WORKING PAPERS

N° 638

Revised version June 2021

The effects of taxation on the individual consumption of sugar-sweetened beverages

Céline Bonnet and Vincent Réquillart

Toulouse School of Economics

The effects of taxation on the individual consumption of sugar-sweetened beverages

Céline Bonnet*and Vincent Réquillart‡

June 9, 2021

Abstract

When assessing the impact of taxation on the consumption of sugar-sweetened beverages (SSBs), due to data constraint, most studies consider the average consumer. Individual consumption is, however, very heterogeneous. In this paper, we propose a three-step methodology to evaluate the impact of SSB taxation on individual consumption. First, we use a disaggregation method to recover individual consumption from observed household consumption. Second, we estimate the demand for different categories of households. Finally, we simulate the impact of a tax policy on individual consumption. We find a high level of heterogeneity in consumption. Adults, both men and women, consume a greater quantity of SSBs than children. More importantly, for any given age category, the average consumption of SSBs increases with body mass index (BMI). Among heavy consumers of SSBs, obese and overweight people are over-represented. In France, a €0.20/l tax on SSBs might decrease sugar intake by more than 1kg per year on average and by more than 2.5 kg, roughly 1.5 teaspoons/day, for 5% of the adult population. Moreover, overweight and obese men and women, who represent 41% of the adult population, represent 59% of the last five percentiles of the distribution of the variation in sugar intake. This is a key result because the objective of taxation is to decrease the consumption of individuals who are more at risk, that is those who are overweight and obese. Finally, we estimate that a €0.20/l tax on SSBs might avoid about 300 deaths (about 1% of the considered diseases) as a consequence of the decrease in SSB consumption.

JEL codes: D12, H31, Q18, I18

Key words: soft drinks, excise tax, differentiated products, individual consumption.

*Toulouse School of Economics, Université Toulouse Capitole (INRA), celine.bonnet@tse-eu.fr.

†Toulouse School of Economics, Université Toulouse Capitole (INRA), 21 Allée de Brienne, 31015 Toulouse Cedex 6, France. Corresponding author, vincent.requillart@tse-eu.fr

‡We thank participants to the XVth EAAE Congress and 11th IHEA congress for their helpful comments on an earlier version of this paper. We also thank Olivier de Mouzon for his assistance on the demand model. We acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010, and from INRAE metaprogramme DIDIT.

1 Introduction

Obesity and its adverse health consequences, such as hypertension and diabetes is a worldwide public health problem. In 2014, the age-standardized prevalence of obesity reached 10.8% in men, and 14.9% in women (NCD Risk Factor Collaboration, 2016). According to Finkelstein et al. (2005), the direct and indirect costs of obesity in the United States (US) might be as high as \$139 billion in 2003, that is 1.2% of the gross domestic product (GDP). In France, the social cost of overweight and obese individuals in 2012 was estimated to be as high as €20 billion, about 1% of GDP, an amount equivalent to the social cost of tobacco (Caby, 2016). The obesity epidemic is multi-causal but there is increasing evidence about the role of sugar-sweetened beverages (SSBs) consumption.[1] In line with the WHO, which recommends developing policies to limit the intake of products high in free sugars (WHO, 2015) more than 40 countries have implemented the taxation of SSBs by August 2020 (Global Food Research Program , 2020). The logic of such taxation is to limit the intake of sugar which is the main caloric ingredient of SSBs.

Ex-post evaluations of SSB taxation suggest that consumers do react to the tax. As examples, a tax led to a 12% decrease in the consumption of SSBs in Mexico (Colchero et al., 2015), a 10% decrease in Berkeley (California), (Silver et al., 2017) and a 21% decrease for low-income populations (Falbe et al., 2016). However, focusing on the average impact of a nutritional tax is not sufficient. First, there is evidence that consumption is highly heterogeneous (e.g. (Gustavsen and Rickertsen, 2011) for Norway, (Etilé and Sharma, 2015) for Australia). Second, given this heterogeneity in consumption, the impact of a tax is likely to differ greatly among the population. Using quantile regression, Gustavsen and Rickertsen (2013) showed that taxation "will have the highest percentage effect among low-purchasing households but the absolute effect is highest among high-purchasing households". Finkelstein et al. (2013) and Etilé and Sharma (2015) found similar results in the US and Australian markets, respectively. Dubois et al.

