t-1}^h, v_1^h \cdots v_{t-1}^h, inv_0^h \right), \tilde{\Theta}^h \right) \end{aligned}$$

After combining both components of households' purchase decisions, calculating households' choice probabilities is fairly straightforward.

Step 4 (Choice Probability) For each subset of household and parameter draws, the choice probabilities conditional on household h 's beginning-of-the-period inventory is given by:

$$\begin{aligned} & \tilde{L}_{sim}^h \left(d_1^h \cdots d_T^h | \phi_t^h, inv_0^h, \tilde{\Theta}^h \right) \\ &= \left[\prod_t \Pr \left(j_t^h | \phi^h, s_t^h, \tilde{\Theta}^h \right) \right] \Pr \left(s_1^h \cdots s_T^h | \phi_t^h, inv_0^h, \tilde{\Theta}^h \right) \end{aligned} \quad (11)$$

This conditional choice probability is equal to the probability of observing a sequence of product choices conditional on a sequence of observed volume purchases multiplied by the probability of observing that particular sequence of volume purchase. To remove some of the conditioning arguments, integrate the above over initial inventories and consumption shocks for each parameter draw:

$$\begin{aligned} & \tilde{L}_{sim}^h \left(d_1^h \cdots d_T^h | \phi_t^h, \tilde{\Theta}^h \right) \\ &= \int \int_{v \text{ } inv_0} \left(d_1^h \cdots d_T^h | \phi_t^h, inv_0^h, \tilde{\Theta}^h \right) dF \left(inv_0^h \right) dF \left(v_1^h \cdots v_T^h \right) \end{aligned}$$

Therefore, each household has a set of N choice probabilities, one for each parameter draw.¹⁴ The objective of the calculations is to find the parameters that maximize the total likelihood.

Step 5 (Log Likelihood Function) To compute the total likelihood function, randomly select a subset of the parameter draws along with their corresponding choice probabilities for each household. The simulated maximum likelihood function is a weighted average of all the randomly selected choice

¹⁴Recall, N is the number of parameter draws specified by the econometrician. In the estimation, I use $N = 5000$.

probabilities across all households. That is,

$$\tilde{L} = \prod_{h=1}^H \left\{ \frac{1}{\#\text{draws}} \sum_{\#\text{draws}} \left[\tilde{L}_{sim}^h \left(d_1^h \dots d_T^h | \phi_t^h, \tilde{\Theta}^h \right) \frac{h \left(\tilde{\Theta}^h | \mu_{\Theta}, \Sigma_{\Theta} \right)}{g \left(\tilde{\Theta}^h | \mu_{\Theta}^0, \Sigma_{\Theta}^0 \right)} \right] \right\} \quad (12)$$

where $h \left(\tilde{\Theta}^h | \mu_{\Theta}, \Sigma_{\Theta} \right)$ denotes the estimated distribution over random coefficients and over initial inventories. The component $h \left(\tilde{\Theta}^h | \mu_{\Theta}, \Sigma_{\Theta} \right) / g \left(\tilde{\Theta}^h | \mu_{\Theta}^0, \Sigma_{\Theta}^0 \right)$ can be seen as a weight on the likelihood. Parameters are allocated more weight if they explain households' decisions well. Note that as all the choice probabilities are calculated prior to maximization, $\tilde{L}_{sim}^h \left(d_1^h \dots d_T^h | \phi_t^h, \tilde{\Theta}^h \right)$ is nothing more than a large matrix of values. Moreover, as $g \left(\tilde{\Theta}^h | \mu_{\Theta}^0, \Sigma_{\Theta}^0 \right)$ is predefined, the maximization routine updates only the weight $h \left(\tilde{\Theta}^h | \mu_{\Theta}, \Sigma_{\Theta} \right)$.

Akerberg (2009) proposes using all parameter tries in the calculation of each household's choice probabilities (i.e., the same set of value function calculations is applied to each household). However, I find that the accuracy of the parameter estimate depends on the number of shared parameter draws across households. That is, estimates become less accurate if a single parameter draw is used across all households. I use a randomly drawn subset of parameters in computing the value functions for each household in the optimization routine.

