Published in final edited form as:
Am J Agric Econ. 2014 January 1; 96(1): 1–25. doi:10.1093/ajae/aat049.

Predicting the Effects of Sugar-Sweetened Beverage Taxes on Food and Beverage Demand in a Large Demand System

Chen Zhen,
Research Triangle Institute

Eric A. Finkelstein,
Duke-Singapore National University Medical School

James Nonnemaker,
Research Triangle Institute

Shawn Karns, and
Research Triangle Institute

Jessica E. Todd
U.S. Department of Agriculture, Economic Research Service

Abstract

A censored Exact Affine Stone Index incomplete demand system is estimated for 23 packaged foods and beverages and a numéraire good. Instrumental variables are used to control for endogenous prices. A half-cent per ounce increase in sugar-sweetened beverage prices is predicted to reduce total calories from the 23 foods and beverages but increase sodium and fat intakes as a result of product substitution. The predicted decline in calories is larger for low-income households than for high-income households, although welfare loss is also higher for low-income households. Neglecting price endogeneity or estimating a conditional demand model significantly overestimates the calorie reduction.

Keywords

obesity; sugar-sweetened beverage tax; censored demand; price endogeneity

---

The objective of this study is to predict the effect of sugar-sweetened beverage (SSB) taxes on demand for 23 categories of packaged foods and beverages, accounting for a third of daily energy requirement, and the associated changes in calories, fat, and sodium intake and consumer welfare. Obesity rates in the United States remain remarkably high at about 36% for adults and 17% for children and adolescents in 2009 to 2010 (Ogden et al. 2012). The costs of obesity are staggering; the medical costs alone are estimated to be at least $147

Send all correspondence to: Chen Zhen, Research Triangle Institute, 3040 Cornwallis Road, P.O. Box 12194, Research Triangle Park, NC 27709, Voice: (919) 541-7023, Fax: (919) 541-6683, czhen@rti.org.

Chen Zhen and James Nonnemaker are research economists, and Shawn Karns is an analyst, at Research Triangle Institute, NC. Eric Finkelstein is professor in Health Services and Systems Research at the Duke-National University of Singapore Graduate Medical School. Jessica Todd is an economist at the U.S. Department of Agriculture (USDA) Economic Research Service (ERS).

billion per year (Finkelstein et al. 2009). In efforts to curb the obesity epidemic, a number of public policy strategies, ranging from reshaping the food environment to changing relative food prices, have been proposed and debated at local and national levels. SSB taxes have been proposed by public health advocates and some policy makers as a way to discourage SSB consumption and to raise government revenue that could fund health-related initiatives (Brownell et al. 2009). Predicting the effect of SSB taxes and other targeted food and beverage taxes is difficult for at least two reasons. First, although various proposals on large SSB taxes in the magnitude of 0.25 cent per ounce or more have been discussed at the federal, state, and local levels, no jurisdiction has enacted such large taxes. The absence of existing large taxes on SSBs makes it impossible for researchers to examine ex post the impact of such taxes on nutrition and health using quasi-experimental designs such as difference-in-difference models. Second, increase in body weight is a result of an imbalance in energy intake and expenditure. To investigate the likely effect of SSB taxes on energy intake, demand interrelationships between SSBs and a plethora of foods and other beverages need to be determined to perform policy simulations. Estimation of consumer demand entails quantity and price data on food acquired from all sources, nutritional content data, and an adequate consumer demand model that handles a large number of goods. These requirements on data and method have proven to be quite challenging to meet.

Researchers studying the effects of taxing SSBs have undertaken two general approaches. The first exploits current state-level taxes on beverages to identify effects of these taxes on beverage consumption and weight status (e.g., Fletcher, Frisvold and Tefft 2010). This line of research has found little to no relationship between existing tax rates and body weight. This is not unexpected because current taxes on soft drinks are not designed to cut caloric intake, and the rates are trivial compared with retail prices and with taxes on tobacco, the commodity whose consumption in the United States has been significantly reduced by large taxes. For example, the average sales tax on soft drinks is about 5% among states that levy such taxes (Bridging the Gap 2011). In comparison, current federal and average state excise taxes on cigarettes combined are about $2.47/pack, or 44% of the retail price of cigarettes, with large variations across states (Campaign for Tobacco-Free Kids 2011). The second approach relies on estimating reduced-form demand (Finkelstein et al. 2010; 2013) or utility-theoretic demand systems (Zhen et al. 2011; Lin et al. 2011; Dharmasena and Capps 2012) and using elasticity estimates to simulate the impact of a large SSB tax on SSB consumption and weight outcomes. These studies found that SSB taxes would result in net reduction in energy intake from beverages.

Opponents of SSB taxes argue that a large SSB tax may cause consumers to switch to other foods with high caloric content. This slippage effect of SSB taxes is at least theoretically plausible and, if sufficiently large, may substantively offset the reduction in beverage energy. Finkelstein et al. (2013) examined this possibility by estimating reduced-form household demand equations for 19 SSB and food categories in a two-part model. They did not find evidence that an SSB tax would result in substitution to sugary foods and predicted that complementarity between SSBs and foods would further contribute to the reduction in total calorie intake (Finkelstein et al. 2013). That study did not investigate the effect of an SSB tax on consumer welfare.

Besides obesity, health risks associated with excessive quantities of fats and sodium consumption are also serious public health issues. Since the 1960s, scientific evidence linking excessive dietary intake of saturated fat and cholesterol to heart disease and cancer risks continues to accumulate (e.g., Van Horn et al. 2008). For decades, government, the medical community, and the food industry have been actively engaged in disseminating this information to the general public through official dietary guidelines, package labeling, and advertising. As a result, overall intake of saturated fat and total fat has fallen over time in most segments of the population (Ippolito and Mathios, 1995).

Prevalence of hypertension in the U.S. population stayed at 28 to 30% between 1999 and 2006 (Ostchega et al. 2008) and is estimated to be responsible for 395,000 deaths in the United States annually (Danaei et al. 2009). There is convincing evidence that high sodium intake is causally associated with risk for hypertension (Institute of Medicine [IOM] 2010a). Numerous initiatives launched between 1969 and the present have not succeeded in reducing the population’s sodium intake (IOM 2010b). The average American consumes far more sodium than the quantity the medical profession advises. The 2010 Dietary Guidelines for Americans recommends a daily sodium intake of no more than 2,300 mg for the general population and 1,500 mg/d for at-risk subpopulations. In comparison, the mean daily sodium intake was 3,513 mg for the general population and 3,264 mg for at-risk groups in 2005 to 2008 (Centers for Disease Control and Prevention [CDC] 2011). So far the literature on SSB taxes has focused on caloric intake due to the critical role of energy imbalance in causing obesity. However, if SSB taxes induce consumers to substitute more foods with high fat or sodium content for SSBs, overconsumption of fats and sodium may be an unintended consequence of SSB taxes that should be factored into the debate.

We make four contributions to the literature on consumer demand estimation and using targeted food and beverage taxes to reduce obesity. First, we directly confront the issue of corner solutions in micro data by estimating a censored demand system. Previous studies sidestepped this issue by either aggregating micro data into representative households (Zhen et al. 2011; Lin et al. 2011; Dharmasena and Capps 2012) or estimating a two-part model of reduced-form demand (Finkelstein et al. 2010; 2013). We use the utility-theoretic Exact Affine Stone Index (EASI) demand system, recently introduced by Lewbel and Pendakur (2009), to characterize household food and beverage preferences. One of the ex ante motivations for estimating a utility-theoretic demand system is to facilitate welfare analysis, which is essential for economic evaluation of any tax policy. To our knowledge, this is the first study extending the EASI model to censored data.

Second, we developed an approach that uses instrumental variables to account for potential price endogeneity arising from omitted variables and measurement errors in censored demand. We are not aware of previous work on micro flexible demand systems that instrument potentially endogenous prices. Our empirical results show that neglecting price endogeneity could overestimate the effect of an SSB tax on overall calorie reduction by over 100% on average.

Third, we demonstrate the importance of estimating an incomplete demand system with a composite numéraire good that represents all goods and services not individually modeled in the system, as opposed to a conditional demand system that treats the group expenditure (i.e., total expenditure on the included goods) fixed. Previous studies have documented the biases in welfare measures based on conditional demand systems (LaFrance and Hanemann 1989; Hanemann and Morey 1992). We demonstrate empirically that, in the context of beverage and food demand, a conditional demand model fails to identify important substitutions between SSBs and other foods and tends to overpredict the reduction in total calories by a large margin.

Fourth, using the household-based Nielsen Homescan consumer purchase panel, we explicitly modeled household purchases of a set of 23 food and beverage categories, including three major SSB categories: regular carbonated soft drinks (CSDs), sports/energy drinks, and fruit drinks. These foods and beverages account for over 50% of total household expenditures on packaged foods and 48% of at-home caloric intake, representing the largest set of products ever examined in the SSB tax literature.[1] By including foods as well as beverages in the analysis, we provide key insights into the potential health impacts of such a tax on obesity and other important nutritional outcomes such as fat and sodium intake.

We used the demand elasticity estimates to simulate the effects of a half-cent per ounce increase in SSB prices, induced by an excise tax, on household purchases of calories and two key nutrients: fat and sodium. Results based on the preferred demand specification that accounts for price endogeneity and includes a composite numéraire good suggest that, across households of all incomes, about one-half of the reduction in SSB calories is offset by increases in calories from other foods and beverages on average. The net effect of a price increase of this magnitude is a reduction of 7.9 kcal per capita per day. The hypothetical SSB tax is also predicted to increase daily per capita fat and sodium intakes by 0.2 g and 49.8 mg, respectively. These results suggest that unintended consequences of SSB taxes are possible, and the expected benefits of any proposed SSB tax has to be evaluated against its potential costs.

The remainder of this paper is structured as follows. In the next section, we present the censored incomplete EASI demand system and discuss the potential sources of endogeneity bias. The subsequent section focuses on the household scanner data used to model consumer demand and describes the construction of price indices and instrumental variables. This is then followed by discussions of the empirical results. Finally, the last section discusses our conclusions.

The Demand Model

Unlike many empirical studies of consumer food demand that invoke the weak separability assumption and estimated conditional demand systems, we model household preferences for the 23 foods and beverages using an incomplete demand system. We do this for two reasons. First, LaFrance and Hanemann (1989) proved that an incomplete demand system provides

[1]A possible exception is Okrent and Alston (2012), who developed an equilibrium displacement model of farm and retail food demand and supply that includes nine exhaustive but highly aggregated food and beverage categories to evaluate the effects of various agricultural and food policies on obesity. Among the policies examined in their study is a sugar tax on all foods and beverages. The authors grouped SSBs and non-SSBs into a single nonalcoholic beverage category and, therefore, did not examine the effect of a targeted SSB tax on obesity.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the exact and correct measures of welfare change, while the conditional demand model does not have this property (Hanemann and Morey 1992). Second, group expenditure in conditional demand equations cannot be correctly assumed exogenous in estimation or fixed in the counterfactual simulation of the SSB tax. The econometric bias could be corrected using proper estimation techniques (LaFrance 1991). However, holding group expenditure constant at the pre-tax level in the simulation will generate misleading predictions if expenditure responds to the tax. Moreover, in most policy analyses, income (not group expenditure) elasticities of demand are the quantities of interest. By including a composite numéraire good in an incomplete demand system to represent all other goods and services and substituting total household income for group expenditure, we obtain unbiased measures of welfare and unconditional predictions of demand responses to the SSB tax.

The Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980) and its variants have been widely used in empirical studies because they are utility based and have approximate versions that can be estimated by linear regression. Of the existing utility-theoretic models of aggregate beverage demand, Lin et al. (2011) and Dharmasena and Capps (2012) estimated a conditional nonlinear AIDS and a conditional linearized quadratic AIDS (Banks, Blundell and Lewbel 1997), respectively, while Zhen et al. (2011) used a dynamic nonlinear AIDS under two-stage budgeting to produce unconditional predictions of demand.

However, a limitation of the basic AID model is that its Engel curves are linear in real expenditures. Although the quadratic AIDS accommodates somewhat more flexible Engel curves, conventional demand system models cannot handle more complex Engel curves. This may not be a major shortcoming in studies that use aggregate data because much of the variation in income or total expenditures is smoothed out by aggregation over consumer units. The ability to fit more complicated Engel curves is valued in modeling demand at the micro level.

The EASI model not only shares all of the desirable properties of the AID model but also provides additional benefits. First, it is not subject to the rank three limitation of Gorman (1981) and allows the Engel curves to take arbitrary shapes (Lewbel and Pendakur 2009). We find that the ability to specify and test for Engel curves that are more flexible than quadratic ones has a substantive impact on price coefficient estimates, which ultimately affects the SSB simulation outcomes. Second, the EASI error term can be interpreted as unobserved consumer heterogeneity, while the AID residual does not have this interpretation. This is important for welfare studies that use consumer-level data because much of the demand variation cannot be explained by observed consumer demographics and price changes and is left in the error term.