[1] Fifteen years ago, Malik et al. (2006), who conducted a review of studies on the link between SSBs and weight gain, concluded: "sufficient evidence exists for public health strategies to discourage consumption of sugary drinks as part of a healthy lifestyle". This statement is endorsed by the World Health Organization (WHO): "There is increasing concern that intake of free-sugars - particularly in the form of sugar-sweetened beverages - increases overall energy intake and may reduce the intake of foods containing more nutritionally adequate calories leading to an unhealthy diet, weight gain and increased risk of NCDs" (WHO, 2015).

(2020) analysed on-the-go purchases of beverages in the UK. Their results suggest that "soda taxes are relatively effective at targeting the sugar intake of the young, but are less successful at targeting the intake of those with high total dietary sugar". In an ex post evaluation of the impact of the French SSB tax, Capacci et al. (2019) found evidence of a larger response by the sub-sample of heavy purchasers.

The justification for SSB taxation relates to the negative health impact of 'excessive' sugar consumption which increases the risk of obesity and related negative health consequences. Estimates by Caby (2016) suggest that the social cost of obesity increases with the degree of obesity. Thus, in France, the annual social cost of an overweight individual is estimated at €360 and the annual social cost of an obese individual is estimated at €1,300. Finkelstein et al. (2010) also found that although people with a BMI greater than 35 represent only 37% of all obese people in the US, they account for 61% of the costs. Most of this social cost is an 'internality', that is, a cost that will be supported by the same individual in the future, and the literature suggests that consumers do not adequately take these internalities into account.[2] In a context of heterogeneous consumers and internalities, Griffith et al. (2017) show that the optimal tax rate should be defined as a function of "the average internality plus an adjustment based on the covariance of internalities and the (absolute value of) the slope of demands". In other words, as suggested by Allcott et al. (2014) what matters is the elasticity-weighted internality among consumers who adjust their consumption in response to a tax. The tax policy is more effective if the consumers who incur the highest internalities are the more price responsive.

The studies that deal with the heterogeneity of consumption conclude that, even if heavy consumers of SSBs are less price responsive than low consumers, the impact of a tax on consumption is the greater for heavy consumers. However these studies do not indicate who the heavy consumers are: in particular, whether or not they are obese, that is, if they are likely to suffer from 'high' internalities. Thus, one cannot conclude from these studies that the tax has a greater effect on individuals who are likely to experience higher internalities. Studies on the social cost of obesity suggest that the internality rises

[2] According to Caby (2016), for obese people, about 40% of the cost is paid by the taxpayer which corresponds to an externality and the 60% remaining to an internality, mostly due to the exclusion of women from the job market. However, the evaluation of the internality does not include utility losses due to a decrease in life expectancy and the decrease in quality of life.

with the severity of obesity. However, this does not mean that the internality rises with the level of SSB consumption as obesity rates are not determined only by SSB consumption which represents only a small portion of calorie consumption. Then, to discuss the merits of a SSB tax policy, it is important to have better estimates of the consumption of individuals according to their obesity status. The tax will be more justified if obese consumers limit their consumption to a greater extent than non-obese consumers.

