# Pre-school children's demand for sugar sweetened beverages: Evidence from stated-preference panel data

Anthony Scott, Peter Sivey, Ou Yang


## Abstract

The aim of this paper is to examine the impact of price changes on children's consumption of sugar sweetened beverages. Using micro-level panel data obtained from a stated preference experiment, we specify a two-sided censoring demand system model with fixed effects. Given our assumption of a two-step hierarchical structure of the censoring generating process, we propose a new test for selectivity bias and a new consistent two-step semiparametric estimation framework. The economic restrictions implied by consumption theory are imposed through a consistent and asymptotically efficient GMM estimator. We analyse the consumption behaviour of subjects through estimated expenditure and price elasticities. The partial elasticities of demand with respect to attributes of soft drinks are also estimated and examined. Our results indicate that after accounting for the income effects, the own-price elasticities for Fizzy, Juice and Cordial are respectively -0.727, -0.112 and -0.918, and therefore, are all price-inelastic. All compensated cross-price elasticities are positive, indicating that the drinks are net substitutes. The cross-drink effects of attributes show that in general, healthier Juice drinks will significantly crowd out the consumption of Fizzy drinks. Fizzy drinks containing no added colours or preservatives will crowd out Juice but with a relatively small effect. However, there seems to be no consistent crowd-out effect between Fizzy drinks and Cordial or between Juice and Cordial.


*Key words*: sugar sweetened beverages, consumption behaviour, panel data, demand system, censoring

# I. Introduction

Consumption of sugar sweetened beverages (SSBs) exhibits strong associations with weight gain, obesity, and dental caries, especially in young children and for children of low socio-economic status (Malik, Schulze, and Hu 2006). These problems affect about one-third of children of pre-school age, with 13% of children aged 2-3 years old consuming SSBs every day (Wake et al. 2006; Dubois et al. 2007).

There are strong arguments, and numerous examples, of taxes on SSBs (Brownell and Frieden 2009). The use of taxes to improve population health is controversial. The evidence of a net welfare gain is mixed, and depends on the effects on the consumption of other foods and beverages (Sharma et al. 2014). Arguments as to whether such taxes are regressive depend on how the price elasticity of demand varies across sub-groups of the population (Sharma et al. 2014). Recent previous studies of the impact of taxation on consumption have either estimated average price elasticities (e.g. Finkelstein et al. 2013, Zhen et al. 2014, Briggs et al. 2013), or have examined heterogeneity amongst moderate and high consumers (Etilé and Sharma 2015) or different income groups (Sharma et al. 2014). Examining the impact of changes in price on high risk populations is therefore important in examining the overall effectiveness of taxation on population health.

The aim of this paper is to examine the impact of price changes on children's consumption of SSBs. We examine price and cross-price elasticities across SSBs. Usual datasets use household scanner data or aggregated data for small areas and so do not have information on the consumption of SSBs by children within households due to aggregation assumptions. Data disaggregated to below household level is generally not available. We use unique micro-data from a stated preference experiment administered to parents of children from a birth cohort study of 500 children (de Silva-Sanigorski et al. 2011). Stated preference experiments use hypothetical choices of goods to examine the impact of prices and other characteristics on choices. Unlike stated preference discrete choice experiments which focus on choosing one good from several alternatives, our consumption experiment was designed to capture, first, the number of bottles of cordial, fruit juice and fizzy drink bought for the household, conditional on their price and other characteristics. Second, respondents were asked how many glasses of each were consumed by children in the household. This provides a continuous measure of consumption suited to analysis using a demand system approach that allows for i) selection (where no soft drinks are consumed at all), ii) censoring (zero consumption of at least one SSB), iii) panel data (multiple scenarios per respondent). We also therefore contribute to the literature on the analysis of stated preference experiments. A particular advantage of such an experiment is that prices are presented to respondents exogenously. In addition, an experimental design is used to ensure that the variation in the attributes is orthogonal and that standard errors are minimised.

The plan of this paper is set out as follows. Section II introduces the consumption experiment and describes the data. Section III presents the model specification and estimation strategy. The estimation results and corresponding discussions are given in Section IV. The last section concludes this paper.

## II. A consumption experiment and data

The consumption experiment (CE) consists of presenting survey respondents, who are parents of 24-month-old children, with a series of hypothetical scenarios about the quantities of alternative drink types for their family's and children's consumption. The CE is a labelled design, where respondents choose consumption levels for four broad categories of drinks: Fizzy Drink, Juice, Cordial and Tap Water. The soft drink categories are characterised by four attributes: price, sugar content, added vitamins and no added colours or preservatives. The tap water category is not described by any attributes.

We undertook an extensive pre-piloting phase with in-depth interviews of 32 families to develop the four labelled drink categories, the attributes of the drinks, and the nature of the choice task. The pre-pilot was an iterative process, where initial designs were drafted, presented to potential respondents during interviews, and attributes and labels refined before being presented again to potential respondents. This process broadly followed the recommendations of Coast et al. (2012) in that we avoided describing the latent construct (eg "the drink is tasty" or "the drink is healthy"), used in-depth interviews and broadly followed a constant-comparative approach to qualitative data collection and analysis. More details of the qualitative approaches used are detailed in de Silva-Sanigorski et al. (2011) and Hoare et al. (2014).

The choice context, attributes and levels were informed by three considerations. Firstly, some attributes were of particular policy interest, including price and sugar content of drinks. Secondly, we conducted an investigation of the websites of major Australian supermarket chains. This was a key step as it enabled us to ensure the hypothetical choices were as close as possible to real-world choices that parents would be making whilst shopping for drinks. Thirdly, all of our decisions were informed verified and modified from the iterative process of the qualitative interviews.

Our consumption experiment is set in the context of the main 'family shop' (e.g. Saturday shop in a supermarket). It was recognised in qualitative work that young children's drink consumption was particular to context and was particularly idiosyncratic out of the household (on trips or visiting friends and family) and on special occasions (Hoare et al. 2014), however it would be difficult to model consumption in all of these alternative contexts comprehensively. The regular family shop provides a well-understood context which accounts for a large proportion of a child's drink intake.

Our design takes into account that the supermarket shop typically involves a choice of drinks for the family, not just for the child. So, for example, a large bottle of juice could be bought with the intention of providing drinks for adults and older children in the household as well as for young children. For this reason we ask responding parents to make two sequential consumption choices in each scenario: first they must decide how many bottles of each soft drink to buy for the week for the whole family; secondly, they must decide how many glasses of each drink they would give to their young child to drink for the week.

Our four categories of drinks (fizzy drink, juice, cordial and tap water) were chosen as the most common broad categories of drinks given to young children. A decision was made early to exclude milk and milk-based drinks as they form a separate category of drinks which can be consumed for nutritional reasons. Tap water is included as a labelled drink category but is not described by the attributes. We assume tap water is regarded as free of charge and homogeneous to the families. The other free drink types can be described by all four attributes: price, sugar content, added vitamins, and no added colours or preservatives.

Price is a key determinant of choice, displayed prominently in supermarkets, mentioned by interviewees as determining their choice and is of policy and academic interest. The three price levels chosen, \$0.90, \$2.95, and \$4.98 per two litre bottle were designed to cover the full range of prices encountered in supermarkets.

The sugar content attribute is another key policy attribute in the study. The attribute has two only two levels, 'Diet-No Sugar' or blank, implying 'with sugar'. We chose this wording to match real-life labelling of drinks, 'Diet' or 'No Sugar' or very similar variants were used on the packaging of sugar free drinks, whereas highly sugar sweetened drinks were not labelled with regard to sugar content. One exception to this wording was for the 'Juice' drinks category, for which we used the wording 'No added sugar' instead of 'Diet-No Sugar', again matching the labelling most often used in supermarkets. The final two binary attributes represent common health claims made by soft drink labels: "Extra vitamins A and C" and "No added colours or preservatives". Each of these attributes is blank when there are no extra vitamins or when there may be added colours or preservatives.