A major obstacle in estimating a household-level continuous demand system model with highly disaggregated goods like ours is dealing with the significant presence of zero expenditure on some goods. In our data, the degree of purchase censoring at quarterly frequencies ranges from 9% for snacks to 71% for sports and energy drinks. The literature offers several approaches to censored data. Wales and Woodland (1983) constructed the likelihood function based on the Kuhn-Tucker conditions of maximization of a stochastic direct utility function subject to budget and nonnegativity constraints. Lee and Pitt (1986) extended the Kuhn-Tucker model to the dual indirect utility maximization and used virtual prices to identify corner solutions. However, because the likelihood function of both models consists of multiple integrals of censored demand regimes, this method quickly becomes computationally infeasible as the number of censored goods grows. A less structural alternative applies the Tobit estimator to estimate latent demand, which may take negative values. This approach circumvents the empirical difficulties of imposing nonnegativity restrictions under the Kuhn-Tucker framework and is the empirical strategy pursued in this study.

We modified the EASI incomplete demand system to account for censoring as follows:

$$w^*_{hit}=\sum_{j=1}^{J}a_{ij}\ln p_{hjt}+\sum_{r=1}^{L}b_{ir}y^r_{ht}+\sum_{k=1}^{K}g_{ik}z_{hkt}+u_{hit},\qquad h=1,\ldots,H;i=1,\ldots J-1;t=1,\ldots,T; \tag{1}$$

where w^*_{hit} is the latent budget share of category i in period t for household h; J is the number of goods; the J th good is the composite numéraire good; H is the number of households; y_{ht} is real household income; L is the highest order of polynomial in y_{ht} to be determined empirically; p_{hjt} is price index of the j th good; K is the number of exogenous demand shifters; z_{hkt} is the k th demand shifter with z_{h1t} being a constant; a_{ij}, b_{ir}, and g_{ik} are parameters; and u_{hit} is the residual. The latent share w^*_{hit} is related to observed budget share w_{hit} according to $w_{hit}\equiv\max\{0,w^*_{hit}\}$, where w_{hit} is calculated as the category expenditure divided by quarterly household income. To simplify analyses already complicated by censoring and endogeneity, we estimated equation (1) as an approximate EASI model, where y_{hit} is specified as the Stone price-deflated real income: $\ln x_{ht}-\sum_{j=1}^{J}w_{hjt}\ln p_{hjt}$, and x_{ht} is nominal quarterly household income.[2] Lewbel and Pendakur (2009) found that the linear approximate EASI and full nonlinear EASI models generate extremely close parameter estimates.

In addition to a constant, we specify the demand shifters z_{hkt} to include household size and 11 binary indicators: three Census regions; presence of female household head; female household head below age 35; female household head with college degree; black, Asian, other race, or Hispanic household head; and children.

Endogeneity

There are two potential sources of endogeneity in equation (1). First, because budget shares w_{hit} are used to construct y_{ht}, real income and its polynomials are endogenous. However, we have found this form of endogeneity to be numerically unimportant when an incomplete demand model is estimated. This is consistent with results reported in Lewbel and Pendakur (2009). Second, and more importantly, prices could be endogenous. In general, price endogeneity may be caused by supply-demand simultaneity, omitted variables, and

[2]In the fully nonlinear EASI model, y_{ht} is the affine transform of the Stone price-deflated real income. Unlike the Stone price index in the linear approximate AID model where it serves as an approximation to the true income deflator, the Stone price in the EASI model is the correct and exact deflator of income by design.

measurement errors. Although supply-demand simultaneity is a concern in studies of aggregate demand (e.g., Dhar, Chavas and Gould 2003), it may not be a major issue with micro data because individual household purchase decisions may not significantly affect market equilibrium prices. However, omitted variables and measurement errors could be important in micro data, because measurement errors and differences in household preferences cannot be averaged out as in aggregate data. Omitted variable bias may occur when households with preferences for certain foods and beverages are better than other households at finding lower prices for identical products; or when households who value quantity over quality would rather purchase less expensive private label products than more expensive name brands. This is analogous to the bias from using unit values as prices in consumer demand models (Cox and Wohlgenant 1986; Deaton 1988).

This quantity–quality trade-off within a food or beverage category is accounted for, to a certain degree, by our use of superlative Fisher ideal price indices calculated based on brand-level prices and quantities. Diewert (1976) proved that the Fisher ideal price index is consistent with a second-order approximation to an arbitrary twice-continuously differentiable linear homogenous cost function. This means that we could use Fisher ideal price indices to measure category-level price variations without explicitly modeling within-category product demand. However, because each Fisher ideal price index is only an approximation to the unknown true cost function, additional bias reduction measures are desirable.

Just like any social economic datasets used by economists, Homescan data has a degree of measurement error. Zhen et al. (2009) found discrepancies in reported expenditures between Homescan and the Consumer Expenditure Survey with the largest differences found in foods without a Universal Product Code (UPC). Although the Consumer Expenditure Survey contains reporting errors of its own and should not be the gold standard for expenditure comparison, the study by Einav, Leibtag and Nevo (2010) comparing purchase transactions recorded in Homescan and in a large grocery chain's database of the same households indicated that Homescan prices are measured with errors. Their simple illustrative example suggests that neglecting measurement errors in Homescan prices has the potential to substantively bias demand elasticity estimates.

The direction of the price endogeneity bias depends on which source of endogeneity dominates. Although measurement error, like demand-supply simultaneity, tends to attenuate coefficient estimates, the omitted-variable bias may have the opposite effect. Including observed demographic variables as additional regressors in budget share equations does not eliminate omitted-variable bias caused by unobserved household heterogeneity. Therefore, we used instrumental variables to account for the omitted-variable bias and measurement errors.

Data and Construction of Variables

We used household food purchase data from the 2006 Nielsen Homescan panel, a large national consumer panel maintained by the Nielsen Company.[3] Geographically, Homescan covers 52 Nielsen markets and 9 remaining areas in the contiguous United States (see figure

1 for market coverage). Each household was provided with a handheld scanner and instructed to scan the UPCs of products purchased at retail outlets. The household was also required to record purchase quantities and coupons used and to identify the retailer from which the product was purchased. Expenditures were recorded in one of two ways. If the purchases occurred at stores affiliated with Nielsen's scanner data program, the household did not need to provide prices or expenditures. Instead, the price and, hence, expenditure on a particular UPC were imputed by Nielsen using the store-provided average price of the UPC from the same week. If the purchase was made at an outlet that did not share its sales information with Nielsen, the household was asked to record dollars spent on the item. This information was then uploaded through the Internet or a landline phone to Nielsen on a weekly basis. Each year, Nielsen prepares a dataset, called the static panel, which includes transactions by households who consistently reported purchases during at least 10 months of the previous calendar year. This static panel is the data used in most academic studies, including ours.

In the static panel, each transaction record contains data on quantity purchased, dollars paid, the UPC number, and product attributes such as package size, brand, and Nielsen-assigned product module code. Although Homescan does not collect information on the nutrition content of products, the UPCs can be linked with third-party nutrient databases to create a nutrition profile of each household's food purchases. For this purpose, we acquired a UPC-level nutrient database from Gladson Interactive that contains information on calorie, fat, and sodium content per unit of product by UPC.

Compared with household scanner data, conventional paper-based expenditure surveys such as the Consumer Expenditure Survey are not suited for our purposes for two reasons. First, most expenditure surveys only collect category-level expenditure and sometimes quantity data. Studies using these data are often forced to approximate prices by unit values and are vulnerable to the unit value bias. Second, targeted food and beverage taxes focus on narrowly defined food and beverage products. Conventional expenditure surveys lack the specificity needed to distinguish otherwise similar products only differentiated by their nutrient content. For example, the Consumer Expenditure Survey does not separate diet soft drinks from regular soft drinks.

We defined 23 categories of packaged foods and beverages in Homescan that are either significant sources of energy, fat, or sodium or may be close substitutes or complements to SSBs. Most of the included foods and beverages are convenience foods that require little to no preparation. Many of these foods have been analyzed individually or as groups in the economics literature because of their importance in the food sector.

The excluded at-home foods include packaged ingredients (e.g., oil, baking mixes, and flour) that require nontrivial preparation and fresh foods (e.g., fresh fruits, vegetables and meats). Although Nielsen maintained a fresh foods panel of about 8,000 households that reported packaged food as well as random-weight food purchases through 2006, findings by Zhen et al. (2009) suggest that random-weight foods might be more vulnerable to

---

[3] In 2010, Nielsen merged its Homescan panel with SymphonyIRI Group's consumer panel to form the National Consumer Panel.

underreporting in Homescan than packaged foods. To be able to use a larger sample, reduce measurement errors, and keep the number of food categories manageable, we excluded most not ready-to-eat packaged foods and random-weight foods that did not have a UPC. The implication of this approach is that the predicted effects on calories, fat, and sodium may not be generalized to the excluded foods.

To obtain nutrient data for the selected food products, we first mapped each category with a set of Nielsen product modules. Next, we matched the Gladson nutrient data with foods or beverages in these modules by UPC. The Gladson nutrient data matched approximately 30% of the UPCs in the 23 categories in Homescan, which account for roughly 63% of the dollar sales for the 23 categories combined. To impute nutrient content of unmatched Homescan products, we calculated the mean nutrient value per unit (e.g., kcal/oz) of the matched products by Homescan descriptive fields including brand, product, type, and flavor. These mean nutrient values were then mapped to the unmatched Homescan products by these descriptive fields. Finally, nutrient values for the remaining unmatched Homescan products were filled in with values for similar products in the U.S. Department of Agriculture (USDA) National Nutrient Database for Standard Reference (USDA 2009).

The full 2006 Homescan static panel contains 37,786 households, and we aggregated household purchases into quarterly totals. To reduce the incidence of price outliers, we dropped households reporting quarterly category-level unit values (i.e., expenditure divided by quantity) that lie outside five standard deviations from the category-level means. Households residing outside the 52 Nielsen markets were also excluded from our analysis. Using 185% of the 2006 federal poverty line as the low-income group cutoff, the assembled analysis dataset contains 110,560 quarterly observations from 7,936 low-income and 19,704 high-income households.

Table 1 presents per capita purchase quantities, expenditures, energy, fat, and sodium content by food/beverage category and income group. Among the 23 foods and beverages, cakes and cookies and snacks (including salty snacks and other snacks) are the two largest sources of energy. Snacks are the largest source of fat, while the largest quantities of sodium come from canned vegetables. Assuming a 2,000 kcal/day reference diet and no food waste, the 23 foods and beverages provide about 32% of daily energy requirement. Because foods away from home account for about a third of total energy intake for an average person (Guthrie, Lin and Frazao 2002), about 52% of at-home food energy would come from store-purchased foods not individually modeled in our demand system. In other words, the composite numéraire good contains all foods away from home, 52% of at-home foods, as well as all other goods and services.

For every food or beverage category, low-income households paid less per ounce than high-income households. Across all categories, the unit value for low-income households is about 8% lower than that for high-income households on average. As documented in previous research (e.g., Broda, Leibtag and Weinstein 2009), this lower cost for low-income households is achieved by buying lower-priced brands and private-label products, products on sale, and products at supercenters and discount stores. Comparing per-ounce quantities of energy, fat, and sodium by category between income strata, low-income households' foods and beverages do not always contain more or less calories or nutrients than those of high-income households. However, when averaged across all categories, foods and beverages purchased by low-income households have 0.37 kcal more energy, 0.05 g more fat, and 1.39 mg more sodium per ounce than those bought by high-income households.

Category-level Price Indices and Instrumental Variables

Each food and beverage category comprises a large number of brands that individually account for tiny shares of the market.[4] To reduce the number of brands in each category price index, we aggregated all nonprivate-label brands with less than 0.5% of the category market volume share into fewer composite brands by product module and all private-label brands into fewer private-label brands by product module. Not all households purchased all brands in all quarters. To impute missing brand-level prices for nonpurchasing households, we regressed observed brand-level prices on market, brand, and quarter indicators; the interactions between brand and Census region indicators, between brand and quarter indicators, and between quarter and market indicators; and household demographics, including presence of female head; female head under age 35; college-educated male and female heads; and children, marital status, race, access to the Internet, income, and income squared. The regression was separately estimated for each food and beverage category. The predicted prices were then used to replace missing prices in the construction of price indices.