The limits of the existing studies relate to the data availability. Because data on the at-home consumption of food is mostly available at the household level, individual purchases are frequently estimated by dividing household purchases by the size of the household.[3] However, SSB consumption by individuals in a household is likely to be heterogeneous and related to the characteristics of the individuals. Because the consequences of excessive SSB consumption on health are an individual issue, it is important to improve the way in which individual consumption is estimated when the information currently available relates to households and not to individuals.[4] The objective of this paper is to fill the gap and to assess the impact of SSB taxation on individual consumption using household purchase data. To do so, we develop a three-stage methodology. First, we use a disaggregation method to recover individual consumption from the observed household consumption of non-alcoholic beverages, distinguishing individuals according to characteristics, such as age, gender, and body mass index (BMI). Second, we model the household demand using a random utility approach and estimate the price elasticities of demand according to households' characteristics. Third, using the estimated price elasticities and the estimated individual consumption of the different SSBs, we simulate the impact of a tax policy on individual consumption. Stages two and three are based on standard procedures, while conversely, the first stage of the approach is original and is based on the few papers using this method in the empirical literature. Chesher (1998) used a non-parametric method to decompose the nutrient consumption of households into individual consumption. Vasdekis and Trichopoulou (2000) and Allais and Tressou (2009) adopted

[3]This is the way in which Finkelstein et al. (2013) and Etilé and Sharma (2015) estimate individual consumption from household data. Dubois et al. (2020) have individual data covering out-of-home consumption.

[4]Wada et al. (2015) used an alternative approach: they merged the 24-hour dietary recall data from the US National Health and Nutrition Examination Surveys with data on soda prices. The data set directly provides individual data, thus solving the problem. This has a cost, however. First, consumption is observed over a very short period of time, which may impact the robustness of the results. Second, prices faced by each consumer are not observed, making it difficult to accurately estimate the demand.

an additive non-parametric approach to recover individual food expenditure, and individual seafood consumption, respectively. Recently, Bonnet et al. (2014) estimated individual nutrient consumption as a function of individual BMI and other individual variables. In this paper, we follow a similar method in order to decompose the household purchase of non-alcoholic beverages into individual purchases. We assume that in a given household, consumption by an individual depends on his own characteristics, and on household characteristics.

According to our results, at-home consumption of SBBs by adults, both men and women, is greater than that of children. In most cases, the average consumption of a beverage for a given age category increases with BMI status. We also find that at a given age and BMI status, the consumption of regular soft drinks in the last decile of the distribution is at least twice the average consumption, while the consumption in the last centile is at least four times the average consumption. From the consumption of the different beverages, we deduce the sugar intake, which, as explained above, is the targeted nutrient of a SSB tax policy. On average, sugar intake from beverages is about 5 kg per year for adults and slightly less than 2 kg per year for children. However, there is a large heterogeneity of consumption. For example, consumption is higher than 11 kg per year for 5% of adults (higher than 18 kg per year for 1% of adults).[5] To give an order of magnitude, 11 kg of sugar per year (from SSBs) is equivalent to 30 g/day, which is about seven teaspoons. In comparison, the WHO recommends limiting the calorie intake of free-sugars (from all food products) to 10% of the energy intake; that is about 50 g/day (or 12 teaspoons) for a standard diet. This means that at-home consumption of this single food product category already provides 60% of the maximum recommended sugar intake. An important result is that those who are overweight and obese are over-represented in the last percentiles of consumption; that is, the proportion of overweight and obese individuals tends to increase with the percentile class. This is particularly the case for adults. For example, obese men and women represent 4.6% and 5.5% of consumers, respectively, in the 50th percentile, but 8.4% and 10.5%, respectively, in the last five percentiles. We find price elasticities of demand in the range of -1.4 to -2.7 for the different beverages. Finally, we show that a €0.20/l tax on

[5]For children, it is higher than 4 kg per year for 5% of them and higher than 7 kg for 1%.

SSBs might decrease the sugar consumption of adults by about 1 kg/year (about 20%) and that of 5% of the adult population by more than 2.5 kg (roughly by 1.5 teaspoons/day as compared with the maximum recommended amount of 12 teaspoons/day). More importantly, the decrease in consumption is higher for overweight and obese individuals than for normal weight adults. This is an important result, which means that individuals who are likely to have larger (negative) internalities will be more impacted by the tax. This is exactly what a tax should do in order to be as effective as possible. Our results thus provide additional support to SSB taxation because the tax is likely to affect more people who are at risk.

The paper is structured as follows. Section 2 presents the data used in this study. Section 3 discusses the different stages of the methodology. Section 4 provides the results of the estimated individual consumptions, price elasticities, and the impact of a tax policy on individual consumption. Section 5 concludes.