In our final panel data set, there are 204 parents whose consumption choices for their pre-school children are observed for nine hypothetical scenarios which are different in terms of attributes and prices of soft drinks. Therefore, there are in total 1836 observations in our sample when laid out as one long cross-section, 47 of which contain missing values and are excluded.

The demand system literature features modelling demand in budget share form on goodness of fit grounds, which also helps avoid heteroscedasticity (Leser 1963). The first problem with using budget share form is that for respondents who decided not to give their children any of the three soft drinks their total expenditure on soft drinks is zero. Budget shares are not defined or defined as missing values, and therefore a budget allocation analysis framework should not include these observations, given that the framework is built on the condition of positive total expenditure.

These zero observations on total soft-drink expenditure can be regarded as a result of the first-stage budget allocation problem in a multi-stage budgeting framework (Deaton and

Muellbauer 1980a, Deaton and Muellbauer 1980b, Edgerton 1997). Specifically, assuming weak separability between soft drinks and water at the first stage, the subject first makes a decision on whether or not he/she will purchase some soft drinks and how much money in total he/she will spend on soft drinks. If he/she decides to purchase some soft drinks, i.e. spend a positive amount of money on total soft drinks, he/she then proceeds to the second stage to make a decision about how to allocate the total soft drink budget among the three drinks. Hence, although all recorded as zeros in expenditure form, it is clearly seen while in share form that these zeros should be treated differently while modelling their generating process.

Table 1 and Table 2 summarise the number of total, positive and zero observations and sample mean and standard deviation for total expenditure on soft drinks and for shares of the three soft drinks considered. As can be seen in Table 1, for each of these four variables, there are a substantial proportion of zero observations which need to be accunted for in the econometric analysis.

**Table 1 Number of positive and zero observations for total expenditure and shares of soft drinks**

| | Total obs. | Positive obs. | Zero obs. |
|---|---|---|---|
| Total expenditure | 1789 (100%) | 868 (48.52%) | 921 (51.48%) |
| Fizzy share | 868 (100%) | 167 (19.24%) | 701 (80.76%) |
| Juice share | 868 (100%) | 730 (84.10%) | 138 (15.90%) |
| Cordial share | 868 (100%) | 304 (35.02%) | 564 (64.98%) |

*Note*: observation is abbreviated to obs.

**Table 2 Summary of total expenditure and shares of soft drinks for total, positive and zero observations**

| | Total obs. | | Positive obs. | | Zero obs. | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| Total expenditure | 0.688 | (1.359) | 1.419 | (1.665) | 0.000 | (0.000) |
| Fizzy share | 0.103 | (0.260) | 0.534 | (0.348) | 0.000 | (0.000) |
| Juice share | 0.750 | (0.377) | 0.892 | (0.206) | 0.000 | (0.000) |
| Cordial share | 0.147 | (0.304) | 0.419 | (0.388) | 0.000 | (0.000) |

*Note*: observation is abbreviated to obs. Standard deviation are given in parentheses.

## III. Model Specification and Two-step Estimation Strategy

As is shown in the last section, there exist a substantial proportion of zero observations for each of the four dependent variables. In the literature, the two principal reasons for zero expenditures in microeconomic expenditure data are consumers at a corner solution for the commodity in question (Wales and Woodland 1983), and limited survey periods leading to infrequency of purchase (Deaton and Irish 1984). To our knowledge, most of the econometric techniques in the literature are developed to model economics non-consumption (for example Yen and Lin 2006, Meyerhoefer, Ranney, and Sahn 2005, Yen 2005, Perali and Chavas 2000, Heien and Wessells 1990). The only exception is Deaton and Irish (1984). Since our longitudinal data come from a consumption experiment, the zero expenditure observations, in the case of our experiment, represent a genuine corner solution where the subject deliberately chooses not to consume particular goods given the attributes of the soft drinks, the appearance of tap water and the prices of the soft drinks in scenarios.

This study employs a fixed-effects censored demand system analysis framework, to account for the reported zero expenditure observations on certain soft drinks (i.e. choose to or not to purchase and give their children certain soft drinks). The fixed effect censored demand system model is estimated using our micro-level panel data. With the increasing availability of micro-data, the use of such individual-level data is preferable, since it avoids the problem of aggregation over individuals and often provides a large and statistically rich sample (Heien and Wessells 1990).

Much of the recent empirical effort on censored demand system has been concerned with circumventing the "curse of dimensionality" associated with the theoretically consistent models proposed by Wales and Woodland (1983) and Lee and Pitt (1986, 1987). For example, Heien and Wessells (1990), Shonkwiler and Yen (1999) and Yen, Kan, and Su (2002) adopt a two-step procedure to reduce the computational burden from using a full information maximum likelihood estimator. Nonetheless, Arndt, Liu, and Preckel (1999) claimed that this procedure and its application to corner solutions are unable to account for the role of reservation prices. Instead, Arndt (1999) proposed to address this difficulty using maximum entropy (ME) techniques, and generate a simpler framework for the imposition of regularity conditions. However, the fact that the asymptotic properties of this estimator are not well understood in nonlinear applications limits its feasibility.

More recently, Yen, Lin, and Smallwood (2003) use the simulation technique, as well as a quasi-maximum likelihood procedure, to facilitate the estimation of a censored demand system based on the Amemiya-Tobin general model structure. Yen and Lin (2006) adopted a sample selection approach to estimating a system involving a small number of commodities using a full information maximum likelihood estimator. Perali and Chavas (2000) have developed a consistent approach to the problem based on generalized method of moments (GMM) techniques. While all of the above studies provide an approach to obtaining consistent estimates of disaggregated demand models, they are designed for cross-sectional data and thereby, they suffer from limited ability to control for heterogeneous preferences and limited variation in real price. To the best of our knowledge, Meyerhoefer, Ranney, and Sahn (2005) is the only work which extends this literature to the context of panel data. They proposed a consistent correlated random-effects GMM estimation framework for censored demand system applications using panel data, and controlled for unobserved heterogeneity using a correlated random-effects specification.

Given the longitudinal structure of our micro-level data, it seems natural for us to follow Meyerhoefer, Ranney, and Sahn (2005)'s estimation strategy. However, a general flexible demand system analysis model, such as AIDS and QUAIDS, requires positive expenditure to be observed for at least one of the three soft drinks; in other words, subjects' total expenditure on all the three soft drinks has to be positive. Even though Meyerhoefer, Ranney, and Sahn (2005)'s censored demand system model is able to handle zero expenditure observations for certain goods, if a subject is observed to have purchased nothing, this observation has to be excluded from the estimation. This is because the logarithm of total expenditure is used as an explanatory variable in the system specification and also because a normal formula to calculate expenditure elasticity involves the logarithm of total expenditure.

As shown in Table 1, 51.48% of the total 1789 non-missing observations show zero total expenditure on soft drinks. Employing Meyerhoefer, Ranney, and Sahn (2005)'s correlated random-effects censored demand system analysis framework will exclude these observations from estimation, which one might find similar to an incidental truncation problem. If a subject's decision about whether or not to give their children any soft drink is not systematically related to their decision about how much of each soft drink to give to their children, estimates conditional on the truncated sample (or equivalently, conditional on positive total expenditure on soft drinks) are still consistent; otherwise, a sample selection bias might result. Accordingly, a statistical test for this potential selection bias is necessary to implement.