The Fisher Ideal price index for category (j=1,…, J−1) was calculated as

$$p_{hjt} = \sqrt{\frac{\sum p_{kht} q_{k0}}{\sum p_{k0} q_{k0}} \frac{\sum p_{kht} q_{kht}}{\sum p_{k0} q_{kht}}} \quad (2)$$

where p_{kht} and q_{kht} are the price and purchase quantity of brand k for household h in period, respectively, and p_{k0} and q_{k0} are the base price and base quantity of k set at their national means in 2006. The $\Sigma p_{k\,0} q_{kht}$ term becomes zero if demand for category j by household h is censored in period t, in which case we substitute $\Sigma p_{kht}/\Sigma p_{k0}$ for $\Sigma p_{kht} q_{kht}/\Sigma p_{k0} q_{kht}$ in equation (2).

The Bureau of Labor Statistics (BLS) reports consumer price index for the four Census regions and some areas. The BLS price index is normalized to 100 at the base period for each area. It allows cost-of-living comparisons over time within an area but not across areas. Therefore, the BLS consumer price index cannot be used to represent the price of the composite numéraire good in our demand model, which uses time-series and cross-sectional data. Instead, we calculated the price index for the numéraire good using quarterly prices from the Council for Community and Economic Research (C2ER) for 56 goods and services items[5] in more than 300 U.S. urban areas. A quarterly Laspeyres cost-of-living index[6] was first constructed for all C2ER urban areas using 2006 national average prices as the base and

[4]There are three levels of aggregations in Homescan. The lowest is the UPC, followed by product module and then by product group. If more than one UPC is associated with a brand name, the brand could appear in more than one product module. In this study, a unique brand is defined as a unique pair of brand name and product module.

[5]These goods and services are classified into six C2ER consumption categories: grocery, housing, utilities, transportation, health care, and miscellaneous goods and services.

[6]The C2ER data cannot be used to build a Fisher Ideal index because quantity data are not collected.

item weights derived by C2ER from the 2006 Consumer Expenditure Survey. A typical C2ER urban area corresponds to a Metropolitan or Micropolitan Statistical Area (MSA) or one or a group of counties. Each Homescan household was matched to a C2ER urban area based on census tract number. A direct match was obtained if the household lives inside an urban area. Households outside all C2ER urban areas were assigned to their nearest urban areas based on the Euclidean distance between the household's tract and the urban area boundary. Let CPI_{ht} denote the Laspeyres cost-of-living index for all goods and services for household *h* in period *t*. The price index for the numéraire good was obtained by solving

$$\ln CPI_{ht} = \sum_{j=1}^{J} w_{hjt} \ln p_{hjt} \text{ for } p_{hJt}$$

(Wohlgenant 1989, p.172).

The Homescan data contain the census tract number of each household's residence. Using ESRI's ArcGIS version 10.0, we calculated Euclidian distances from the centroid of each tract to the centroids of all other tracts within a Nielsen market. The instrument for each category-level household-specific price index was calculated as the weighted average of the price indices of the same food/beverage category for all other households, excluding those living in the same census tract as the target household, in the same Nielsen market and quarter. The inverse of the Euclidian distance between the target household and everyone else was used as the weight. The instrument for the price of the numéraire good was similarly constructed by averaging the prices of neighboring C2ER urban areas located in the same Nielsen markets.[7]

The concept of using prices of adjacent locations to instrument endogenous prices originated from Hausman (1997). The identifying assumption is that, after controlling for mean household valuations of foods and beverages and household demographic effects, household-specific demand shocks and measurement errors are independent across households in different tracts. This type of instrument is quite useful in overcoming endogeneity bias when researchers lack supply-side variables that possess the degree of specificity required to identify variation in relative prices of highly disaggregated goods. Exclusion of households residing in the same tract as the target household from calculation of the price instrument is intended to avoid the situation when prices faced by households in a tract are affected by common demand shocks.

The instrument for real income y_{ht} was calculated as

$$\tilde{y}_{ht} \equiv \ln x_{ht} - \sum_{j=1}^{J} \overline{w}_j \ln \tilde{p}_{hjt}, \quad (3)$$

[7]The average number of C2ER urban areas in a Nielsen market is 12.4. Using C2ER prices to construct the price of the numéraire good for low-income households creates additional measurement errors because C2ER prices are collected to reflect costs of goods and services for households in the top quintile of the income distribution. Unlike random price reporting errors, the latter errors are caused by differences in preferences between low- and high-income households. To some extent, the income variable and its polynomials in the demand model account for this bias.

where $\tilde{y}_{ht}$ is the instrument for y_{ht}, $\bar{w}_j$ is average budget share of the j th good across all households, and $\tilde{p}_{hjt}$ is the Hausman-type instrument for p_{hjt}.

Estimation and Results

We extended Amemiya's generalized least squares (AGLS) estimator for a single censored equation (Amemiya 1979; Newey 1987), which is efficient among a class of limited information estimators, to the system of censored incomplete EASI demand equations to account for endogenous prices and real income.[8] To our knowledge, this is the first extension of AGLS estimation to a system of censored equations. Therefore, we present the estimation details and SAS codes in the supplementary online appendix to facilitate potential applications of this estimator in future consumer demand studies.

To determine the proper degree of income polynomials, starting from L= 2 we added one higher degree of polynomial at a time and tested the joint significance of the b_{iL} coefficients by minimum distance (Wooldridge 2002, p. 444). Under the null that the Lth degree of polynomial is excludable, the test statistic is asymptotically distributed as $\chi^2(J-1)$. At L=5 the test statistic is 157.1 with a p-value < 0.00. We also estimated a model with L= 6 but experienced nonpositive definite variance-covariance matrices for some Tobit equations as the result of an ill-conditioned data matrix containing high degree polynomials.[9] Therefore, we determined that a fifth polynomial in y_{ht} is sufficient to capture the curvature of the Engel curves. Interestingly, Lewbel and Pendakur's (2009) also found L= 5 to be appropriate in their analysis of Canadian household demand for much more aggregated categories of goods. Figure 2 plots the Engel curves for CSDs, milk, bread, and juice at L=5 for a two-person non-Hispanic white household in the Northeast with a college-educated female head over age 35 and no children. An inspection of figure 2 suggests that the Engel curve shapes are too complicated to be adequately represented by a linear or quadratic function.

An important question that we address is whether household-level prices can be treated as exogenous in micro demand system models. For comparison, we estimated the approximate incomplete EASI model assuming price exogeneity.[10] We performed the exogeneity test developed by Durbin (1954), Wu (1973), and Hausman (1978) (hereafter, DWH) to determine whether parameters estimated assuming exogenous prices are statistically

---

[8]A reviewer raised the point that a panel data estimator is desirable because our dataset is a panel. Meyerhoefer, Ranney and Sahn (2005) developed a correlated random effect panel estimator for censored equation systems. Their estimator includes household-specific means of those covariates that vary over time as additional explanatory variables in the censored equations to control for *time-invariant* sources of endogeneity. In principle, one could augment our demand equations by adding household-specific average prices as additional regressors. This would in effect make the AGLS estimator a correlated random effect estimator for panel data. However, doing this adds too many variables and eliminates much of the price variation required to identify price effects. Note that when instrumental variables are available, our estimator is more general than the correlated random effect panel estimator because we do not need to assume endogeneity to be time-invariant for the AGLS estimator to be consistent.

To examine the effect of not using a panel data estimator on the results, we aggregated the data to annual frequencies to create a cross-sectional dataset. Results based on this alternative dataset were implausible in that several estimated own-price elasticities were positive. This is partly caused by the much lower correlation between the instruments and prices at annual frequencies compared with quarterly frequencies. When the category-specific price is regressed on its instrument, the coefficient on the instrument has a t-value of 63.0 and 148.2 at the annual and quarterly frequencies, respectively. Although a t-value of 63.0 is far from suggesting a weak instrument, the large number of disaggregated goods in the incomplete demand model suggests that the instruments have to be quite strong for the price effects to be individually identified. Therefore, we concluded that the use of quarterly data is appropriate despite the fact that a panel data estimator is not used.

[9]We followed Lewbel and Pendakur's (2009) suggestion of centering y_{ht} by its sample mean to reduce numerical problems in polynomial regressions. At L=5 the estimation did not encounter any numerical difficulty.

different from parameters estimated without the price exogeneity assumption. Under the null hypothesis that price exogeneity is a proper assumption, the DWH test statistic has a chi-square distribution and a degree of freedom equal to the number of structural parameters of the EASI model.[11] With 736 degrees of freedom, the test statistic is 8,075.8 (p-value <0.00) and thereby decisively rejects price exogeneity.

Price and Income Elasticities

Price and income elasticities are the centerpiece of most demand studies. In an uncensored approximate EASI model, the Hicksian elasticity of demand for good *i* with respect to price of good *j* is

$$h_{ij}=\frac{a_{ij}}{w_i}+w_j-\delta_{ij}, \quad (4)$$

where δ_{ij} = 1 if $i= j$, and 0 otherwise, and the household and time subscripts are withheld for notational brevity. The J×1 vector of income elasticities was calculated as

$$E=(diag(W))^{-1}\left[\left(I_J+BP'\right)^{-1}B\right]+1_J, \quad (5)$$

where W is the J×1 vector of observed budget shares, B is a J×1 vector whose i th element equals $\sum_{r=1}^{L} rb_{ir}y^{r-1}$, P is the J×1 vector of log prices, and 1_J is a J×1 vector of ones. The Marshallian price elasticity, e_{ij}, is recovered from the Slutsky equation $e_{ij} = h_{ij} - w_j e_i$, where e_i is the ith element of E. Equation (5) takes into account the fact that the budget shares w_i appears on both sides of the demand equation through y_{ht} and its polynomials.

When demand is censored at zero, meaningful elasticity estimates cannot be obtained at zero demands. One solution is to calculate expected elasticities by replacing W with conditional means of observed budget shares and substituting marginal effects of log prices and real income polynomials on these conditional means for a_{ij} and b_{ir} in equations (4) and (5). We calculated expected price and expenditure elasticities at all observations. The standard error for each point estimate was generated by taking 100 random draws from a multivariate normal distribution with the mean vector and variance-covariance matrix set to their estimated values.

Table 2 presents median Marshallian price elasticities from the EASI incomplete demand model accounting for endogenous prices. Because of the large sample size, we used the 1% level for all statistical significance tests. Of the 576 median price elasticities, only 38 are

[10]In the incomplete demand model assuming exogenous prices, y_{ht} and its polynomials were treated as exogenous variables in estimation. This is supported by the fact that a regression of y_{ht} on the instrument $\left(\ln x_{ht}-\sum_{j=1}^{J}\overline{w}_j \ln p_{hjt}\right)$ and its polynomials returns a R^2 of 100%. Note that when prices are assumed exogenous, the only source of endogeneity in y_{ht} comes from the budget shares.

[11]The number of structural parameters equals the dimension of the γ vector excluding the λ_i coefficients in the online technical appendix.

statistically insignificant. All own-price elasticities are statistically significant and negative at the median. Six of the 24 categories have median own-price elasticities less than unity in absolute value and therefore have inelastic demand. Figure 3 is a histogram of the uncompensated own-price elasticities pooled across all 24 categories over all observations. It shows that the vast majority of the own-price elasticity point estimates are negative. All three categories of SSBs are own-price elastic, indicating that intake of SSB calories is responsive to SSB price changes, although overall energy intake also depends on the signs and magnitudes of cross-price elasticities between SSBs and other foods.

Many cross-price elasticity estimates are consistent with a priori expectations. For example, substitution is found between regular CSDs and sports/energy drinks, between whole and reduced-fat/skim milk, and between whole-grain and white bread. Peanut butter, which is mainly used as a sandwich spread, is estimated to be a complement to white bread (made from refined grains) but a substitute for whole-grain bread. Peanut butter has the highest rates of energy and fat per ounce of all 23 food and beverage categories. If consumers of whole-grain bread are less likely to use peanut butter as a spread, it would explain the lack of complementarity between the two categories of foods.

Aside from peanut butter, cheese and lunch meat are also common ingredients for making a sandwich. Consistent with expectations, cheese and lunch meat are found to be complements to white bread. However, similar to peanut butter, cheese is found to be a substitute for whole-grain bread. Among the three sandwich ingredients, peanut butter, cheese, and lunch meat are all estimated to be substitutes for each other. The results indicate that candy, cakes and cookies, and snacks are substitutes, which is plausible given these are common nondairy snack categories. In addition, candy, ice cream, and cakes and cookies—the three sweet snack categories—are estimated to be substitutes.

When so many cross-price elasticities are estimated with minimal functional form restrictions, some elasticities will be less intuitive and less straightforward to interpret than others. Among the three SSB categories, although regular CSDs is a substitute to sports/energy drinks and juice drinks, sports/energy drinks and juice drinks are found to be complements. Because sports/energy drinks represent about 4% of total SSB calories in Homescan, the complementary relationship between sports/energy drinks and juice drinks is possibly driven by consumer demand for variety in noncarbonated beverages.