2 Data on SSB purchases of households

We use the 2011 data from a French representative consumer panel data of 27,291 households collected by KANTAR Worldpanel, a home-scan data set providing detailed information on at-home purchases of non-alcoholic beverages. For each purchase, the data set provides characteristics of the good such as the brand or the type of sweetener used, the quantity purchased and the expenditure. The data set also provides information on households, such as the socio-economic status, as well as the age, gender and BMI of each person within the household. Because beverages are mainly non-perishable products that can be stored for several months (except for some varieties of fruit juice), we assume there is no loss, implying that consumption and purchases are equal.[6] Our data set provides at-home consumption and thus does not cover the whole consumption of non-alcoholic beverages. According to INCA surveys, at-home consumption of non-alcoholic beverages represented 76.9% of the whole consumption in 2006 and 70.6% in 2014-15.[7]

The market of non-alcoholic beverages includes soft drinks (SDs), fruit juice, nectar, and bottled wa-

[6] Because we analyse yearly consumption, changes in inventory play a minor role.
[7] INCA stands for "Etude Individuelle Nationale des Consommations Alimentaires".

ter.[8] SDs include colas, iced tea, flavoured water, and an aggregate of other soft drinks (tonic, lemonade, sport drinks, energy drinks, and fruit drinks). We exclude drinking milk from the non-alcoholic market, as econometric analysis of the French market suggests that substitutions between milk and non-alcoholic beverages are small (Allais et al., 2010). As a consequence, the market is composed of six groups of products (the four SDs, nectar, and fruit juice), plus bottled water. For SDs, we distinguish the regular version (which contains added sugar) from the diet version (which does not contain added sugar). In the following, we will refer to seven product categories: regular colas, regular iced tea, regular flavoured water, other regular SDs, diet SDs, nectar, and fruit juice.[9] Finally, we define SSBs as regular SDs and nectar as they contain added sugar.

Table 1 provides descriptive statistics on consumption for the whole sample and for different types of households. It provides the proportion of households that do not consume a given product category and the average per-capita consumption on the basis of households that consume rather than on the total population.[10] We distinguish households according to the presence of children and to the BMI status of the main shopper.

Most households consume some SSBs. Thus, only 10% do not consume any SSBs, and 13% do not consume any regular SDs. Obviously, the percentage of non-purchasers for a given sub-category is higher. Fruit juice is also consumed by most households. The proportion of non-purchasers is higher for households without children than for households with children. This is particularly the case for regular SDs or for regular colas. This might be due to a 'size' effect. Thus, assuming that any consumer has her own preferences for the type of beverages to drink, an increase in the size of the household is likely to increase the number of different products purchased by the household. This might also be due to the specific preferences of children compared to adults. We also observe that the proportion of non-purchasers of a particular product tends to decrease when the BMI status of the shopper 'increases',

[8] The fruit juice category aggregates pure fruit juice (60% of purchases), and juice prepared with fruit purée (40%). We assume that fruit juice does not contain added sugar.

[9] In the descriptive analysis, we distinguish the four diet SDs. However, due to the number of observations, to estimate individual consumption, we consider an aggregate of diet SDs.

[10] The per-capita consumption is the household consumption divided by the number of persons in the household.

however, this effect remains much smaller than the previous one. The average consumption (for those who consume) of regular SDs is 35 litres per person per year, which is relatively low compared to the reported level of consumption in some other countries.[11] The average consumption of diet SDs is 27 litres per person per year, greater than the average consumption of fruit juice and nectar (21 litres per person per year).