Before proceeding to the introduction of a statistical test for selection bias in the current context, let us first derive the share equations for a censored demand system model whereby price and expenditure elasticities can be estimated. Conditional on positive total expenditure on soft drinks, the subject makes decisions on how to allocate the total expenditure among individual soft drinks in scenarios given the price and attributes of each drink. In accordance with neoclassical consumption theory, define the direct utility function as $U(q_{jt}, q_{jt}^{w}; d_{1jt}, \ldots, d_{Ljt}, \varphi_j)$, where $t$ $(=1,\ldots,T)$ indexes scenarios, $j$ $(=1,\ldots,J)$ denotes subjects or decision makers, $q_{jt} = (q_{1jt}, \ldots, q_{Kjt})'$ is a vector containing subject $j$'s demand levels for the $k$th soft drink in scenario $t$, $q_{jt}^{w}$ denotes the quantity of tap water chosen by subject $j$ in scenario $t$, $d_{ljt}$ denotes the realisation of the $l$ th $(=1,\ldots,L)$ attribute for subject $j$ $(=1,\ldots,J)$ at scenario $t$ $(=1,\ldots,T)$, and $\varphi_j$ is a time invariant individual specific effect representing unobserved heterogeneity across subjects.

It is further assumed that $U(.)$ represents a preference ordering of the PIGLOG form. Then, according to duality theory (Deaton and Muellbauer 1980b), the indirect utility function corresponding to Deaton and Muellbauer (1980a) can be specified as:

$$V_{jt}^{*} = \frac{\log c_{jt} - \alpha_0 - \sum_k \alpha_k \log p_{kjt} - \sum_k \sum_l \lambda_{kl} \log p_{kjt} d_{ljt} - \sum_k \omega_k \log p_{kjt} \log q_{jt}^{w}}{\beta_0 \prod_k p_{kjt}^{\beta_k}} - \frac{\frac{1}{2} \sum_k \sum_i \tilde{\gamma}_{ki} \log p_{kjt} \log p_{ijt} - \sum_k \psi_k \log p_{kjt} \varphi_j}{\beta_0 \prod_k p_{kjt}^{\beta_k}} \tag{3.1}$$

where $\log c_{jt}$ represents the total expenditure on soft drinks at scenario $t$ for household $j$ and $p_{kjt}$ denotes the price of soft drink $k$ observed at scenario $t$ by subject $j$.

The attributes of soft drinks, the subject's individual specific effects, the quantity of tap water and stochastic error items are embedded into the demand model following a procedure named "demographic translating". This procedure is general in the sense that the demographically extended demand system are still theoretically plausible, if the initial demand system is theoretically plausible (Pollak and Wales 1981, Pollak and Wales 1992).

Demand equations are most easily represented in share form, to be more consistent with an assumption of homoscedasticity and to remove dependence on the numeraire (Fry, Fry, and McLaren 1996). Applying the logarithm version of Roy's Identity, the Marshallian uncompensated demand share equations of the demographically extended Almost Ideal Demand System (AIDS) are obtained and their econometric specification is shown as follows:

$$w_{njt}^{*} = \alpha_n + \sum_l \lambda_{nl} d_{ljt} + \omega_n \log q_{jt}^{w} + \sum_k \gamma_{nk} \log p_{kjt} + \beta_n (\log c_{jt} - \log P_{jt}) + \rho_{nj} + u_{njt} \qquad (3.2)$$

where $w_{njt}^{*}$ is the expenditure share of soft drink $n\ (=1,\ldots K)$ at scenario $t$ for household $j$,

$$\log P_{jt} = \alpha_0 + \sum_k \alpha_k \log p_{kjt} + \sum_k \sum_l \lambda_{kl} \log p_{kjt} d_{ljt} + \sum_k \omega_k \log p_{kjt} \log q_{jt}^{w} + \frac{1}{2}\sum_k \sum_i \gamma_{ki} \log p_{kjt} \log p_{ijt} + \sum_k \psi_k \log p_{kjt}\varphi_j,$$

$\rho_{nj} = \psi_n \varphi_j$, $\gamma_{ki} = {}^{1}\!/_{2}(\tilde{\gamma}_{ki} + \tilde{\gamma}_{ik})$, and it is assumed that $u_{njs}$ and $u_{njt}$ are identically distributed conditional on $\left(\rho_{nj},\, x_{j1},\ldots,\, x_{jT}\right)$ for any $s,\ t \le T$, where $x_{js}$ and $x_{jt}$ denote observable explanatory variables in (3.2) at scenarios $s$ and $t$.

In order to linearize the above budget share equation and circumvent the problem that incorporating demand shifters in the intercepts renders the AIDS model invariant to units of measurement (Alston, Chalfant, and Piggott 2001). One way to solve the problem is to use a scale-invariant log-linear Laspeyres index, $\log P_{jt}^{S} = \sum_{k} w_{k}^{o} \log p_{kjt}$ where $w_{k}^{o}$ is the mean share for soft drink $k$ across all the subjects and all the scenarios, to replace $\log P_{jt}$ in the AIDS model, which has been shown by Moschini (1995) and Buse (1998) to have good approximation properties. This new price index can also reduce the potential for severe multicollinearity problem while reducing the burden of estimation. Homogeneity and symmetry restrictions implied from consumption theory can be imposed on the demand equations through restrictions on certain parameters as follows: $\sum_{k} \gamma_{ik} = 0$ and $\gamma_{ki} = \gamma_{ik}$.

The adding-up condition is not imposed *a priori*, because although the observed budget shares add up to one, the latent shares need not, which should have limited impact on the price coefficients since they sum up to zero across equations by default once symmetry and homogeneity restrictions have been imposed. Later on, it will also be seen from the results that our estimates approximately satisfy the adding-up condition, even if adding-up has not been explicitly imposed *a priori*.

The share equations in (3.2) can be regarded as latent share equations (Wales and Woodland 1983). In reality, demand shares are bounded between zero and unity. Thus, observed shares $w_{njt}$ relate to latent shares $w_{njt}^{*}$ such that

$$w_{njt} = \begin{cases} 0 & \text{if} \quad w_{njt}^{*} < 0 \\ w_{njt}^{*} & \text{if} \quad 0 \le w_{njt}^{*} \le 1 \\ 1 & \text{if} \quad w_{njt}^{*} > 1 \end{cases},$$

From (3.2), it can be clearly seen that any observation with total expenditure, $c_{jt}$, being zero will be excluded from the estimation.

To test the significance of the potential sample selection bias, a variable addition test, similar in spirit to Wooldridge's (1995) variable addition tests for selection bias (also see Wooldridge 2010a), is proposed and applied in this study. Specifically, the selection mechanism is specified as an equation of the Tobit form, as follows:

$$\begin{aligned} c_{jt}^{*} &= \alpha_0 + \sum_l \alpha_l d_{ljt} + \sum_k \beta_k \log p_{kjt} + \gamma q_{jt}^{w} + \eta_j + \varepsilon_{jt} \\ c_{jt} &= \max(0,\ c_{jt}^{*}) \\ \varepsilon_{jt} &\sim N\left(0,\ \sigma_{\varepsilon}^{2}\right) \end{aligned} \tag{3.3}$$

Combining the latent equations in (3.2) and (3.3), for each soft drink $n\ (=1,\dots K)$ introduces the following fixed-effects selection system:

$$w_{njt}^{*} = \alpha_n + \sum_l \lambda_{nl} d_{ljt} + \sum_k \gamma_{nk} \log p_{kjt} + \beta_n (\log c_{jt} - \log P_{jt}) + \rho_{nj} + u_{njt} \tag{3.4}$$

$$c_{jt}^{*} = \alpha_0 + \sum_l \alpha_l d_{ljt} + \gamma q_{jt}^{w} + \sum_k \beta_k \log p_{kjt} + \eta_j + \varepsilon_{jt} \tag{3.5}$$