The cross-price elasticities between regular CSDs and diet CSDs are estimated to be positive, small, and statistically insignificant. In comparison, all past demand system studies of aggregate beverage demand using Homescan found regular and diet CSDs to be statistically and economically significant gross complements (Zhen et al. 2011; Lin et al. 2011; Dharmasena and Capps 2012). We suspect that the use of micro data, correcting for the price endogeneity bias and estimating extremely flexible Engel curves are collectively responsible for the difference in our estimated regular-diet CSD cross-price relationship from previous research.

Canned soup, along with canned vegetables and frozen dinners, is a leading source of sodium among the 23 food and beverage categories included in this study. We found that canned soup is a complement to regular CSDs but a substitute for sports/energy drinks and juice drinks. Although we did not set an a priori expectation on the relationship between SSBs and canned soup, it perhaps would have made more sense for the estimated cross-price elasticities between canned soup and individual SSB categories to take the same sign. Table 2 reveals that canned soup is also a complement to diet CSDs but a substitute for bottled water. If one assumes the demand model is correctly specified, carbonation may play a role in this disparity in cross-price relationships because carbonation is one of the few main factors distinguishing CSDs from other SSBs and bottled water. Unlike canned soup, the sign for the cross-price elasticities of demand for canned vegetables and frozen dinners with respect to SSB prices is uniformly positive across individual SSB categories.

The benefit of using income and its polynomials as covariates in the demand model is that the slope of the Engel curve and values of the income elasticity at different income levels are more dictated by data and less by functional form. Table 3 presents the median income elasticity by category and income status. Consistent with the intuition that foods are necessities, income elasticities for the 23 foods and beverages are all significantly below unity. Except for sports/energy drinks and whole-grain bread, the income elasticity for high-income households is always lower than the one for low-income households for the same category of food or beverage. This is also a consequence of foods being a necessity. Sports/energy drinks is the most expensive category of SSBs and is less affordable for low-income households than for their high-income counterparts. Besides being more expensive than white bread, whole-grain bread is recommended for increased consumption by health professionals because of its fiber, vitamin, and mineral content. High-income households' higher income elasticity for whole-grain bread may suggest that high-income households are more attentive to the nutritional value of foods than low-income households. Note that high-income households' median income elasticity for white bread is very close to being zero.

Within each income stratum, for pairs of categories differentiated by healthfulness but otherwise similar, the income elasticity is always higher for the healthier category than for the less healthy category. This is true for CSDs, milk, bread, and juice. This difference in income elasticities for healthy and less healthy options is larger for high-income households than for low-income households. This further supports the notion that higher income consumers are more conscious of the nutrition and health impacts of their food choices. Interestingly, the negative sign on the income elasticity for whole milk for high-income households suggests that it is an inferior good for this income class, although this elasticity is not statistically significant.

Counterfactual Simulation

We simulated a scenario in which a half-cent per ounce SSB excise tax is imposed on regular CSDs, sports/energy drinks, and juice drinks. Assuming the tax is passed one for one onto retail prices, the increase in average retail SSB prices is about 26%. Instead of the much higher penny-per-ounce tax that had been proposed in New York State, we selected the half-cent per-ounce tax for simulation because the price elasticities may be more suited for predicting marginal effects of small to moderate price changes than large price changes. Table 4 summarizes the simulated average effects of the tax on per capita quantity, calories, fat, and sodium by category and income status based on price elasticities from the incomplete demand model that corrected for price endogeneity bias.

On average, a half-cent per ounce SSB price increase is predicted to reduce per capita SSB purchases in Homescan by 113 ounces per quarter (54 ounces of regular CSDs, 16 ounces of sports/energy drinks, and 43 ounces of juice drinks). The SSB reduction is larger for low-income households than for high-income households because low-income households reported higher quantities of SSBs in Homescan. The SSB price increase also induces declines in milk, 100% juice, and bottled water but an increase in diet CSDs. Thus, we found that SSBs as a group are a net complement to non-SSBs except diet CSDs. In contrast, Zhen et al. (2011), Lin et al. (2011), and Dharmasena and Capps (2012) all predict that taxing SSBs would increase demand for 100% juice and whole or low-fat milk, and the first two studies also predict an increase in bottled water demand by high-income households.

The effect of the price increase on foods depends on whether complementarity dominates substitution or vice versa. Food categories predicted to experience an increase in demand after the SSB price increase include white and whole-grain bread, cheese, cereals, canned/dried fruits, canned vegetables, frozen dinners, canned soup, candy, and snacks. Frozen dinners account for the largest increase in calories followed by cheese and white bread. With respect to fat intake, cheese and frozen dinners experience the highest and second highest increase, respectively. Canned soup and canned vegetables are the largest and second largest contributors to increases in sodium intake. Overall, per capita quantities of calories, fat, and sodium from these foods are predicted to increase more for high-income households than for low-income households because an average high-income household purchases more of these foods combined at baseline.

In a closely related study using reduced-form two-part demand models and a subset of the foods and beverages examined here, Finkelstein et al. (2013) found canned soup to be the only food category that is a statistically significant substitute for SSBs. Given that both studies correct for price endogeneity, the functional form difference between the two is likely a major driver of the difference in findings.

Table 5 presents the predicted total effects of the tax on calories, fat, and sodium intake and on consumer welfare. In addition to results from the preferred model, table 5 also contains results from the incomplete model assuming price exogeneity and from two conditional demand models, one of which corrected for price endogeneity and the other did not.[12] For the conditional model, we set $L= 4$ based on results from the same minimum distance test applied earlier to the incomplete demand model. An application of the DWH exogeneity test to the conditional demand model again rejected price exogeneity with a test statistic of 4,059.2 and a p-value<0.00 at 671 degrees of freedom.

[12]In estimating both conditional demand models, we created an instrument for real group expenditure by substituting ln $\hat{x}_{ht}$ for ln x_{ht} in equation (3), where ln x_{ht} is now the log group expenditure on the 23 food and beverage categories and ln $\hat{x}_{ht}$ is the fitted value from a regression of ln x_{ht} on log income and log mean quarterly total expenditures on packaged foods excluding the quarter being instrumented and their polynomials to the fifth order, the demographic variables z_{hkt} from equation (1), and six dummies for residence types. In the conditional demand model under the assumption of price exogeneity, $\tilde{p}_{hjt}$ in equation (3) is replaced by p_{hjt}.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Within the incomplete or conditional demand framework, correction for the price endogeneity bias substantially reduces the magnitude of predicted declines in calories from SSBs and from the 23 foods and beverages; it also reverses the predicted change in sodium intake from a net reduction to a net increase. However, the conditional demand specification fails to predict the increase in food calories that partially compensates for the reduced SSB calories. This results in an estimated overall reduction in dietary energy that is even higher than the decline in SSB calories. In fact, significant calorie compensation is detected only when both price endogeneity correction and an incomplete demand framework are employed. With respect to overall fat intake, the incomplete demand model predicts small net increases in fat, while the conditional model predicts small net decreases, regardless of whether price instruments are used.

Turning now to welfare estimates, the EASI log change in the cost-of-living index (Lewbel and Pendakur 2009, p. 835) was calculated as

$$\ln\left(x^1/x\right)=\left(P^1-P\right)^{'}W+0.5\left(P^1-P\right)^{'}\Gamma\left(P^1-P\right), \quad (6)$$

, where x^1 is post-tax income necessary to maintain utility at the pre-tax level, P^1 is the J×1 vector of new log prices after the tax is imposed, and Γ is a $J\times J$ matrix of parameters whose element Γ_{ij} equals a_{ij} in equation (1). By using observed rather than predicted budget shares in equation (6), the EASI model is able to explicitly incorporate unobserved household heterogeneity into the welfare analysis.

Equation (6) captures two effects of the SSB tax on welfare. The first term of this index is the Stone price effect that ignores any changes in budget shares of the taxed goods. By including observed rather than predicted budget shares, unobserved heterogeneity that is assigned to the error term in estimation is incorporated in the welfare analysis. The second term measures the effect of changing budget shares. The total effect will be smaller than the Stone price effect if budget shares of the taxed goods decrease in response to the tax. In table 5, the compensating variations (CVs) associated with the SSB price increase are decomposed into the Stone price effect and total effect for the incomplete demand model and conditional demand model.

The preferred model (incomplete demand and correction for endogenous prices) predicts the annual CV to be about −$24 per household, which is about 2% lower in absolute value than the Stone price effect. This suggests that a back-of-the-envelope calculation of CV using pre-tax budget shares and price changes provides a good approximation to the true welfare change. The welfare loss for low-income households is about $5 per household per year more than high-income households because low-income households reported higher SSB purchases in Homescan. This difference in welfare loss between low- and high-income households reinforces the regressive nature of an SSB tax.

CV estimates from the conditional demand model are lower than those from the incomplete demand model in absolute value. However, Hanemann and Morey (1992) showed that the CV derived from a conditional demand model is a lower bound for the true CV only when

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the separability assumption is correct. When goods in the conditional demand cannot be assumed separable from the outside goods, the conditional CV estimates are of no value (Hanemann and Morey 1992). Because there is no obvious reason to expect that the 23 food and beverage categories are separable from the excluded foods, goods, and services, the conditional CV estimates are neither upper nor lower bounds of the true CVs.

Finally, the preferred model predicts that a half-cent per ounce SSB tax would generate about $20 per year per household on average in tax revenue. The conditional demand model and the incomplete demand model without accounting for price endogeneity significantly underestimate the potential SSB tax revenue that could be earmarked to fund obesity and healthy eating–related public health campaigns and initiatives.

Conclusion

In this study, we estimated an approximate EASI incomplete demand system containing 23 packaged food and beverage categories and a composite numéraire good. Instrumental variables were used to correct for endogeneity biases caused by omitted variables and measurement errors. The preferred demand specification predicts that almost half of the reduction in SSB calories caused by an increase in SSB prices is compensated for by an increase in calories from other foods. We further found potential unintended consequences of an SSB price increase on sodium and fat intake. Because energy intake is just one of many dimensions of nutrition, the results on sodium and fat highlight the complexity of using targeted food and beverage taxes to improve nutrition outcomes.

An increase in the price of SSBs of one half-cent per ounce, possibly induced by an excise tax on SSBs, is expected to reduce per capita daily calorie intake by 13.2 kcal for the low-income population and 5.6 kcal for the high-income population. Applying these estimates to the dynamic energy-weight loss model used in Lin et al. (2011) predicts weight reductions of 0.37 and 0.16 kg/person in 1 year and 0.70 and 0.31 kg/person in 10 years for low- and high-income adults, respectively.

These findings have broader implications for applications of utility-theoretic consumer demand models to food and nutrition policy research. Our results suggest that the practice of estimating a conditional demand model should be abandoned in favor of an incomplete demand model whenever feasible. In welfare analysis, the main rationale for an incomplete demand specification has been to obtain unbiased welfare estimates. However, at least for SSB price changes, we found that the utility-theoretic estimates of CV are numerically close to simple calculations based on observed budget shares (with income as the denominator) and expected price changes. A much larger and economically significant difference between the conditional and incomplete demand model is in their predictions of the effect of an SSB price increase on nutrition. A conditional model is unable to identify important substitutions among SSBs and other foods and provides an overly optimistic prediction of the potential positive nutritional effect of increasing SSB prices. It is equally important to account for endogenous prices in an incomplete demand model: without a price endogeneity correction, the incomplete demand model alone cannot predict household compensation for reduced SSB consumption.

Notwithstanding several data and methodological enhancements, our results have to be interpreted with caution. Without accounting for nutrients from away-from-home consumption and at-home food categories embedded in the composite numéraire good, the predicted effects on calories, fat, and sodium are specific to the group of 23 food and beverage categories. However, the estimated welfare losses associated with the SSB tax are likely to be robust to future expansions of food categories due to the theoretical results of LaFrance and Hanemann (1989) on incomplete demand systems. Although our preferred demand specification produces more plausible predictions on the overall effects of an SSB tax than several alternative models, a number of estimated cross-price elasticities at the category level have signs that are difficult to explain with intuition. This may not be surprising because cross-price elasticities are the most difficult to identify among all price and income (or expenditure) elasticities estimated with flexible functional forms. For future research, identification of the cross-price effects may benefit from using longer time series, alternative specifications of censoring, and product aggregation schemes guided by formal tests of weakly separable preferences for elementary food items (e.g., Reed, Levedahl and Hallahan 2005). Finally, if food purchases are underreported in Homescan, the predicted reduction in calories would be a lower bound on the true effect, assuming the extent of underreporting is uniform across the 23 food and beverage categories.