Because we are interested in the intake of sugar due to beverage consumption, we compute the added sugar consumption and the total sugar consumption using the average sugar content of the different categories of beverages. For those who consume SSBs, the average intake of added sugar is 3.5 kg per person per year. For those consuming SSBs or fruit juice, the average intake of sugar (whether the sugar is added or not) from beverages is estimated at 4.8 kg per person per year.[12]

Table 2 provides information on prices. Prices are computed using observed prices (unit-value) of 60 products sold under 40 different brands purchased from six different retailers.[13] Prices differ between categories, with fruit juice, nectar, and other SDs (both regular and diet) being the most expensive products. There is no clear ranking between the prices of the diet versions of a product compared to the regular ones. For example, on average, diet cola is less expensive than regular cola but diet iced tea is more expensive than regular iced tea. As shown by the standard deviations, within a category, there is some heterogeneity in price, which is mainly due to the heterogeneity of price across brands in a given category. Finally, households with children tend to pay lower prices than households without children.

3 Method

To assess the impact of SSB taxation on individual consumption, we develop a three-stage methodology. First, we use a disaggregation method that allows us to recover individual consumption from individual characteristics and household consumption. Second, we model and estimate household demand for differentiated products in the non-alcoholic beverage market and deduce own- and cross-price elasticities

[11]For example, Etilé and Sharma (2015) report a 90 litres per year at-home consumption of SDs in Australia.

[12]In comparison, Han and Powell (2013) report an annual intake of sugar from SSBs greater than 20 kg per year in the US.

[13]To build Table 2, for each product category, we compute a monthly average price as the weighted average, using market shares as weights, of the price of all purchases from all retailers of all products in each category.

of demand. Third, using the estimated price elasticities and estimated individual consumption of the different beverages, we simulate the impact of a tax policy on the individual consumption of beverages and sugar.

3.1 Disaggregation method

In the following, we describe the method used to estimate individual consumption from data on household consumption. We use at-home purchase data described above. For each household, we compute her annual purchase for seven product categories (four regular SDs, an aggregate diet SD, nectar, and fruit juice). For each product category, the annual consumption is equal to the sum over the year of every purchases in the six considered retailers. We apply the method of disaggregation for each product category independently. We first present conditions under which individual consumption can be identified and estimated, and then explain and justify the specification used for each category.

3.1.1 *Identification*

Let us assume that for a person p in the household i the individual consumption y_{ip}^b of beverage b is defined by:

$$y_{ip}^b = \beta^b(x_{ip}) + u_{ip}^b, \tag{1}$$

where x_{ip} is a vector of individual characteristics of the person p in household i, $\beta^b(.)$ is a semi-parametric function, and u_{ip}^b is a deviation for this person's consumption. Then, the household consumption y_i^b of beverage b is given by:

$$y_i^b = \sum_{p=1}^{P(i)} y_{ip}^b = \sum_{p=1}^{P(i)} \beta^b(x_{ip}) + \varepsilon_i^b, \tag{2}$$

where $\varepsilon_i^b = \sum_{p=1}^{P(i)} u_{ip}^b$ and $P(i)$ is the number of persons in the household i.
Assuming that $\forall p$, i, t:

$$E\left(u_{ip}^b | x_{i1}, .., x_{iP(i)}\right) = 0, \tag{3}$$

implies that:

$$E\left(\varepsilon_i^b | x_{i1}, .., x_{iP(i)}\right) = 0,$$

allowing us to identify β^b consistently.

Assumption (3) implies that $\beta^b(x_{ip})$ can be interpreted as the average consumption of beverage b by an individual with characteristics x_{ip}. Finally, u_{ip}^b is interpreted as the deviation from the mean consumption of this individual. As shown in Table 1, household consumption in any beverage b is highly heterogeneous. As $\beta^b(x_{ip})$ represents the average individual consumption of beverage b across the households in our sample, we lose a part of the heterogeneity of consumption. To deal with this heterogeneity in household consumption, we compute for each person in a household his estimated share of consumption (using the estimated consumption of the household defined as the sum of the estimated individual consumption of all persons in the household). Formally, we have $\widehat{y}_{ip}^b = \widehat{\beta}^b(x_{ip})$ the estimated individual consumption and $\widehat{y}_i^b = \sum_{p=1}^{P(i)} \widehat{y}_{ip}^b$ the estimated household consumption. $\frac{\widehat{y}_{ip}^b}{\widehat{y}_i^b}$ is the estimated share of consumption of person p in household i. We define the individual consumption as $\widetilde{y}_{ip}^b = \frac{\widehat{y}_{ip}^b}{\widehat{y}_i^b} y_i^b$.