Since the unobservable individual specific effect $\eta_j$ in (3.5) is expected to be correlated with individual tap water consumption, using a Mundlak-type model, as illustrated in (Wooldridge 2010a), this correlation can be modelled as a linear projection of $\eta_j$ on the average tap water consumption across all the scenarios, denoted by $\overline{q_j^{w}}$ :

$$\eta_j = \lambda \overline{q_j^{w}} + \nu_j \tag{3.6}$$

where $\nu_j$ is assumed to be independent of the exogenous regressors, $\varepsilon_{jt}$ and $u_{njt}$, for any $n$, and is distributed $N(0, \sigma_{\nu}^{2})$. Substituting in $\eta_j$, the selection equation (3.5) can be written as:

$$c_{jt}^{*} = \alpha_0 + \sum_l \alpha_l d_{ljt} + \gamma q_{jt}^{w} + \sum_k \beta_k \log p_{kjt} + \lambda \overline{q_j^{w}} + \xi_{jt} \tag{3.7}$$

where $\xi_{jt} = v_j + \varepsilon_{jt}$. As indicated in the last section, this equation (3.3) can be considered as a reduced form of a first/upper-stage budget allocation problem, and it should also be mentioned that this test is under the assumption that the latent variable determining selection can be observed whenever it is nonnegative, but for the purpose of test, the selection mechanism does not have to be correctly specified in any sense, as it simply serves as a vehicle for obtaining a sensible test (Wooldridge 1995)

If it is assumed that there is no selectivity bias, since $w^*_{njt}$ in (3.4) is only partially observed, a normal linear fixed-effects estimation strategy for (3.4) produces inconsistent estimates. Therefore, Alan et al. (2014)'s semi-parametric estimator for two-sided censoring models with fixed effects is employed. Denote the observed explanatory variables in (3.4) as $x_{jt}$ and let $x_j \equiv \left(x_{j1}, \ldots, x_{jT}\right)'$ and $\xi_j = \left(\xi_{j1}, \ldots \xi_{jT}\right)'$. Under the assumptions that for any $n$, $u_{njt}$ is identically distributed conditional on $\left(\rho_{nj},\, x_j\right)$ and $E\left(u_{njt} \mid \rho_{nj}, v_j,\, x_j, \xi_j\right) = 0$, the semi-parametric estimator is consistent and asymptotically normal and there is no selectivity bias (Alan et al. 2014). This equation also suggests a useful alternative that implies selectivity bias. The simplest such alternative is

$$E\left(u_{njt} \mid \rho_{nj}, v_j,\, x_j,\, \xi_j\right) = \theta_n \varepsilon_{jt} = \theta_n \left(\xi_{jt} - v_j\right), \qquad t = 1, 2, \ldots, T, \tag{3.8}$$

for some unknown scalar $\theta_n$.

Under the alternative (3.8), we have

$$\begin{aligned} w^*_{njt} &= \alpha_n + \sum_l \lambda_{nl} d_{ljt} + \sum_k \gamma_{nk} \log p_{kjt} + \beta_n (\log c_{jt} - \log P_{jt}) + \rho_{nj} + \theta_n \left(\xi_{jt} - v_j\right) + \omega_{njt} \\ &= \alpha_n + \sum_l \lambda_{nl} d_{ljt} + \sum_k \gamma_{nk} \log p_{kjt} + \beta_n (\log c_{jt} - \log P_{jt}) + \theta_n \xi_{jt} + \rho_{nj} - \theta_n v_j + \omega_{njt} \end{aligned} \tag{3.9}$$

where $\omega_{njt} = u_{njt} - \theta_n\left(\xi_{jt} - v_j\right)$. From (3.9), it follows that if we could observe $\xi_{jt}$, when $c^*_{jt} > 0$, then we could test the null hypothesis by including the $\xi_{jt}$ as an additional regressor in the semi-parametric fixed-effects estimation and testing $\mathrm{H}_0:\ \theta_n = 0$ using standard methods. While $\xi_{jt}$ is not observable, it can be estimated whenever $c^*_{jt} > 0$ because $\xi_{jt}$ is simply the error in a Tobit model. Therefore, the following test for selection bias when $c_{jt} > 0$ is proposed:

*Step 1:* Estimate the equation (3.7) by pooled Tobit.

*Step 2:* When $c_{jt} > 0$, calculate the Tobit residuals:

$$\hat{\xi}_{jt} = c_{jt} - \left(\hat{\alpha}_0 + \sum_l \hat{\alpha}_l d_{ljt} + \hat{\gamma} q^w_{jt} + \sum_k \hat{\beta}_k \log p_{kjt} + \hat{\lambda}\overline{q^w_j}\right)$$

*Step 3:* Estimate the equation

$$\begin{aligned} w^*_{njt} = \alpha_n + \sum_l \lambda_{nl} d_{ljt} + \sum_k \gamma_{nk} \log p_{kjt} + \beta_n(\log c_{jt} - \log P_{jt}) \\ + \theta_n \hat{\xi}_{jt} + \left(\rho_{nj} - \theta_n v_j\right) + \omega_{njt}, \end{aligned} \tag{3.10}$$

using those observations for which $c_{jt} > 0$.

*Step 3:* Test $\mathrm{H}_0:\ \theta_n = 0$ using the *t*-statistic for $\hat{\theta}_n$.

As mentioned above, $w^*_{njt}$ is only partially observed, the normal linear fixed-effects estimation produces inconsistent estimates. Therefore, Alan et al. (2014)'s consistent semi-parametric estimator for two-sided censoring models with fixed-effects is employed to estimate parameters in equation (3.10). In particular, let $\delta_n$ denote coefficients in (3.10) to be estimated and $x_{jt}$ denote the vector of all the explanatory variables in (3.10) including $\hat{\xi}_{jt}$.

Since, under the null hypothesis, $\omega_{njs}$ and $\omega_{njt}$ are identically distributed conditional on $\left(\rho_{nj}, x_j, v_j, \xi_{jt}\right)$ for any $s,\ t \leq T$, $\delta_n$ can be consistently estimated as follows,

$$\hat{\delta}_n = \arg\min_{\delta} \sum_{j=1}^{J} \sum_{1<s<t<T_j} \frac{1}{T_j} U\left(w_{njt}, w_{njs}, \left(x_{jt} - x_{js}\right)' \delta\right) \tag{3.11}$$

where

$$U\left(y_1, y_2, d\right)\begin{cases} 1+2c_1+c_1^2-2c_3c_1+2c_3c_2+\left(y_1-y_2-c_2\right)^2 & \text{for } d<-1 \\ -2d-d^2+2c_1+c_1^2-2c_3c_1+2c_3c_2+\left(y_1-y_2-c_2\right)^2 & \text{for } -1\leq d<c_1 \\ -2c_3d+2c_3c_2+\left(y_1-y_2-c_2\right)^2 & \text{for } c_1\leq d<c_2 \\ \left(y_1-y_2-d\right)^2 & \text{for } c_2\leq d<c_3 \\ -2c_2d+2c_2c_3+\left(y_1-y_2-c_3\right)^2 & \text{for } c_3\leq d<c_4 \\ -d^2+2d+c_4^2-2c_4-2c_2c_4+2c_2c_3+\left(y_1-y_2-c_3\right)^2 & \text{for } c_4\leq d<1 \\ 1+c_4^2-2c_4-2c_2c_4+2c_2c_3+\left(y_1-y_2-c_3\right)^2 & \text{for } d\geq 1 \end{cases}$$

and

$c_1 = \min\left\{-y_2, y_1-1\right\}$, $c_2 = \max\left\{-y_2, y_1-1\right\}$, $c_3 = \min\left\{1-y_2, y_1\right\}$ and $c_4 = \max\left\{1-y_2, y_1\right\}$.