Acknowledgments

We are grateful for constructive suggestions and detailed comments from two anonymous reviewers and the journal's editor Brian Roe. We thank ERS for access to the Homescan data; Melanie Ball, Jeremy Bray, and Matthew Farrelly for the C2ER data; and David Chrest and Kibri Everett for GIS analyses. The views expressed here are those of the authors and cannot be attributed to ERS or USDA. Grants support from the Robert Wood Johnson Foundation's Healthy Eating Research (65062) and the National Heart, Lung, and Blood Institute, National Institute of Health (1R21HL108121) is gratefully acknowledged. All errors are our own.

References

Amemiya T. The Estimation of a Simultaneous-Equation Tobit Model. International Economic Review. 1979; 20:169–181.

Banks J, Blundell R, Lewbel A. Quadratic Engel Curves and Consumer Demand. Review of Economics and Statistics. 1997; 79:527–539.

Bridging the Gap. Bridging the Gap Program. Chicago, IL: University of Illinois at Chicago; 2011. State Sales Tax on Regular, Sugar-Sweetened Soda. http://www.impacteen.org/statetaxdata/BTG_State_Soda_Sales_Tax_Jan012011_publuse_091911.pdf

Broda C, Leibtag E, Weinstein DE. The Role of Prices in Measuring the Poor's Living Standards. Journal of Economic Perspectives. 2009; 23:77–97. [PubMed: 20052301]

Brownell KD, Farley T, Willett WC, Popkin BM, Chaloupka FJ, Thompson JW, Ludwig DS. The Public Health and Economic Benefits of Taxing Sugar-Sweetened Beverages. New England Journal of Medicine. 2009; 361:1599–1605. [PubMed: 19759377]

Campaign for Tobacco-Free Kids. State Cigarette Excise Tax Rates and Rankings. 2011. http://www.tobaccofreekids.org/research/factsheets/pdf/0097.pdf

CDC (Centers for Disease Control and Prevention). Usual Sodium Intakes Compared with Current Dietary Guidelines—United States, 2005–2008. Morbidity and Mortality Weekly Report. 2011; 60:1413–1417. [PubMed: 22012113]

Cox TL, Wohlgenant MK. Prices and Quality Effects in Cross-Sectional Demand Analysis. American Journal of Agricultural Economics. 1986; 68:908–919.

Danaei G, Ding EL, Mozaffarian D, Taylor B, Rehm J, Murray CJL, Ezzati M. The Preventable Causes of Death in the United States: Comparative Risk Assessment of Dietary, Lifestyle, and Metabolic Risk Factors. PLoS Medicine. 2009; 6(4)

Deaton AS. Quality, Quantity, and Spatial Variation of Price. American Economic Review. 1988; 78:418–430.

Deaton AS, Muellbauer J. An Almost Ideal Demand System. American Economic Review. 1980; 70:312–326.

Dhar T, Chavas JP, Gould BW. An Empirical Assessment of Endogeneity Issues in Demand Analysis for Differentiated Products. American Journal of Agricultural Economics. 2003; 85:605–617.

Dharmasena S, Capps O Jr. Intended and Unintended Consequences of a Proposed National Tax on Sugar-Sweetened Beverages to Combat the U.S. Obesity Problem. Health Economics. 2012; 21:669–694. [PubMed: 21538676]

Diewert WE. Exact and Superlative Index Numbers. Journal of Econometrics. 1976; 4:115–145.

Durbin J. Errors in Variables. Review of the International Statistical Institute. 1954; 22:23–32.

Einav L, Leibtag E, Nevo A. Recording Discrepancies in Nielsen Homescan Data: Are They Present and Do They Matter? Quantitative Marketing and Economics. 2010; 8:207–239.

Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual Medical Spending Attributable to Obesity: Payer- and Service-Specific Estimates. Health Affairs. 2009; 28(5):w822–w831. [PubMed: 19635784]

Finkelstein EA, Zhen C, Bilger M, Nonnemaker J, Farooqui AM, Todd JE. Implications of a Sugar-Sweetened Beverage (SSB) Tax When Substitutions to Non-Beverage Items Are Considered. Journal of Health Economics. 2013; 32:219–239. [PubMed: 23202266]

Finkelstein EA, Zhen C, Nonnemaker J, Todd JE. Impact of Targeted Beverage Taxes on Higher- and Lower-Income Households. Archives of Internal Medicine. 2010; 170:2028–2034. [PubMed: 21149762]

Fletcher JM, Frisvold D, Tefft N. The Effects of Soft Drink Taxes on Child and Adolescent Consumption and Weight Outcomes. Journal of Public Economics. 2010; 94:967–974.

Guthrie JF, Lin BH, Frazao E. Role of Food Prepared Away from Home in the American Diet, 1977–78 versus 1994–96: Changes and Consequences. Journal of Nutrition Education and Behavior. 2002; 34:140–150. [PubMed: 12047838]

Hanemann M, Morey E. Separability, Partial Demand Systems, and Consumer's Surplus Measures. Journal of Environmental Economics and Management. 1992; 22:241–258.

Hausman J. Specification Tests in Econometrics. Econometrica. 1978; 46:1251–1271.

Hausman, J. Valuation of New Goods under Perfect and Imperfect Competition. In: Bresnahan, TF.; Gordon, RJ., editors. The Economics of New Goods, Studies in Income and Wealth. Vol. 58. Chicago: National Bureau of Economic Research; 1997.

IOM (Institute of Medicine). A Population-based Policy and Systems Change Approach to Prevent and Control Hypertension. Washington, DC: The National Academies Press; 2010a.

IOM (Institute of Medicine). Strategies to Reduce Sodium Intake in the United States. Washington, DC: The National Academies Press; 2010b.

Ippolito PM, Mathios AD. Information and Advertising: The Case of Fat Consumption in the United States. American Economic Review Papers and Proceedings. 1995; 85:91–95.

LaFrance JT. When Is Expenditure 'Exogenous' in Separable Demand Models? Western Journal of Agricultural Economics. 1991; 16:49–62.

LaFrance JT, Hanemann WM. The Dual Structure of Incomplete Demand Systems. American Journal of Agricultural Economics. 1989; 71:262–274.

Lee LF, Pitt MM. Microeconometric Demand System with Binding Nonnegativity Constraints: The Dual Approach. Econometrica. 1986; 54:1237–1242.

Lewbel A, Pendakur K. Tricks with Hicks: The EASI Demand System. American Economic Review. 2009; 99:827–863.

Lin BH, Smith TA, Lee JY, Hall KD. Measuring Weight Outcomes for Obesity Intervention Strategies: The Case of a Sugar-Sweetened Beverage Tax. Economics and Human Biology. 2011; 9:329–341. [PubMed: 21940223]

Meyerhoefer CD, Ranney CK, Sahn DE. Consistent Estimation of Censored Demand Systems Using Panel Data. American Journal of Agricultural Economics. 2005; 87:660–672.

Newey WK. Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables. Journal of Econometrics. 1987; 36:231–250.

Ogden, CL.; Carroll, MD.; Kit, BK.; Flegal, KM. NCHS data brief, no 82. Hyattsville, MD: National Center for Health Statistics; 2012. Prevalence of Obesity in the United States, 2009–2010.

Okrent AM, Alston JM. The Effects of Farm Commodity and Retail Food Policies on Obesity and Economic Welfare in the United States. American Journal of Agricultural Economics. 2012; 94:611–646.

Ostchega, Y.; Yoon, SS.; Hughes, J.; Louis, T. NCHS Data Brief No 3. Hyattsville, MD: National Center for Health Statistics; 2008. Hypertension Awareness, Treatment, and Control—Continued Disparities in Adults: United States, 2005–2006.

Reed AJ, Levedahl JW, Hallahan C. The Generalized Composite Commodity Theorem and Food Demand Estimation. American Journal of Agricultural Economics. 2005; 87:28–37.

USDA (U.S. Department of Agriculture), Agricultural Research Service, USDA Nutrient Data Laboratory. USDA National Nutrient Database for Standard Reference. 2009. Release 22

Van Horn L, McCoin M, Kris-Etherton PM, Burke F, Carson JAS, Champagne CM, Karmally W, Sikand G. The Evidence for Dietary Prevention and Treatment of Cardiovascular Disease. Journal of the American Dietetic Association. 2008; 108:287–331. [PubMed: 18237578]

Wales TJ, Woodland AD. Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints. Journal of Econometrics. 1983; 21:263–285.

Wohlgenant, MK. Effects of the Changing Composition of Beef Consumption on the Elasticities for Beef and Poultry. In: Buse, Rueben C., editor. The Economics of Meat Demand. Madison WI: University of Wisconsin; 1989. p. 170-186.

Wu D. Alternative Tests of Independence between Stochastic Regressors and Disturbances. Econometrica. 1973; 41:733–750.

Zhen C, Taylor JL, Muth MK, Leibtag E. Understanding Differences in Self-Reported Expenditures between Household Scanner Data and Diary Survey Data: A Comparison of Homescan and Consumer Expenditure Survey. Review of Agricultural Economics. 2009; 31:470–492.

Zhen C, Wohlgenant MK, Karns S, Kaufman P. Habit Formation and Demand for Sugar-Sweetened Beverages. American Journal of Agricultural Economics. 2011; 93:175–193.

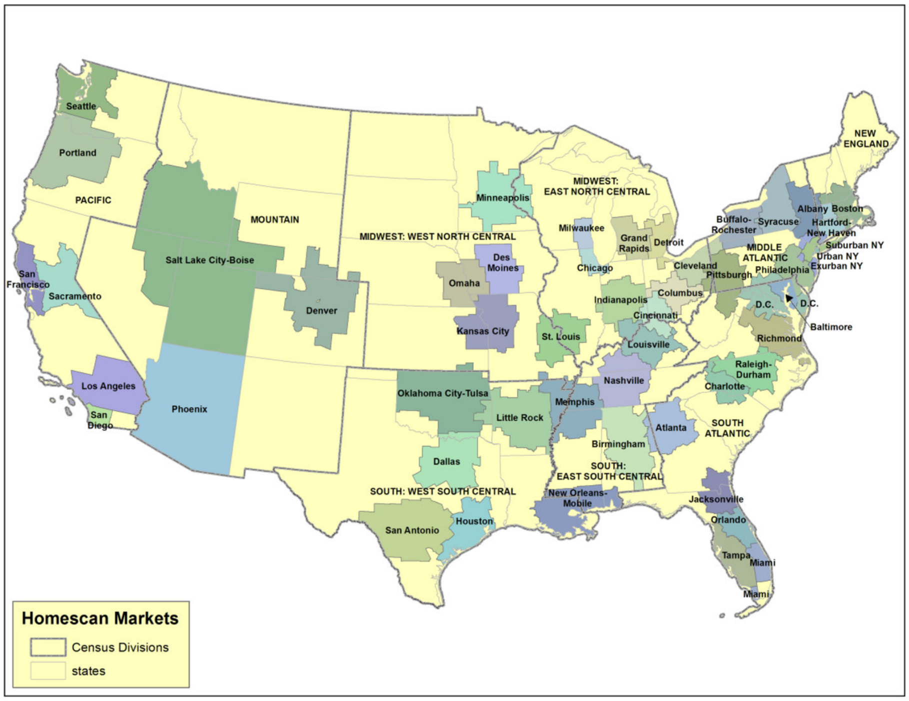

Figure 1. Homescan market map

Source: Created based on Nielsen's county list.