3.1.2 Specification

We consider that the consumption of an individual is affected by individual characteristics (age, gender and BMI) and by characteristics of the household. The household characteristics (Table 11,Appendix 6.2) are selected using a statistical approach (see details in the Appendix 6.2).

Because disaggregation models differ slightly across beverage categories, we detail one specific case (regular cola) and indicate how this specification is modified in the other cases. We consider that the consumption of an individual is affected by his own individual characteristics (gender, age, and BMI status) and by selected household characteristics. We consider a semi-parametric function to represent individual consumption. We assume that the age and gender of the individual and the household characteristics allow us to discretize the household consumption. In addition, we introduce a multiplicative specification for the individual BMI, as in Bonnet et al. (2014). A change in BMI proportionally affects individual consumption. We consider three age categories ($x_{ip}^1 = a$ with $a \in [\leq 10; 11-17; \geq 18]$), gender ($x_{ip}^2 = g$ with $g \in [male; female]$), three income levels ($x_{ip}^3 = s$ with $s \in [poor; intermediate; rich]$), six regions ($x_{ip}^4 = r$ with $r \in [Paris; East; North; West; Centre; South]$), and two types of area ($x_{ip}^5 = l$ with $l \in [Urban; Rural]$). Thus, in the case of the regular cola category, we estimate individual consumption using the following specification of the function β:

$$\beta\left(x_{ip}^1, x_{ip}^2, x_{ip}^3, x_{ip}^4, x_{ip}^5, z_{ip}\right) = \sum_{a=1}^{3}\sum_{g=1}^{2}\sum_{s=1}^{3}\sum_{r=1}^{6}\sum_{l=1}^{2} 1_{\left\{x_{ip}^1=a, x_{ip}^2=g, x_{ip}^3=s, x_{ip}^4=r, x_{ip}^5=l\right\}} \beta_{ag}^{srl}\left[\delta_0^g + \delta^g\left(x_{ip}^1\right)\left(\frac{z_{ip} - \overline{z}_{a,g}}{\sigma_{a,g}}\right)\right], \quad (4)$$

with $\delta^g\left(x_{ip}^1\right) = 1_{\left\{x_{ip}^1 \leq 10\right\}}\delta_1^g + 1_{\left\{11 < x_{ip}^1 \leq 17\right\}}\delta_2^g + 1_{\left\{x_{ip}^1 \geq 18\right\}}\delta_3^g$. z_{ip}, $\overline{z}_{a,g}$, and $\sigma_{a,g}$ are, respectively, the BMI of person p in household i at the beginning of the year, the mean, and the standard deviation of the BMI for individuals of age a and gender g. With this specification, the continuous part of the function β in the BMI is intended to be an age and gender specific linear function of the standardized BMI by gender and age.[14] This specification applies to regular colas, fruit juice and other SDs. For nectar, the variable region is replaced by the variable SPC. For iced tea, the variable type of area is excluded. For flavoured water, the income variable is replaced by the education variable and the type of area is omitted. Finally, for diet products, the education variable replaces the income variable, as summarized in Table 14.

3.2 The demand model: a random coefficient logit model

To model the purchasing behaviour in the soft drink market, we opt for a random coefficient logit model that allows getting flexible consumer substitution patterns (Berry et al., 1995; McFadden and Train, 2000). We consider that households face 60 differentiated products, conditional on which retailer they visit.[15] The 60 differentiated products are the main national brands and an aggregate of private label products for each soft drink categories. We define the average price of the 60 differentiated products in each of the five main retailers and an aggregate of the other supermarkets and hypermarkets. This allows variability in product prices and a clear identification of the households' price sensitivity. To get flexible

[14] As we consider the BMI at the beginning of the year (that is, before purchases occur), we do not encounter a reverse causality problem between BMI and beverage consumption.