The rationale behind this estimator is that for example, if $E\left(\varepsilon x\right)=0$, then one has the moment conditions $E\left[\left(y^*-x'\beta\right)x\right]=0$, where $y^*$ denotes the latent variable. However, with censoring, $y-x'\beta$ will not have the same properties as $\varepsilon$. The idea employed in Alan et al. (2014), and some others such as Powell (1986), Honoré (1992) and Honoré and Powell (1994), is to apply additional censoring to $y-x'\beta$ in such a manner that the resulting re-censored residual satisfies the conditions assumed on $\varepsilon$. The minimisation problem (3.11) is convex and has as first-order condition the sample analogue of moment conditions as follows:

$$E\left[\sum_{1<s<t<T_j}\frac{1}{T_j}u\left(w_{njt},\mathrm{w}_{njs},\Delta x_j{}'\delta_n\right)\Delta x_j\right]=0 \tag{3.12}$$

where $\Delta x_j \equiv x_{jt} - x_{js}$

and

$$u\left(y_1,\mathrm{y}_2,d\right)=\begin{cases}0 & \text{for } d<-1\\ 1+d & \text{for } -1\le d<c_1\\ \min\left\{1-y_2,\ y_1\right\} & \text{for } c_1\le d<c_2\\ y_1-y_2-d & \text{for } c_2\le d<c_3\\ \max\left\{y_1-1,\ -y_2\right\} & \text{for } c_3\le d<c_4\\ d-1 & \text{for } c_4\le d<1\\ 0 & \text{for } d\ge 1\end{cases}$$

Under $\mathrm{H}_0:\ \theta_n=0$,

$$\sqrt{J}\left(\hat{\delta}_n-\delta_n\right)\xrightarrow{d}N\left(0,\Gamma^{-1}V\Gamma^{-1}\right) \tag{3.13}$$

where $\Gamma$ and $V$ are consistently estimated as follows (Alan et al. 2014):

$$\hat{\Gamma}=\frac{1}{J}\sum_{j=1}^{J}\left[\sum_{s<t}\frac{1}{T_j}1\left\{-1<\left(x_{js}-x_{jt}\right)'\hat{\delta}_n<1\right\}\begin{pmatrix}1\left\{-1<\left(x_{js}-x_{jt}\right)'\hat{\delta}_n<w_{js}-1\right\}-\\ 1\left\{0<\left(x_{js}-x_{jt}\right)'\hat{\delta}_n<w_{js}\right\}-\\ 1\left\{-w_{jt}<\left(x_{js}-x_{jt}\right)'\hat{\delta}_n<0\right\}+\\ 1\left\{1-w_{jt}<\left(x_{js}-x_{jt}\right)'\hat{\delta}_n<1\right\}\end{pmatrix}\left(x_{js}-x_{jt}\right)\left(x_{js}-x_{jt}\right)'\right] \tag{3.14}$$

and

$$\hat{V} = \frac{1}{J}\sum_{j=1}^{J}\hat{v}_j\hat{v}_j{}'$$

with

$$\hat{v}_j = \sum_{s<t}\frac{1}{T_j}u\left(w_{js}, w_{jt}, \left(x_{js} - x_{jt}\right)'\hat{\delta}_n\right)\left(x_{js} - x_{jt}\right).$$

In cases where the null hypothesis is rejected, the model has to be corrected for selection bias. In particular, the addition item $\hat{\xi}_{jt}$ is kept in the equation, and an adjustment to the asymptotic variance of $\sqrt{J}\left(\hat{\delta}_n - \delta_n\right)$ is needed. This is because, letting $\tau$ denote coefficients in (3.7), with the estimation involving $\hat{\xi}_{jt}$, $\hat{\delta}_n$ is a function of $\hat{\tau}$. Therefore, the asymptotic variance estimator of $\hat{\delta}_n$ must account for the asymptotic variance of $\hat{\tau}$. It can be shown that

$$\sqrt{J}\left(\hat{\delta}_n - \delta_n\right) \xrightarrow{d} N\left(0, A^{-1}DA^{-1}\right).$$

The formulae of $A$ and $D$, as well as their estimators, can be found in Appendix.

***Generalized Method of Moments Estimation framework***

Once the consistent equation-by-equation estimates are obtained for each soft drink, following Meyerhoefer, Ranney, and Sahn (2005), the cross-equation homogeneity and symmetry restrictions on $\gamma_{nk}$'s, implied from the consumption theory, are imposed through a minimum distance estimator using the sample analogue of moment conditions in (3.12), to derive consistent structural parameter estimates. Specifically, denote the drink-by-drink reduced-form parameter estimates for all share equations as $\delta = \left(\delta_1', \ \delta_2', \ \delta_3'\right)'$. The structural parameters, denoted by $\pi$, can be consistently estimated as:

$$\min_{\pi} \left(\hat{\delta} - m(\pi)\right)' \mathrm{W} \left(\hat{\delta} - m(\pi)\right)$$

where $\hat{\delta}$ are consistent estimates of the reduced-form parameters $\delta$, which are obtained from drink-by-drink estimation, and $\mathrm{W}$ is the weighting matrix measuring the distance between the sample moments and the corresponding population moments. $m(.)$ is a function mapping $\pi$ into $\delta$, which is used to impose restrictions implied from demand theory on the reduced form parameters. $\pi$ can be efficiently estimated if $\mathrm{W} = \Xi^{-1}$, where $\Xi$ is the asymptotic covariance matrix of $\hat{\delta}$. It can be shown that $\Xi = \mathrm{H}^{-1} S \mathrm{H}^{-1}$ (Wooldridge 2010b).

Let $S_j = \left(S_{1j}', S_{2j}', S_{3j}'\right)'$ denote subject $j$'s univariate scores of equation (3.11) for all the soft drinks and $\mathrm{H}_{nj}$ denote the univariate Hessian for soft drink $n$. Then, define $\mathrm{H}^{-1} = diag\left\{E\left(H_{1j}\right)^{-1}, \ldots, E\left(H_{Nj}\right)^{-1}\right\}$ and $S = E\left(S_j S_j'\right)$. $\Xi$ can be consistently estimated by substituting in sample analogues.

### *Elasticity Formulae*

The total expenditure and uncompensated price elasticities for any demand system are given by

$$E_n = \frac{\partial w_n}{\partial \log c} \frac{1}{w_n} + 1 \tag{3.15}$$

and

$$e_{ni} = \frac{\partial w_n}{\partial \log p_i} \frac{1}{w_n} - \delta^*_{ni}, \tag{3.16}$$

where $\delta^*_{ni}$ is the Kronecker delta, and the compensated price elasticities are derived using the Slutsky relationship as $\tilde{e}_{ni} = e_{ni} + s_i E_n$. Since, as explained in Honor (2008), the parameter estimates for the fixed-effect models can be converted to marginal effects by multiplying them by the fraction of observations that are not censored. $E_n$ and $e_{ni}$ can be estimated by

$$\hat{E}_n = \hat{\beta}_n F_n \frac{1}{\bar{w}_n} + 1 \tag{3.17}$$

and

$$\hat{e}_{ni} = \hat{\gamma}_{ni} F_n \frac{1}{\bar{w}_n} - \delta^*_{ni}, \tag{3.18}$$

where $F_n$ denotes the fraction of observations that are not censored for soft drink $n$; $\bar{w}_n$ denotes the sample mean of shares of drink $n$; $\hat{\beta}_n$ and $\hat{\gamma}_{ni}$ are the estimates of $\beta_n$ and $\gamma_{ni}$ as defined in (3.9).