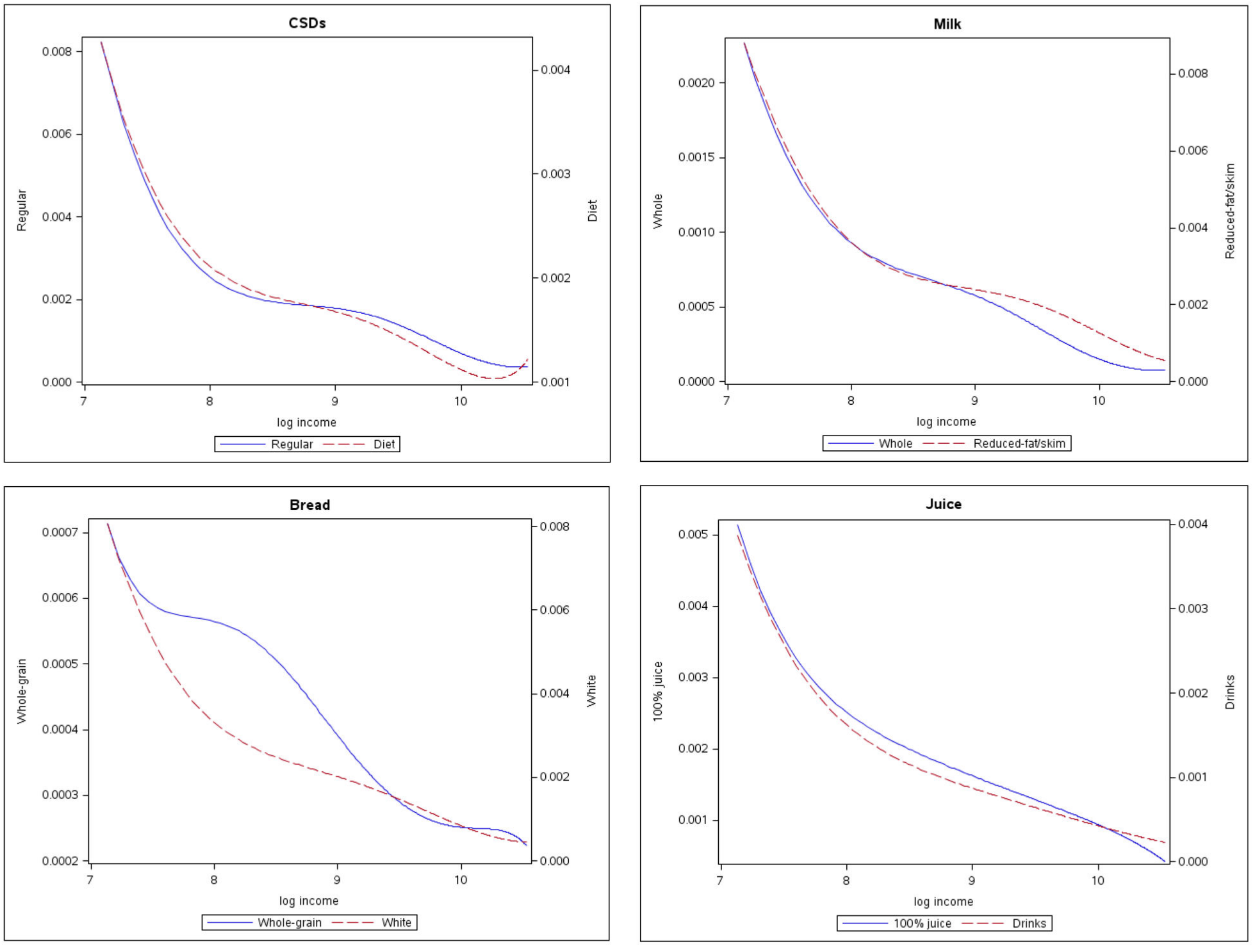

Figure 2. Estimated Engel curves for selected foods and beverages

Note: Incomplete demand, correction for price endogeneity, *L*=5, two-person non-Hispanic white household in the Northeast with college-educated female head over age 35 and no children.

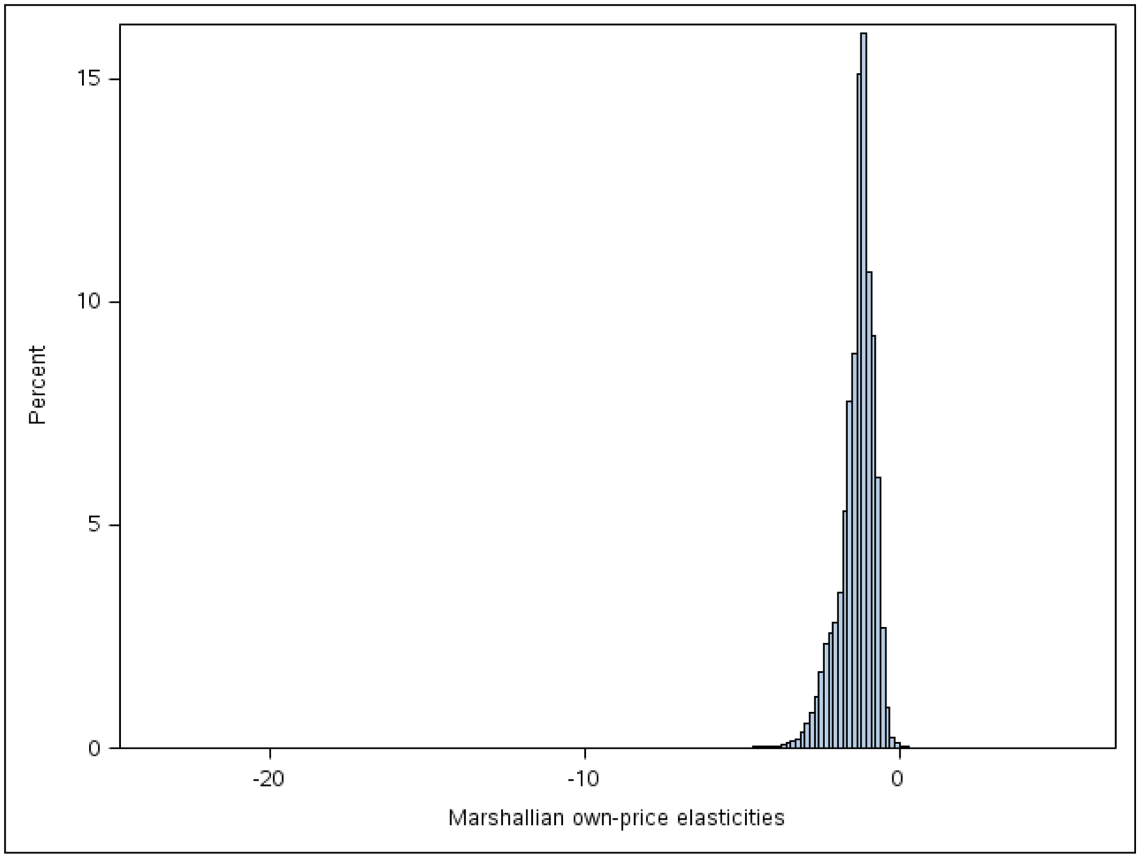

Figure 3. Distribution of own-price elasticity point estimates, all categories
Note: Incomplete demand, correction for price endogeneity, *L*=5.

Table 1

Per Capita Expenditures, Purchase Quantities, Calories, Nutrients, and Unit Values by Food Category and Income Status

	Low income						High income					
Food/beverage category	Expenditure ($/quarter)	Volume (oz/quarter)	Energy (kcal/day)	Fat (g/day)	Sodium (mg/day)	Unit value (cents/oz)	Expenditure ($/quarter)	Volume (oz/quarter)	Energy (kcal/day)	Fat (g/day)	Sodium (mg/day)	Unit value (cents/oz)
1. Regular CSD	6.33	333	46	0.0	14	1.9	5.53	279	38	0.0	12	2.0
2. Sports/energy drinks	1.39	40	3	0.0	5	3.5	2.08	57	4	0.0	7	3.6
3. Whole milk	2.39	100	21	1.0	17	2.4	1.70	68	14	0.7	12	2.5
4. Reduced-fat/skim milk	5.76	259	42	1.2	46	2.2	6.64	291	45	1.1	52	2.3
5. Whole-grain bread	0.89	10	8	0.1	15	9.0	1.26	13	10	0.1	20	9.8
6. White bread	5.23	68	55	0.7	110	7.7	5.72	62	51	0.6	100	9.2
7. Cheese	7.19	40	37	2.9	97	17.8	8.83	45	42	3.2	105	19.5
8. 100% juices	4.09	117	19	0.0	2	3.5	5.01	139	22	0.0	2	3.6
9. Juice drinks	3.12	218	31	0.0	15	1.4	3.26	180	25	0.0	11	1.8
10. Peanut butter	0.79	8	16	1.3	11	9.3	0.83	9	16	1.3	12	9.6
11. Cereals	6.38	46	53	0.6	67	13.9	7.18	48	55	0.6	69	15.0
12. Yogurt	1.85	22	6	0.1	4	8.3	2.64	31	8	0.1	5	8.6
13. Diet CSD	3.55	185	0	0.0	8	1.9	4.85	247	0	0.0	10	2.0
14. Bottled water	2.13	203	0	0.0	1	1.0	2.90	268	0	0.0	1	1.1
15. Canned/dried fruits	2.32	28	19	0.1	5	8.2	2.58	28	19	0.1	5	9.3
16. Canned vegetables	4.52	82	13	0.1	134	5.5	4.91	79	13	0.1	123	6.2
17. Frozen dinners	11.67	84	49	2.2	115	13.9	12.68	79	45	2.0	105	16.0
18. Canned soup	2.49	34	7	0.2	67	7.4	3.18	41	8	0.3	79	7.8
19. Candy	6.19	33	46	1.6	12	18.7	7.00	37	52	1.9	13	19.1
20. Ice cream	3.24	65	26	1.3	10	5.0	3.35	60	24	1.2	9	5.5
21. Cakes and cookies	7.45	52	68	2.9	56	14.3	8.06	50	64	2.7	55	16.1
22. Lunch meat	3.69	19	12	0.8	64	19.4	3.89	18	11	0.7	58	22.2
23. Snacks	7.02	41	66	3.4	102	16.9	8.52	46	72	3.5	111	18.4
Total	99.68	2,089	642	20.6	976		112.61	2,174	638	20.3	976	

Note: Summary statistics weighted by Nielsen's projection factor. Beverages and canned soup are measured in fluid ounces. Average household size is 2.6 for both low- and high-income households. Total per capita expenditures on packaged foods were $187 and $217 per quarter for low- and high-income households, respectively. Cereals include ready-to-eat and ready-to-heat cereals. Frozen dinners combine frozen entrée, frozen pizza, and frozen French fries. Snacks include salty snacks, popcorn, snack mixes, nutrition/energy/protein bars, and other snacks.