A.3 Identification

I briefly and informally discuss the identification of the model. The identification of the static parameters follows the standard arguments (see Bajari et al., 2012, for a detailed discussion on the identification in random coefficient logit models). Parameters governing households' sensitivities to prices and advertising are identified from variations over time in prices and promotional activities. As price reductions are not perfectly correlated with products on feature or on display, the effects can be separately identified. The unobservable persistent household taste parameters are identified off households' purchase behaviors as prices and promotions change over time. For instance, consider the case where we observe a household consistently purchasing regular Coke even during most sales of regular Pepsi. The household switches to regular Pepsi rarely and only when the price of regular Coke is very high in comparison to regular Pepsi. For households like this, we can conclude that the household has a strong preference for regular Coke.

Similar to Hendel and Nevo (2006b) the identification of the dynamic parameters is more subtle. The complication arises from the fact that neither the consumption nor the inventory is observed. If both were observed, then the identification would follow standard arguments (see Aguirregabiria, 2005; Magnac and Thesmar, 2002; and Rust, 1996). To recover the utility of consumption and the cost of inventory, I rely on the relationship between current purchases, previous purchases, and prices. Given the process of prices and advertising, the preferences, storage costs, and initial inventories determine consumer purchases. For a given storage cost function, the utility of consumption determines the level of demand. That is, households with a high utility of consumption will purchase more on average. Hence, if we observe a household making large purchases frequently, we may conclude that the household has a high preference of consumption. For a given preference, storage costs determine interpurchase duration and the extent to which consumers can take advantage of price reductions. Here consider the case of no storage, or very high storage cost. Such households cannot exploit the price reduction opportunities and stock up on their preferred products. Their interpurchase durations are dependent only on current prices but not past prices. Hendel and Nevo (2006b) provide a more in-depth discussion of the identification of dynamic parameters.

A.4 Heterogeneity

Heterogeneity in product preferences or marginal impacts of prices and promotions can lead to interesting dynamics. For instance, consider a household with a high unobservable preference for regular Coke. This household is likely not to react to sales on other products and instead chooses to wait for regular Coke to go on sale. If the inventory of this household runs low, it may purchase small amounts of an alternative while waiting for the price of regular Coke to decrease. Due to the unobservable persistent preferences, such households are not likely to switch to other products even after the tax is imposed and simply stock up more on its favorite drinks during sales. From an earlier section, we know the data suggests that there is a segment of consumers who have strong preferences for certain product characteristic. Hence, to correctly predict the outcome of the taxes, it is important to allow such unobservable persistent shocks on product preferences and other factors.

Previous models were unable to accommodate this feature due to the computation intensive nature of the problem. As pointed out by Hendel and Nevo (2006b), in allowing random effects in product preferences or marginal impacts of prices, taste parameters will be solutions to dynamic

programming problems instead of being static parameters. That is, to compute the choice probabilities associated with the model unconditional on the unobservable type implies integrating over the distribution of the types. The probability of product choice conditional on size purchased and conditional on the type (of consumer) will still be independent of the various dynamic components. However, to compute the brand choice probability conditional on size (but unconditional on persistent components) we would have to integrate over the distribution of tastes conditional on the size purchased. Calculating this distribution requires solving the dynamic problem.

A.5 Own-Price and Cross-Price Elasticity

Tables 12 and 13 together document the full matrix of own-price and cross-price elasticities. Each cell corresponds to the percentage change in quantity of the column product subject to a 1% change in price of the row product. For instance, the second cell of the first row (0.8252) of table 12 represents a 0.82% increase in the quantity sold of Regular Pepsi subject to a 1% price increase in Regular Coke.