## IV. Results

All tests and estimations were carried out using the R programming language. The codes can be obtained from the authors upon request. As the null hypothesis that there is no selection bias is only rejected for the share equation of Fizzy, but not for Juice or Cordial. The correction procedure is only implemented for Fizzy. Table 3 shows the structural parameter estimates. Based on these estimates, expenditure and price elasticities are derived and reported in Table 4. We have also estimated and examined the partial elasticities of demand for drink $n$ (=Fizzy, Juice or Cordial) with respect to the *j*th attribute of drink $i$, which are defined as:

$$P_{n,ij} \equiv \frac{\partial \log x_n}{\partial d_{ij}} = \frac{\partial w_n}{\partial d_{ij}} \frac{1}{w_n} \tag{4.1}$$

where $x_n$ denotes the demand for drink $n$, and can be estimated by

$$\hat{P}_{n,ij} = \hat{\lambda}_n F_n \frac{1}{\bar{w}_n}, \tag{4.2}$$

where $\hat{\lambda}_n$ is the estimate of $\lambda_n$ as defined in (3.9). The estimates and their standard errors are reported in Table 5. These results are important evidence for soft drink tax policy concerns and health-related, such obesity and dental health, campaigns.

As shown in Table 4, uncompensated own-price elasticities for all the three drinks are expectedly all negative. According to the sign of the uncompensated cross-price elasticities, Fizzy and Juice and Juice and Cordial are treated as gross complements. After accounting for the income effects, compensated own-price elasticities of all three goods are still negative, as expected, and all compensated cross-price elasticities are positive, indicating that all the three drinks are net substitutes.

As has been noticed, the adding-up condition is not imposed *a priori*. It is interesting to see if this condition is still, at least approximately, satisfied. Since the adding-up condition is equivalent to the Engel aggregation constraint: $\sum w_n E_n = 1$ (Deaton and Muellbauer 1980b), given our estimated income elasticities $\hat{E} = (1.493,\ 1.119,\ 0.378)'$ and the mean shares of drinks $\bar{w}=(0.101,\ 0.749,\ 0.150)$, $\bar{w}'\hat{E}=1.046$, which is very close to 1. Therefore, it can be seen that our estimates approximately satisfy the adding-up condition, even if it has not been explicitly *a priori* imposed in the model specification.

The partial elasticities w.r.t attributes, shown in Table 5, give us some more interesting insights. For instance, it can be seen that subjects prefer to consume more Fizzy if it contains extra vitamins. They would also consume more Juice if it has no added colours or preservatives and consume more Cordial if it is diet or has no added colours or preservatives. These findings are all consistent with our expectation. As for cross-drink effect of attributes, in general, it is shown that *ceteris paribus*, healthier Juice will significantly crowd out the consumption of Fizzy. Fizzy containing no added colours or preservatives will crowd out Juice but with a relatively small effect. However, there seems to be no consistent crowd-out effect between Fizzy and Cordial or between Juice and Cordial.

**Table 3 Structural Parameter Estimates**

| Drink | Variable | Coef. | | S.E. |
|---|---|---|---|---|
| Fizzy | | | | |
| | Log of water consumption | 0.083 | | 0.415 |
| | Fizzy Diet | 0.055 | | 0.072 |
| | Fizzy Vitamins | 0.093 | * | 0.052 |
| | Fizzy Nocolours | 0.003 | | 0.048 |
| | Juice Diet | -0.161 | *** | 0.048 |
| | Juice Vitamins | -0.035 | | 0.081 |
| | Juice Nocolours | -0.221 | *** | 0.066 |
| | Cordial Diet | -0.054 | | 0.050 |
| | Cordial Vitamins | -0.051 | | 0.052 |
| | Cordial Nocolours | -0.028 | | 0.044 |
| | Log Fizzy Price | 0.081 | *** | 0.024 |
| | Log Juice Price | -0.096 | *** | 0.022 |
| | Log Cordial Price | 0.014 | | 0.016 |
| | Log Real Total Expenditure | 0.331 | *** | 0.062 |
| Juice | | | | |
| | Log of water consumption | 0.024 | | 0.332 |
| | Fizzy Diet | 0.151 | | 0.153 |
| | Fizzy Vitamins | -0.012 | | 0.032 |
| | Fizzy Nocolours | -0.152 | *** | 0.043 |
| | Juice Diet | -0.011 | | 0.036 |
| | Juice Vitamins | -0.008 | | 0.076 |
| | Juice Nocolours | 0.073 | ** | 0.035 |
| | Cordial Diet | 0.104 | | 0.141 |
| | Cordial Vitamins | 0.054 | | 0.068 |
| | Cordial Nocolours | -0.028 | | 0.033 |
| | Log Fizzy Price | -0.096 | *** | 0.022 |
| | Log Juice Price | 0.125 | *** | 0.029 |
| | Log Cordial Price | -0.029 | | 0.025 |
| | Log Real Total Expenditure | 0.299 | *** | 0.045 |
| Cordial | | | | |
| | Log of water consumption | -0.235 | | 0.256 |
| | Fizzy Diet | 0.092 | ** | 0.042 |
| | Fizzy Vitamins | -0.045 | | 0.066 |
| | Fizzy Nocolours | 0.091 | ** | 0.036 |
| | Juice Diet | 0.090 | ** | 0.035 |
| | Juice Vitamins | -0.019 | | 0.077 |
| | Juice Nocolours | 0.046 | | 0.028 |
| | Cordial Diet | 0.193 | ** | 0.084 |
| | Cordial Vitamins | -0.042 | | 0.040 |
| | Cordial Nocolours | 0.081 | * | 0.046 |
| | Log Fizzy Price | 0.014 | | 0.016 |
| | Log Juice Price | -0.029 | | 0.025 |
| | Log Cordial Price | 0.015 | | 0.035 |
| | Log Real Total Expenditure | -0.368 | *** | 0.039 |

*Note*: * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

**Table 4 Income Elasticities and Uncompensated and Compensated Price Elasticities**

| | Fizzy | Juice | Cordial |
|---|---|---|---|
| Income | 1.493*** | 1.119*** | 0.378*** |
| | (0.092) | (0.018) | (0.066) |
| Uncompensated | -0.879*** | -0.142*** | 0.021 |
| | (0.036) | (0.033) | (0.024) |
| | -0.038*** | -0.950*** | -0.012 |
| | (0.009) | (0.011) | (0.010) |
| | 0.024 | -0.050 | -0.974*** |
| | (0.027) | (0.042) | (0.058) |
| Compensated | -0.727*** | 0.976*** | 0.244*** |
| | (0.034) | (0.085) | (0.028) |
| | 0.075*** | -0.112*** | 0.156*** |
| | (0.009) | (0.017) | (0.010) |
| | 0.062** | 0.234*** | -0.918*** |
| | (0.026) | (0.062) | (0.061) |

*Note*: Elasticities are evaluated at the mean point. Standard errors are in parenthesis. * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

**Table 5 The Partial Elasticities w.r.t. Attributes**

| Drink | Attributes | Coef. | | S.E. |
|---|---|---|---|---|
| Fizzy | | | | |
| | Fizzy Diet | 0.082 | | 0.107 |
| | Fizzy Vitamins | 0.139 | * | 0.077 |
| | Fizzy Nocolours | 0.004 | | 0.071 |
| | Juice Diet | -0.240 | *** | 0.071 |
| | Juice Vitamins | -0.052 | | 0.121 |
| | Juice Nocolours | -0.329 | *** | 0.099 |
| | Cordial Diet | -0.080 | | 0.074 |
| | Cordial Vitamins | -0.076 | | 0.078 |
| | Cordial Nocolours | -0.041 | | 0.066 |
| Juice | | | | |
| | Fizzy Diet | 0.060 | | 0.061 |
| | Fizzy Vitamins | -0.005 | | 0.013 |
| | Fizzy Nocolours | -0.060 | *** | 0.017 |
| | Juice Diet | -0.004 | | 0.014 |
| | Juice Vitamins | -0.003 | | 0.030 |
| | Juice Nocolours | 0.029 | ** | 0.014 |
| | Cordial Diet | 0.041 | | 0.056 |
| | Cordial Vitamins | 0.022 | | 0.027 |
| | Cordial Nocolours | -0.011 | | 0.013 |
| Cordial | | | | |
| | Fizzy Diet | 0.156 | ** | 0.071 |
| | Fizzy Vitamins | -0.076 | | 0.112 |
| | Fizzy Nocolours | 0.155 | ** | 0.061 |
| | Juice Diet | 0.152 | ** | 0.060 |
| | Juice Vitamins | -0.032 | | 0.131 |
| | Juice Nocolours | 0.078 | | 0.048 |
| | Cordial Diet | 0.327 | ** | 0.142 |
| | Cordial Vitamins | -0.070 | | 0.068 |
| | Cordial Nocolours | 0.138 | * | 0.078 |

*Note*: * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

## V. Summary and Conclusion

In this study, using longitudinal data obtained from a soft drink consumption experiment, a fixed-effects censored demand system is specified and estimated. To deal with the difficulty of a substantial proportion of zero observations for the total expenditure on soft drinks and for the expenditure on each drink, a new two-step estimation strategy is developed and a semi-parametric estimator for two-sided censoring models with fixed effects is employed. In addition, a consistent and asymptotically efficient GMM estimator is used to impose economic restrictions on the model and identify the underlying structural parameters.