Table 2

Median Marshallian Price Elasticities, All Income

Elasticity of the quantity of	With respect to the price of											
	1. Regular CSD	**2. Sports/energy drinks**	**3. Whole milk**	**4. Reduced-fat/skim milk**	**5. Whole-grain bread**	**6. White bread**	**7. Cheese**	**8. 100% juice**	**9. Juice drinks**	**10. Peanut butter**	**11. Cereals**	**12. Yogurt**
1. Regular CSD	**−1.035** (−103.2)	0.150 (19.8)	0.025 (3.4)	−0.131 (−20.0)	0.061 (19.6)	0.184 (54.8)	0.167 (32.1)	−0.177 (−32.8)	0.171 (31.7)	−0.039 (−16.8)	−0.078 (−12.7)	0.049 (9.2)
2. Sports/energy drinks	0.262 (19.9)	**−2.363** (−106.9)	−0.793 (−49.2)	0.349 (27.1)	−0.311 (−53.1)	0.175 (24.2)	*0.025* (2.2)	−0.298 (−31.0)	−0.556 (−69.1)	−0.132 (−23.9)	0.209 (22.8)	−0.176 (−19.2)
3. Whole milk	0.048 (3.5)	−0.844 (−46.7)	**−0.900** (−23.9)	0.734 (40.3)	*−0.004* (0.7)	−0.081 (−10.2)	0.060 (5.8)	−0.163 (−14.3)	0.132 (12.0)	0.080 (13.3)	−0.133 (−8.4)	−0.164 (−14.6)
4. Reduced-fat/skim milk	−0.178 (−20.1)	0.270 (27.0)	0.532 (40.7)	**−1.199** (−129.5)	0.053 (9.4)	−0.125 (−25.0)	−0.265 (−49.8)	0.121 (21.3)	−0.135 (−23.1)	0.011 (4.3)	0.125 (16.1)	0.040 (6.6)
5. Whole-grain bread	0.239 (19.8)	−0.695 (−52.8)	*−0.008* (0.6)	0.152 (9.5)	**−1.196** (−95.0)	0.372 (36.5)	0.487 (38.7)	0.485 (30.2)	0.352 (33.8)	0.110 (15.2)	−0.642 (−39.0)	−0.182 (−13.9)
6. White bread	0.330 (54.8)	0.180 (24.3)	−0.079 (−10.3)	−0.166 (−25.0)	0.171 (37.4)	**−0.666** (−103.2)	−0.370 (−47.9)	0.322 (39.1)	−0.158 (−29.7)	−0.070 (−20.3)	−0.287 (−38.4)	0.199 (25.3)
7. Cheese	0.209 (32.5)	*0.017* (2.1)	0.039 (5.7)	−0.245 (−49.8)	0.156 (40.1)	−0.258 (−48.8)	**−0.567** (−64.9)	0.264 (30.4)	0.125 (24.9)	0.210 (50.5)	0.260 (36.6)	−0.151 (−22.1)
8. 100% juice	−0.289 (−33.2)	−0.278 (−30.9)	−0.144 (−14.3)	0.146 (21.2)	0.202 (30.2)	0.292 (39.1)	0.344 (30.4)	**−1.566** (−78.2)	−0.187 (−22.5)	−0.081 (−11.4)	*0.002* (−0.9)	−0.171 (−15.1)
9. Juice drinks	0.353 (32.1)	−0.656 (−68.2)	0.146 (12.0)	−0.205 (−22.9)	0.185 (34.2)	−0.181 (−30.0)	0.206 (25.0)	−0.236 (−22.5)	**−1.192** (−99.0)	−0.217 (−41.6)	0.138 (11.0)	−0.127 (−12.9)
10. Peanut butter	−0.212 (−16.8)	−0.406 (−23.6)	0.231 (13.4)	0.045 (4.4)	0.151 (15.3)	−0.209 (−20.3)	0.906 (49.2)	−0.266 (−11.3)	−0.566 (−40.7)	**−1.466** (−90.5)	−0.626 (−33.7)	1.351 (67.6)
11. Cereals	−0.107 (−12.8)	0.163 (22.9)	−0.098 (−8.5)	0.126 (16.2)	−0.224 (−39.2)	−0.218 (−38.5)	0.283 (36.5)	*0.002* (−0.9)	0.091 (10.9)	−0.159 (−33.8)	**−0.814** (−65.0)	0.042 (5.4)
12. Yogurt	0.120 (9.2)	−0.244 (−19.2)	−0.215 (−14.7)	0.072 (6.6)	−0.113 (−13.9)	0.270 (25.3)	−0.294 (−22.0)	−0.255 (−15.0)	−0.150 (−13.0)	0.610 (70.9)	0.074 (5.4)	**−2.043** (−119.4)
13. Diet CSD	*0.004* (0.0)	0.236 (27.4)	−0.236 (−24.7)	0.302 (42.4)	−0.088 (−24.1)	0.025 (7.0)	−0.024 (−3.4)	0.174 (31.3)	0.023 (2.9)	*0.002* (0.7)	0.150 (25.7)	0.080 (14.1)
14. Bottled water	−0.198 (−12.3)	−0.618 (−55.6)	0.167 (9.5)	*−0.012* (−1.9)	0.054 (13.5)	−0.180 (−33.8)	−0.381 (−43.9)	0.224 (29.3)	−0.124 (−15.9)	0.154 (37.7)	0.315 (37.1)	−0.369 (−50.2)
15. Canned/dried fruits	−0.100 (−10.0)	0.294 (32.0)	0.300 (31.9)	−0.425 (−66.2)	0.133 (24.0)	0.201 (29.7)	−0.337 (−34.0)	−0.201 (−17.3)	0.155 (22.5)	0.042 (7.9)	0.092 (8.3)	−0.175 (−19.6)
16. Canned vegetables	0.260 (27.8)	0.116 (10.0)	0.462 (31.3)	−0.508 (−51.8)	0.156 (27.6)	0.202 (32.9)	0.026 (5.3)	0.494 (42.3)	0.139 (17.8)	0.396 (67.1)	0.262 (25.8)	−0.343 (−31.9)
17. Frozen dinners	0.190 (28.4)	0.462 (78.7)	−0.236 (−24.6)	0.167 (30.5)	−0.070 (−22.3)	*0.006* (1.3)	−0.344 (−58.9)	−0.051 (−6.8)	*0.017* (0.3)	−0.031 (−9.9)	0.022 (5.0)	0.032 (7.7)
18. Canned soup	−0.132 (−13.0)	1.218 (121.7)	−0.410 (−35.7)	0.343 (41.7)	0.014 (3.3)	0.179 (31.7)	*−0.023* (−1.8)	−0.620 (−57.0)	0.779 (89.1)	*0.011* (1.7)	0.207 (21.5)	0.067 (7.6)
19. Candy	−0.230 (−23.7)	0.501 (58.8)	0.176 (16.4)	−0.091 (−13.6)	0.080 (19.4)	0.023 (4.4)	−0.479 −59.1)	0.177 (18.5)	0.304 (47.6)	*0.010* (1.2)	0.141 (18.1)	−0.037 (−5.3)
20. Ice cream	−0.121 (−17.6)	−0.462 (−57.9)	0.184 (21.3)	−0.130 (−21.1)	−0.330 (−74.5)	*−0.010* (−0.6)	0.280 (39.0)	0.080 (10.3)	−0.179 (−27.8)	−0.125 (−33.3)	0.044 (5.7)	−0.037 (−4.7)
21. Cakes and cookies	−0.051 (−6.8)	0.043 (6.0)	−0.189 (−21.1)	*−0.005* (−0.3)	−0.108 (−21.4)	0.237 (42.1)	0.117 (15.0)	0.060 (6.7)	−0.095 (−14.3)	−0.171 (−39.7)	−0.132 (−17.8)	0.112 (13.8)
22. Lunch meat	*0.005* (2.4)	−0.346 (−37.3)	0.243 (27.3)	−0.042 (−6.5)	−0.158 (−33.3)	−0.107 (−17.2)	1.008 (99.4)	−0.155 (−17.2)	0.041 (5.2)	0.053 (11.6)	−0.042 (−4.7)	0.972 (92.8)
23. Snacks	0.145 (17.9)	*0.000* (−0.1)	0.080 (8.9)	−0.089 (−13.1)	0.273 (54.0)	−0.535 (−97.7)	0.123 (15.5)	0.316 (28.2)	−0.073 (−8.4)	0.088 (18.3)	−0.289 (−32.3)	0.131 (17.1)
24. All other goods	−0.001 (−18.4)	0.001 (18.9)	0.001 (6.8)	−0.001 (−13.5)	0.000 (−7.4)	0.000 (−16.7)	−0.002 (−62.5)	0.000 (−11.2)	0.000 (−9.6)	−0.001 (−33.8)	−0.001 −55.0)	0.000 (8.0)

Elasticity of the quantity of	With respect to the price of											
	13. Diet CSD	**14. Bottled water**	**15. Canned/dried fruits**	**16. Canned vegetables**	**17. Frozen dinners**	**18. Canned soup**	**19. Candy**	**20. Ice cream**	**21. Cakes & cookies**	**22. Lunch meat**	**23. Snacks**	**24. All other goods**
1. Regular CSD	*0.004* (0.0)	−0.100 (−12.4)	−0.038 (−10.1)	0.132 (27.5)	0.275 (28.4)	−0.055 (−13.0)	−0.183 (−23.8)	−0.065 (−17.7)	−0.043 (−6.9)	*0.003* (2.3)	0.121 (17.8)	0.294 (12.5)
2. Sports/energy drinks	0.401 (27.7)	−0.547 (−55.2)	0.194 (31.6)	0.103 (10.1)	1.169 (77.2)	0.887 (124.9)	0.699 (59.4)	−0.431 (−57.8)	0.064 (6.1)	−0.342 (−38.3)	*0.000* (−0.1)	0.731 (33.7)
3. Whole milk	−0.426 (−24.2)	0.158 (9.5)	0.211 (31.6)	0.439 (30.7)	−0.634 (−24.2)	−0.317 (−34.6)	0.263 (16.4)	0.183 (21.1)	−0.298 (−20.9)	0.257 (27.6)	0.126 (8.9)	1.116 (22.0)
4. Reduced-fat/skim milk	0.396 (42.8)	*−0.008* (−1.9)	−0.217 (−66.7)	−0.350 (−53.3)	0.327 (30.7)	0.193 (41.6)	−0.098 (−13.7)	−0.094 (−21.0)	*−0.007* (−0.3)	−0.033 (−6.5)	−0.100 (−13.1)	0.469 (25.6)
5. Whole-grain bread	−0.332 (−24.3)	0.107 (13.5)	0.196 (23.9)	0.311 (28.1)	−0.395 (−22.3)	0.023 (3.4)	0.251 (19.6)	−0.687 (−74.1)	−0.359 (−21.1)	−0.349 (−33.9)	0.890 (49.9)	0.289 (18.7)
6. White bread	0.044 (7.0)	−0.164 (−34.5)	0.136 (29.5)	0.185 (33.7)	*0.015* (1.3)	0.134 (31.8)	0.033 (4.4)	*−0.010* (−0.7)	0.363 (42.4)	−0.109 (−17.4)	−0.800 (−101.2)	0.609 (44.5)
7. Cheese	−0.029 (−3.4)	−0.241 (−44.6)	−0.159 (−34.6)	0.017 (5.2)	−0.622 (−59.9)	−0.012 (−1.8)	−0.477 (−60.8)	0.186 (38.4)	0.124 (14.9)	0.713 (107.2)	0.128 (15.5)	*0.040* (2.4)
8. 100% juice	0.275 (31.1)	0.185 (29.3)	−0.123 (−17.3)	0.410 (43.7)	−0.119 (−6.7)	−0.421 (−57.6)	0.230 (18.7)	0.069 (10.3)	0.083 (6.7)	−0.144 (−17.2)	0.429 (28.2)	0.422 (37.4)
9. Juice drinks	0.047 (2.9)	−0.129 (−15.9)	0.120 (22.3)	0.146 (17.9)	*0.051* (0.4)	0.669 (89.5)	0.500 (50.4)	−0.197 (−28.1)	−0.167 (−14.2)	0.047 (5.2)	−0.126 (−8.5)	0.457 (25.3)
10. Peanut butter	*0.013* (0.8)	0.418 (37.2)	0.084 (7.9)	1.086 (68.1)	−0.241 (−9.9)	*0.024* (1.6)	*0.044* (1.3)	−0.358 (−32.4)	−0.783 (−38.3)	0.161 (11.6)	0.395 (18.3)	−0.098 (−5.1)
11. Cereals	0.198 (25.7)	0.217 (37.3)	0.047 (8.3)	0.182 (25.8)	0.043 (5.0)	0.118 (21.5)	0.153 (18.1)	0.032 (5.6)	−0.154 (−17.8)	−0.033 (−4.7)	−0.328 (−32.6)	0.139 (12.2)
12. Yogurt	0.190 (14.2)	−0.454 (−51.1)	−0.160 (−19.5)	−0.425 (−32.0)	0.113 (7.7)	0.067 (7.6)	−0.071 (−5.3)	−0.048 (−4.7)	0.233 (13.8)	1.336 (102.7)	0.266 (17.1)	0.649 (38.0)
13. Diet CSD	**−0.959** (−74.3)	−0.119 (−16.2)	−0.021 (−4.6)	−0.114 (−21.4)	−0.245 (−22.2)	−0.034 (−7.7)	0.084 (9.9)	−0.119 (−26.6)	−0.060 (−7.4)	−0.063 (−13.9)	*0.011* (1.5)	0.364 (23.7)
14. Bottled water	−0.228 (−16.4)	**−1.703** (−97.7)	0.254 (50.7)	*0.013* (1.3)	−0.226 (−15.0)	0.413 (63.3)	0.687 (48.8)	0.087 (16.3)	0.395 (35.6)	0.143 (21.5)	−0.142 (−12.7)	0.752 (28.5)
15. Canned/dried fruits	−0.053 (−4.6)	0.341 (50.4)	**−1.222** (−116.2)	−0.285 (−31.0)	*−0.004* (−1.2)	−0.043 (−6.1)	−0.462 (−38.8)	0.044 (5.3)	0.381 (33.0)	−0.024 (−2.9)	0.562 (41.8)	0.537 (32.4)
16. Canned vegetables	−0.217 (−21.4)	*0.013* (1.4)	−0.211 (−30.6)	**−1.516** (−138.9)	−0.152 (−12.8)	−0.035 (−6.7)	−0.522 (−55.7)	0.143 (20.3)	−0.096 (−7.8)	0.402 (58.1)	0.291 (26.6)	0.053 (3.4)
17. Frozen dinners	−0.165 (−22.1)	−0.079 (−15.0)	*−0.001* (−1.3)	−0.054 (−12.9)	**−0.765** (−61.4)	−0.157 (−39.2)	0.071 (12.3)	0.278 (66.1)	0.190 (25.0)	−0.208 (−38.9)	−0.031 (−5.3)	0.465 (43.9)
18. Canned soup	−0.078 (−7.7)	0.502 (63.4)	−0.039 (−6.2)	−0.043 (−6.8)	−0.547 (−39.1)	**−2.472** (−253.8)	−1.599 (−142.1)	0.761 (106.0)	0.118 (10.8)	0.385 (48.2)	0.624 (62.8)	0.440 (25.7)
19. Candy	0.103 (9.9)	0.436 (47.4)	−0.218 (−38.3)	−0.334 (−55.1)	0.129 (12.4)	−0.834 (−120.9)	**−1.485** (−108.4)	0.108 (20.4)	0.102 (9.2)	0.071 (11.7)	0.259 (29.5)	0.808 (26.7)
20. Ice cream	−0.217 (−26.9)	0.083 (16.2)	0.031 (5.3)	0.137 (20.6)	0.756 (69.6)	0.594 (107.3)	0.162 (20.6)	**−1.115** (−160.3)	0.294 (32.7)	−0.259 (−43.0)	−0.416 (−45.4)	0.545 (55.1)
21. Cakes and cookies	−0.068 (−7.4)	0.235 (35.8)	0.168 (32.9)	−0.057 (−7.8)	0.323 (25.0)	0.058 (10.8)	0.095 (9.3)	0.184 (32.5)	**−1.697** (−122.2)	−0.187 (−26.4)	0.313 (30.6)	0.584 (41.9)
22. Lunch meat	−0.108 (−14.0)	0.128 (21.4)	−0.016 (−2.9)	0.362 (61.4)	−0.532 (−37.1)	0.284 (48.1)	0.100 (11.7)	−0.244 (−42.0)	−0.282 (−26.0)	**−1.216** (−139.5)	0.109 (10.6)	−0.332 (−23.2)
23. Snacks	*0.013* (1.5)	−0.086 (−12.6)	0.254 (42.1)	0.178 (26.7)	−0.053 (−5.3)	0.312 (61.8)	0.247 (29.8)	−0.266 (−45.2)	0.320 (30.4)	0.074 (10.6)	**−1.266** (−80.7)	−0.276 (−21.1)
24. All other goods	*0.000* (1.1)	0.001 (11.2)	0.000 (−11.5)	−0.001 (−38.5)	−0.001 −26.0)	0.000 −15.1)	0.000 (3.7)	0.000 (−23.2)	−0.001 −10.8)	−0.002 (−78.1)	−0.003 (−71.6)	**−1.004** −2,031.2)