Table 13: Own-Price and Cross-Price Elasticity, Part I

	Regular Coke	Regular Pepsi	Regular Sprite	Regular 7 Up	Regular Mtn Dew	Regular Dr. Pepper
Regular Coke	-1.2016 (0.0781)	0.8252 (0.0554)	0.6252 (0.0854)	0.4932 (0.0932)	0.5098 (0.0862)	0.2076 (0.0105)
Regular Pepsi	0.9540 (0.0979)	-1.2481 (0.0675)	0.6196 (0.0952)	0.6932 (0.0808)	0.3296 (0.1268)	0.1192 (0.0612)
Regular Sprite	0.7011 (0.1312)	0.3672 (0.0615)	-1.1085 (0.1544)	0.8044 (0.1150)	0.1047 (0.0491)	0.4145 (0.0317)
Regular 7 Up	0.6011 (0.1498)	0.4698 (0.1369)	1.0219 (0.8511)	-1.7685 (0.1393)	0.2457 (0.1309)	0.1531 (0.0763)
Regular Mtn Dew	0.8891 (0.0547)	0.8663 (0.1596)	0.2339 (0.0528)	0.2075 (0.0433)	-0.8284 (0.1838)	0.1185 (0.0303)
Regular Dr. Pepper	0.3964 (0.0976)	0.3651 (0.1630)	0.3455 (0.0811)	0.2939 (0.1241)	0.1291 (0.0967)	-1.2503 (0.1986)
Diet Coke	0.9589 (0.2081)	0.7146 (0.1960)	0.2089 (0.0630)	0.2018 (0.0836)	0.4011 (0.0628)	0.1044 (0.0990)
Diet Pepsi	0.7110 (0.1910)	0.8118 (0.1293)	0.2164 (0.0274)	0.2027 (0.0194)	0.3033 (0.0984)	0.1017 (0.0478)
Diet Sprite	0.3771 (0.0148)	0.1479 (0.0690)	0.7159 (0.2559)	0.6033 (0.0951)	0.1024 (0.0291)	0.1888 (0.0726)
Diet 7 Up	0.3817 (0.0691)	0.2983 (0.0381)	0.8026 (0.1204)	0.8183 (0.0996)	0.2425 (0.0856)	0.2148 (0.0419)
Diet Mtn Dew	0.3701 (0.1018)	0.2483 (0.0862)	0.1081 (0.0128)	0.1028 (0.0137)	1.2668 (0.0137)	0.1044 (0.0413)
Diet Dr. Pepper	0.3294 (0.0452)	0.3817 (0.0807)	0.3916 (0.0501)	0.2174 (0.0829)	0.3183 (0.1082)	0.7902 (0.1329)

Table 14: Own-Price and Cross-Price Elasticity, Part II

	Diet Coke	Diet Pepsi	Diet Sprite	Diet 7 Up	Diet Mtn Dew	Diet Dr. Pepper
Regular Coke	0.8660 (0.1491)	0.6041 (0.1029)	0.4912 (0.0919)	0.3009 (0.0917)	0.4314 (0.0828)	0.2004 (0.0454)
Regular Pepsi	0.3050 (0.0537)	0.7779 (0.1591)	0.2016 (0.0829)	0.2003 (0.0826)	0.0905 (0.0525)	0.0801 (0.0242)
Regular Sprite	0.1033 (0.0384)	0.1099 (0.0836)	0.6284 (0.2318)	0.4016 (0.0862)	0.0818 (0.0776)	0.0901 (0.0756)
Regular 7 Up	0.2891 (0.0852)	0.2075 (0.0695)	0.6117 (0.0399)	0.5591 (0.1738)	0.1521 (0.0968)	0.2532 (0.1067)
Regular Mtn Dew	0.1037 (0.0884)	0.0918 (0.0269)	0.0805 (0.0523)	0.0834 (0.0153)	0.7825 (0.1375)	0.0008 (0.0091)
Regular Dr. Pepper	0.0103 (0.0029)	0.0204 (0.0013)	0.0011 (0.0035)	0.0081 (0.0025)	0.0058 (0.0002)	1.2923 (0.3086)
Diet Coke	-1.0905 (0.1523)	0.8804 (0.0375)	0.0603 (0.0189)	0.0345 (0.0102)	0.0132 (0.0061)	0.0112 (0.0075)
Diet Pepsi	0.8117 (0.0823)	-1.5739 (0.1244)	0.2084 (0.0909)	0.1110 (0.0253)	0.0054 (0.0017)	0.0041 (0.0022)
Diet Sprite	0.1777 (0.2712)	0.1782 (0.0927)	-1.2782 (0.2354)	0.7379 (0.0620)	0.0621 (0.0300)	0.0499 (0.0389)
Diet 7 Up	0.2549 (0.0281)	0.4502 (0.0080)	0.9529 (0.2524)	-1.8193 (0.0999)	0.1218 (0.0286)	0.1342 (0.0147)
Diet Mtn Dew	0.2928 (0.1576)	0.3470 (0.0130)	0.0173 (0.0130)	0.0130 (0.0094)	-0.7369 (0.2026)	0.0055 (0.0095)
Diet Dr. Pepper	0.4272 (0.1770)	0.4928 (0.1541)	0.2918 (0.1480)	0.1094 (0.0846)	0.1392 (0.0901)	-1.2185 (0.0880)