Based on our parameter estimates, the consumption behaviour of subjects is analysed through estimated income elasticities and uncompensated and compensated price elasticities. The partial elasticities of demand with respect to attributes of soft drinks are also estimated, to further understand subjects' choices. These results provide valuable empirical evidence for soft-drink tax policy concerns and health-related, such obesity and dental health, campaigns.

# Appendix

To see that $\hat{\delta}_n$ is a two-step M-estimator of $\delta_n$, it is helpful to rewrite (3.11) as:

$$\hat{\delta}_n = \arg\min_{\delta_n} \sum_{j=1}^{J} \sum_{1<s<t<T_j} \frac{1}{T_j} U\left( w_{njt}, w_{njs}, \left(\tilde{x}_{jt} - \tilde{x}_{js}\right)' \tilde{\delta}_n + \left(c_{jt} - c_{js}\right)\theta_n - \left(\widehat{x}_{jt} - \widehat{x}_{js}\right)' \hat{\tau}\theta_n \right),$$

where $\tilde{x}_{jt}$ denotes the vector of all the explanatory variables in (3.10) except $\hat{\xi}_{jt}$; $\tilde{\delta}_n$ denotes coefficient vector corresponding to $\tilde{x}_{jt}$; $\theta_n$ denotes the coefficient corresponding to $\hat{\xi}_{jt}$, so that $x_{jt} \equiv \left(\tilde{x}_{jt}', \hat{\xi}_{jt}\right)'$ and $\delta_n \equiv \left(\tilde{\delta}_n', \theta_n\right)'$; $\widehat{x}_{jt}$ denotes the vector of explanatory variables in (3.7) and $\hat{\tau}$ denotes corresponding coefficient estimates. Accordingly, the formal adjustment procedure for two-step estimation applies (see for example Wooldridge 2010c). For brevity, we present only the formulae for the asymptotic variance; derivations are available upon request.

Following Wooldridge (1995), let $\hat{\tau}$ be a $\sqrt{N}$-asymptotically normal estimator of $\tau$ with representation

$$\sqrt{J}\left(\hat{\tau} - \tau\right) = J^{-1/2} \sum_{j=1}^{J} r_j + o_p(1), \tag{A.1}$$

where $r_j$ is an i.i.d. sequence with $E\left(r_i\right) = 0$ and is estimated by minus the inverse of the average estimated Hessian (over the entire cross-section) times the estimated score of the Tobit log-likelihood function for observation $j$.

It can be shown that

$$\sqrt{J}\left(\hat{\delta}_n - \delta_n\right) \xrightarrow{d} N\left(0, A^{-1}DA^{-1}\right), \tag{A.2}$$

where $A$ equals $\Gamma$ in (3.13) and $D \equiv E\left[g_j g_j{}'\right]$; $g_j \equiv s_j + Fr_j$, where $r_j$ is from (A.1),

$$s_j \equiv \sum_{s<t} \frac{1}{T_j} u\left(w_{js}, w_{jt}, \left(x_{js} - x_{jt}\right)' \delta_n\right)\left(x_{js} - x_{jt}\right)$$

and

$$F = \frac{dE\left[\sum_{s<t} \frac{1}{T_j} u\left(w_{njt}, \mathrm{w}_{njs}, \left(\tilde{x}_{jt} - \tilde{x}_{js}\right)' \tilde{\delta}_n + \left(c_{jt} - c_{js}\right)\hat{\theta}_n - \left(\widehat{x}_{jt} - \widehat{x}_{js}\right)' \hat{\tau}\hat{\theta}_n\right)\left(x_{jt} - x_{js}\right)\right]}{d\tau}$$

We already know how to consistently estimate $A$: use expression (3.14). To estimate $D$, we need to estimate $F$. An estimator of $F$ can be shown as:

$$\hat{F} = \frac{1}{J}\sum_{j=1}^{J}\left[\sum_{s<t}\frac{1}{T_j}\left\{\begin{array}{l} 1\left\{-1 < \left(x_{js} - x_{jt}\right)' \hat{\delta}_n < 1\right\} \\ \left(\begin{array}{l} 1\left\{-1 < \left(x_{js} - x_{jt}\right)' \hat{\delta}_n < w_{js} - 1\right\} - \\ 1\left\{0 < \left(x_{js} - x_{jt}\right)' \hat{\delta}_n < w_{js}\right\} - \\ 1\left\{-w_{jt} < \left(x_{js} - x_{jt}\right)' \hat{\delta}_n < 0\right\} + \\ 1\left\{1 - w_{jt} < \left(x_{js} - x_{jt}\right)' \hat{\delta}_n < 1\right\} \end{array}\right)\left(x_{js} - x_{jt}\right)\left(-\left(\widehat{x}_{jt} - \widehat{x}_{js}\right)' \hat{\theta}_n\right) + \\ u\left(w_{js}, w_{jt}, \left(x_{js} - x_{jt}\right)' \delta_n\right)\left[\begin{array}{c} -\left(\widehat{x}_{jt} - \widehat{x}_{js}\right)' \\ \mathbb{Z} \end{array}\right] \end{array}\right\}\right]$$

where $\mathbb{Z}$ is a matrix of zeros, with the row dimension being equal to the length of $x_{js}$ and column dimension being equal to the length of $\widehat{x}_{jt}$.

## References

Alan, Sule, Bo E. Honoré, Luojia Hu, and Søren Leth-Petersen. 2014. "Estimation of Panel Data Regression Models with Two-Sided Censoring or Truncation." *Journal of Econometric Methods* 3: 1–20.

Alston, J. M., J. A. Chalfant, and N. E. Piggott. 2001. "Incorporating Demand Shifters in the Almost Ideal Demand System." *Economics Letters* 70 (1): 73–78.

Arndt, C. 1999. "Demand For Herbicide In Corn: An Entropy Approach Using Micro-Level Dataa." *Journal of Agricultural and Resource Economics* 24 (01): 204–21.

Arndt, C., S. Liu, and P.l Preckel. 1999. "On Dual Approaches to Demand Systems Estimation in the Presence of Binding Quantity Constraints." *Applied Economics* 31 (8): 999–1008.

Briggs, A. D. M., O. T. Mytton, R. Tiffin, M. Rayner, P. Scarborough, and A. Kehlbacher. 2013. "Overall and Income Specific Effect on Prevalence of Overweight and Obesity of 20% Sugar Sweetened Drink Tax in UK: Econometric and Comparative Risk Assessment Modelling Study." *BMJ* 347: f6189.

Brownell, K. D., and T. R. Frieden. 2009. "Ounces of Prevention — The Public Policy Case for Taxes on Sugared Beverages." *New England Journal of Medicine* 360 (18): 1805–8.

Buse, A. 1998. "Testing Homogeneity in the Linearized Almost Ideal Demand System." *American Journal of Agricultural Economics* 80 (1): 208–20.