Note: All elasticities and their t-values (in parentheses) are median values over households of all income levels weighted by Nielsen’s projection factor. All elasticities are statistically significant at the 1% level except for the italicized ones. Results are based on the incomplete demand model that accounts for price endogeneity with L = 5.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Median Income Elasticities by Income Status

	Low income	High income
1. Regular CSD	0.726 (93.0)	0.114 (16.5)
2. Sports/energy drinks	0.679 (79.0)	0.683 (56.4)
3. Whole milk	0.553 (45.0)	−0.146 (−0.8)
4. Reduced-fat/skim milk	0.596 (110.6)	0.074 (16.8)
5. Whole-grain bread	0.574 (73.2)	0.603 (53.5)
6. White bread	0.482 (110.9)	0.009 (8.8)
7. Cheese	0.459 (139.0)	0.170 (34.8)
8. 100% juice	0.554 (138.5)	0.369 (63.9)
9. Juice drinks	0.402 (84.1)	0.363 (46.6)
10. Peanut butter	0.784 (89.6)	0.080 (10.7)
11. Cereals	0.533 (142.6)	0.184 (33.7)
12. Yogurt	0.682 (110.4)	0.298 (51.2)
13. Diet CSD	0.738 (135.2)	0.582 (59.8)
14. Bottled water	0.820 (126.6)	0.354 (32.6)
15. Canned/dried fruits	0.379 (84.0)	0.194 (17.0)
16. Canned vegetables	0.337 (66.8)	0.173 (17.0)
17. Frozen dinners	0.540 (128.7)	0.177 (27.0)
18. Canned soup	0.637 (110.7)	0.131 (23.6)
19. Candy	0.482 (76.9)	0.181 (15.9)
20. Ice cream	0.486 (137.9)	0.059 (14.5)
21. Cakes & cookies	0.341 (88.6)	0.201 (16.7)
22. Lunch meat	0.639 (100.7)	0.103 (13.9)
23. Snacks	0.383 (111.8)	0.346 (47.1)
24. All other goods	1.020 (5,157.9)	1.018 (5,148.2)

Note: All elasticities are median values within the corresponding income status weighted by Nielsen's projection factor. Results are based on the incomplete demand model that accounts for price endogeneity with L = 5. t-values in parentheses.

Table 4

Average Effects of a Half-Cent per Ounce Increase in SSB Prices on Per Capita Purchases

	Per capita change in														
	purchase quantity (oz/quarter)[a]						calories (kcal/day)[b]			fat (g/day)[b]			sodium (mg/day)[b]		
	All income		Low income		High income		All income	Low income	High income	All income	Low income	High income	All income	Low income	High income
	Est.	t-Stat	Est.	t-Stat	Est.	t-Stat									
1. Regular CSD	−54.0	−68.5	−65.8	−95.4	−49.3	−57.1	−7.5	−9.2	−6.8	0.00	0.00	0.00	−2.3	−2.7	−2.1
2. Sports/energy drinks	−16.0	−79.3	−9.9	−84.7	−18.3	−76.2	−1.0	−0.6	−1.2	0.00	0.00	0.00	−1.9	−1.1	−2.2
3. Whole milk	−3.3	−6.9	−2.8	−6.2	−3.6	−7.2	−0.7	−0.6	−0.8	−0.03	−0.03	−0.04	−0.6	−0.5	−0.6
4. Reduced-fat/skim milk	−11.9	−14.5	−7.6	−14.8	−14.1	−14.4	−1.9	−1.2	−2.2	−0.05	−0.04	−0.05	−2.1	−1.4	−2.5
5. Whole-grain bread	0.7	14.0	0.5	15.0	0.8	13.4	0.5	0.3	0.6	0.01	0.00	0.01	1.0	0.7	1.2
6. White bread	3.8	33.9	2.6	33.3	4.6	34.1	3.1	2.1	3.8	0.04	0.03	0.05	6.1	4.2	7.5
7. Cheese	3.5	47.6	2.2	48.3	4.2	47.2	3.3	2.0	3.9	0.25	0.15	0.30	8.3	5.2	9.7
8. 100% juice	−19.1	−54.8	−11.7	−54.4	−22.5	−55.0	−3.1	−1.9	−3.6	0.00	0.00	0.00	−0.3	−0.2	−0.4
9. Juice drinks	−43.0	−69.4	−49.2	−96.5	−40.7	−58.4	−6.0	−6.9	−5.6	0.00	−0.01	0.00	−2.8	−3.3	−2.6
10. Peanut butter	−1.9	−44.9	−1.3	−45.1	−2.1	−44.8	−3.5	−2.4	−3.9	−0.28	−0.20	−0.32	−2.5	−1.7	−2.9
11. Cereals	0.5	3.5	0.4	4.0	0.6	3.3	0.6	0.4	0.7	0.01	0.00	0.01	0.7	0.5	0.9
12. Yogurt	−0.8	−6.2	−0.5	−6.2	−1.0	−6.2	−0.2	−0.1	−0.2	0.00	0.00	0.00	−0.1	−0.1	−0.2
13. Diet CSD	7.2	10.8	4.2	10.8	8.4	10.8	0.0	0.0	0.0	0.00	0.00	0.00	0.3	0.2	0.3
14. Bottled water	−36.0	−34.6	−21.9	−34.1	−42.4	−34.9	0.0	0.0	0.0	0.00	0.00	0.00	−0.1	−0.1	−0.1
15. Canned/dried fruits	1.1	15.4	0.8	15.9	1.3	15.1	0.7	0.5	0.9	0.00	0.00	0.00	0.2	0.1	0.2
16. Canned vegetables	8.5	36.3	5.6	36.7	9.9	36.0	1.4	0.9	1.6	0.01	0.01	0.02	13.5	9.2	15.3
17. Frozen dinners	7.8	45.4	5.5	45.1	9.1	45.6	4.5	3.3	5.1	0.20	0.15	0.23	10.5	7.6	12.0
18. Canned soup	10.3	87.0	6.1	90.4	12.3	85.2	2.0	1.2	2.4	0.06	0.04	0.08	20.2	12.1	23.9
19. Candy	2.2	23.7	1.3	25.1	2.6	23.0	3.0	1.8	3.7	0.11	0.06	0.14	0.8	0.5	0.9
20. Ice cream	−7.3	−52.3	−5.1	−51.9	−8.3	−52.5	−2.9	−2.1	−3.3	−0.14	−0.10	−0.16	−1.1	−0.8	−1.3
21. Cakes and cookies	−1.4	−11.5	−0.9	−11.5	−1.6	−11.5	−1.8	−1.2	−2.1	−0.08	−0.05	−0.09	−1.5	−1.0	−1.8
22. Lunch meat	−0.4	−10.2	−0.3	−9.2	−0.5	−10.6	−0.3	−0.2	−0.3	−0.02	−0.01	−0.02	−1.4	−1.0	−1.7
23. Snacks	0.8	10.0	0.5	10.0	1.0	9.9	1.2	0.8	1.6	0.06	0.04	0.08	1.9	1.2	2.4

[a]Changes in purchase quantities and their t-values are medians.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

[b] Calorie, fat and sodium changes are calculated as median quantity change times average per ounce calories, fat and sodium for the corresponding food or beverage category and income class, respectively. Therefore, their t-values are the same as those for the corresponding predicted purchase quantity change.

Note: Nielsen's projection factors are used as weights. Results are based on the incomplete demand model that accounts for price endogeneity with L = 5.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 5

Average Effects of a Half-Cent per Ounce Increase in SSB Prices on Nutrition and Welfare

	Incomplete demand, L = 5						Conditional demand, L = 4					
	Accounting for price endogeneity			Without accounting for endogenous prices			Accounting for price endogeneity			Without accounting for endogenous prices		
	All income	**Low income**	**High income**	**All income**	**Low income**	**High income**	**All income**	**Low income**	**High income**	**All income**	**Low income**	**High income**
Per capita changes in daily												
energy from SSBs (kcal)	−15.1 (−92.9)	−17.5 (−118.4)	−13.9 −79.3)	−18.9 (−662.5)	−20.7 (−855.2)	−18.1 −587.6)	−13.5 (−72.3)	−16.2 (−80.8)	−12.1 (−67.5)	−20.8 (−543.5)	−23.9 (−582.9)	−19.4 (−523.0)
sodium from SSBs (mg)	−7.1 (−110.8)	−7.5 (−134.6)	−6.9 (−98.7)	−7.9 (−608.9)	−8.3 (−785.2)	−7.8 (−549.5)	−6.0 (−80.9)	−6.9 (−86.6)	−5.5 (−77.7)	−8.4 (−472.5)	−9.5 (−513.3)	−7.9 (−454.0)
Per capita changes in daily												
energy from the 23 foods and beverages (kcal)	−7.9 (−23.6)	−13.2 (−55.3)	−5.6 (−15.2)	−17.6 (−281.7)	−20.0 (−473.3)	−16.4 (−243.1)	−15.4 (−48.3)	−18.2 (−60.0)	−14.1 (−43.2)	−22.3 (−548.1)	−26.1 (−645.9)	−20.5 (−521.0)
fat from the 23 foods and beverages (g)	0.2 (9.2)	0.1 (4.9)	0.2 (11.5)	0.0 (17.9)	0.0 (4.9)	0.1 (21.0)	−0.2 (−8.3)	−0.2 (−9.1)	−0.2 (−7.8)	−0.1 (−22.9)	−0.1 (−34.4)	−0.1 (−18.5)
sodium from the 23 foods and beverages (mg)	49.7 (64.4)	29.3 (60.0)	60.0 (66.7)	−2.8 (−23.9)	−5.1 (−70.8)	−1.8 (−14.8)	16.6 (28.6)	14.5 (25.1)	17.6 (30.5)	−7.2 (−48.4)	−9.9 (−70.3)	−6.1 (−41.8)
SSB tax paid per household ($/year)	20.48 (329.7)	23.18 (433.8)	19.22 (291.4)	19.06 (1,785.0)	21.72 (2,557.1)	17.68 (1,567.6)	15.40 (317.3)	15.03 (335.0)	15.58 (309.3)	13.30 (1,281.1)	13.02 (1,362.6)	13.44 (1,249.0)
Compensating variation per household												
stone effect ($/year)[a]	−24.15	−27.74	−22.98	−24.15	−27.74	−22.98	−18.2	−18.5	−18.1	−18.2	−18.5	−18.1
total effect ($/year)	−23.68 (−351.9)	−27.55 (−1,076.6)	−22.40 (−270.9)	−23.51 (−2,275.8)	−27.50 (−6,950.8)	−22.22 (−1,758.2)	−18.2 (−515.0)	−18.4 (−539.7)	−18.1 (−504.9)	−17.5 (−2,834.7)	−17.8 (−2,991.8)	−17.3 (−2,758.1)

[a]The Stone effect is calculated using observed budget shares and expected price changes (see equation (6) and related discussion). Therefore, the Stone effect is nonstochastic for each observation and does not have a t-value.

Note: Per capita calorie and nutrient changes for SSBs and the 23 foods and beverages (including SSBs) are medians of predicted changes over all observations for the corresponding income class. t-values in parentheses.