Coast, Joanna, Hareth Al-Janabi, Eileen J Sutton, Susan A Horrocks, A Jane Vosper, Dawn R Swancutt, and Terry N Flynn. 2012. "Using Qualitative Methods for Attribute Development for Discrete Choice Experiments: Issues and Recommendations." *Health Economics* 21 (6): 730–41.

De Silva-Sanigorski, Andrea, Elizabeth Waters, Hanny Calache, Michael Smith, Lisa Gold, Mark Gussy, Anthony Scott, Kathleen Lacy, and Monica Virgo-Milton. 2011. "Splash!: A Prospective Birth Cohort Study of the Impact of Environmental, Social and Family-Level Influences on Child Oral Health and Obesity Related Risk Factors and Outcomes." *BMC Public Health* 11 (505): 1–9.

Deaton, A., and M. Irish. 1984. "Statistical Models for Zero Expenditures in Household Budgets." *Journal of Public Economics* 23: 59–80.

Deaton, A., and J. Muellbauer. 1980a. "An Almost Ideal Demand System." *American Economic Review*, Staff General Research Papers, 70 (3). JSTOR: 312–26.

———. 1980b. *Economics and Consumer Behavior*. Cambridge; New York: Cambridge University Press.

Dubois, L., A. Farmer, M. Girard, and K. Peterson. 2007. "Regular Sugar-Sweetened Beverage Consumption between Meals Increases Risk of Overweight among Preschool-

Aged Children{A Figure Is Presented}." *Journal of the American Dietetic Association* 107 (6): 924–34.

Edgerton, David L. 1997. "Weak Separability and the Estimation of Elasticities in Multistage Demand Systems." *American Journal of Agricultural Economics* 79 (1): 62–79.

Etilé, F., and A. Sharma. 2015. "Do High Consumers of Sugar-Sweetened Beverages Respond Differently to Price Changes? A Finite Mixture IV-Tobit Approach." *Health Economics*.

Finkelstein, E. A., C. Zhen, M. Bilger, J. Nonnemaker, A. M. Farooqui, and J. E. Todd. 2013. "Implications of a Sugar-Sweetened Beverage (SSB) Tax When Substitutions to Non-Beverage Items Are Considered." *Journal of Health Economics* 32 (1): 219–39.

Fry, J. M., T. R. Fry, and K. R. McLaren. 1996. "The Stochastic Specification of Demand Share Equations: Restricting Budget Shares to the Unit Simplex." *Journal of Econometrics* 73 (2): 377–85.

Heien, D., and C. R. Wessells. 1990. "Demand Systems Estimation with Microdata: A Censored Regression Approach." *Journal of Business & Economic Statistics* 8 (3). JSTOR: 365–71.

Hoare, Alexandria, Monica Virgo-Milton, Rachel Boak, Lisa Gold, Elizabeth Waters, Mark Gussy, Hanny Calache, Michael Smith, and Andrea M de Silva. 2014. "A Qualitative Study of the Factors That Influence Mothers When Choosing Drinks for Their Young Children." *BMC Research Notes* 7: 430.

Honoré, B. E. 1992. "Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects." *Econometrica* 60 (3): 533–65.

———. 2008. *On Marginal Effects in Semiparametric Censored Regression Models*.

Honoré, B. E., and J. L. Powell. 1994. "Pairwise Difference Estimators of Censored and Truncated Regression Models." *Journal of Econometrics* 64 (1-2): 241–78.

Lee, L. F., and M. M. Pitt. 1986. "Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica* 54 (5): 1237–42.

———. 1987. "Microeconometric Models of Rationing, Imperfect Markets, and Non-Negativity Constraints." *Journal of Econometrics* 36 (1-2): 89–110.

Leser, C. E. V. 1963. "Forms of Engel Functions." *Econometrica* 31 (4): 694–703.

Malik, V. S., M. B. Schulze, and F. B. Hu. 2006. "Intake of Sugar-Sweetened Beverages and Weight Gain: A Systematic Review." *American Journal of Clinical Nutrition* 84 (2): 274–88.

Meyerhoefer, C. D., C. K. Ranney, and D. E. Sahn. 2005. "Consistent Estimation of Censored Demand Systems Using Panel Data." *American Journal of Agricultural Economics* 87 (3): 660–72.

Moschini, G. 1995. “Units of Measurement and the Stone Index in Demand System Estimation.” *American Journal of Agricultural Economics* 77 (1): 63–68.

Perali, Federico, and Jean-Paul Chavas. 2000. “Estimation of Censored Demand Equations from Large Cross-Section Data.” *American Journal of Agricultural Economics* 82 (4): 1022–37.

Pollak, R. A., and T. J. Wales. 1992. *Demand System Specification and Estimation*. Oxford University Press.

Pollak, R.A., and T.J. Wales. 1981. “Demographic Variables in Demand Analysis.” *Econometrica* 49 (6): 1533–51.

Powell, J. L. 1986. “Symmetrically Trimmed Least Squares Estimation for Tobit Models.” *Econometrica* 54 (6): 1435–60. doi:10.2307/1914308.

Sharma, A., K. Hauck, B. Hollingsworth, and L. Siciliani. 2014. “The Effects of Taxing Sugar-Sweetened Beverages across Different Income Groups.” *Health Economics* 23 (9): 1159–84.

Shonkwiler, J.S., and S.T. Yen. 1999. “Two-Step Estimation of a Censored System of Equations.” *American Journal of Agricultural Economics* 81 (4): 972–82.

Wake, M, P Hardy, L Canterford, M Sawyer, and J B Carlin. 2006. “Overweight, Obesity and Girth of Australian Preschoolers: Prevalence and Socio-Economic Correlates.” *International Journal of Obesity* 31 (7): 1044–51.

Wales, T.J., and A.D. Woodland. 1983. “Estimation of Consumer Demand Systems with Binding Non-Negativity Constraints.” *Journal of Econometrics* 21: 263–85.

Wooldridge, J. M. 1995. “Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions.” *Journal of Econometrics* 68: 115–32.

———. 2010a. “Chapter 19 Censored Data, Sample Selection, and Atrition.” In *Econometric Analysis of Cross Section and Panel Data*, Second Ed, 775–851.

———. 2010b. “Chapter 14 Generalized Method of Moments and Minimum Distance Estimation.” In *Econometric Analysis of Cross Section and Panel Data*, Second Ed, 525–58.

———. 2010c. “Chapter 12 M-Estimation, Nonlinear Regression, and Quantile Regression.” In *Econometric Analysis of Cross Section and Panel Data*, Second Ed, 397–462.

Yen, S. T., B. H. Lin, and D. M. Smallwood. 2003. “Quasi-and Simulated-Likelihood Approaches to Censored Demand Systems: Food Consumption by Food Stamp Recipients in the United States.” *American Journal of Agricultural Economics* 85 (2): 458–78.

Yen, S.T. 2005. “A Multivariate Sample-Selection Model: Estimating Cigarette and Alcohol Demands with Zero Observations.” *American Journal of Agricultural Economics* 87 (May). JSTOR: 453–66.

Yen, S.T., K. Kan, and S.J. Su. 2002. “Household Demand for Fats and Oils: Two-Step Estimation of a Censored Demand System.” *Applied Economics* 34 (14). Taylor & Francis: 1799–1806.

Yen, S.T., and B.H. Lin. 2006. “A Sample Selection Approach to Censored Demand Systems.” *American Journal of Agricultural Economics* 88 (August): 742–49.

Zhen, C., E. A. Finkelstein, J. M. Nonnemaker, S. A. Karns, and J. E. Todd. 2014. “Predicting the Effects of Sugar-Sweetened Beverage Taxes on Food and Beverage Demand in a Large Demand System.” *American Journal of Agricultural Economics* 96 (1): 1–25.