[15] Some products could be unavailable in some retailers at some time periods.

substitution patterns of households in the soft drink market, we use a random coefficient logit model and assume that the indirect utility function V_{ijt} for household i buying product j in month t is given by:

$$V_{ijt} = \alpha_i p_{jr(i)t} + \mu_{b(j)} + \sum_{c=1}^{C} \gamma_{c(j)} + \varepsilon_{ijt},$$

where $\mu_{b(j)}$ is a brand fixed effect that captures the (time-invariant) unobserved brand characteristics, $p_{jr(i)t}$ is the price of product j in month t in the retailer r that the household i visits, α_i is the marginal disutility of the price for household i, $\gamma_{c(j)}$ represents the mean taste of the category $c(j)$ over the population, and ε_{ijt} is an unobserved individual error term. As households can each have a different price disutility, we take into account unobserved heterogeneity allowing for a random price coefficient: $\alpha_i = \alpha + \sigma \nu_i$, where α is the mean price disutility, σ measures the deviation to the mean disutility, and ν_i is independently distributed as standard normal and captures the unobserved households characteristics.

Rather than consuming one of the considered products, the household can decide to consume an alternative good, named an outside option, thus allowing substitution between the considered products and the alternative. In this study, the outside good is non-flavoured bottled water. The utility a household gets when consuming the outside good is normalised to zero. The indirect utility of choosing the outside good is $V_{i0t} = \varepsilon_{i0t}$.

Own- and cross-price elasticities can be deduced from the demand model estimates. Identification issues and the expression for price elasticities are provided in Appendix 6.4.

3.3 Simulation of the impact of the tax

To evaluate the impact of the tax on individual consumption of beverages and sugar, we proceed as follows:

- We compute the new retail prices as the sum of the initial retail prices and the tax. Thus, retail prices of products are assumed to vary by the amount of the tax. This is a simplification, as some previous work suggests that the pass-through rate of the tax may differ from 100%;[16]

[16] An ex-post study on the pass-through rate of an excise tax on non-alcoholic beverages in France estimated an average full shifting of the tax for sodas, a 94% pass-through rate for fruit drinks and a 62% pass-through rate for flavoured water (Berardi et al., 2016).

- For each household, we compute the change in consumption of each of the 60 products using the initial consumption of the household and the relevant (we have six categories of households) matrix of own- and cross-price elasticities.
- For each household, we then compute the percent change in the household's consumption for each of the seven product categories;
- For any individual, we compute the change in her consumption of the seven product categories by using her initial consumption and the percent change of consumption of the household she belongs;
- From the change in individual consumption, we deduce the change in sugar intake by using the average sugar content of each product category.

4 Results

We first describe the estimated individual consumption of beverages by analysing the distribution of the consumption of SSBs, diet SDs, and fruit juice. We also discuss the distribution of the individual sugar intake from the associated individual consumption of all beverages. Then, we provide the results on the price elasticities and discuss the impact of taxing SSBs on individual consumption.

4.1 Individual consumption of beverages

Figure 1(a-c) presents the distribution of the estimated individual consumption of SSBs, diet SDs, and fruit juice for children, adult women and men who actually consume a positive quantity of the product.[17] There are common features in the consumption of the three groups of beverages. First, adults consume much more than children.[18] Second, women tend to consume slightly more than men, but the difference is small. Third, consumption tends to increase with BMI. The only exception is the consumption of fruit juice by children, which seems unaffected by BMI. The tests of equality of mean consumption (Table

[17] We present aggregate results for children as there are no significant differences in consumption among the different classes of children considered in this study (gender distinction, age < 10 or between the ages of 11 and 17).

[18] Because we compare the distribution of consumption, it means that the consumption of a given decile of adults is larger than the consumption of the same decile of children.

15 in the Annex) and the Kolmogorov-Smirnov tests of differences in the distribution of consumption (Table 16 in the Annex) support the conclusions about the link between the distribution of consumption and BMI. In addition, they also show that the mean consumption of fruit juice does not differ for men and women while it does differ for SSBs and diet SDs. Results on distributions also show that there is a large heterogeneity in the distribution of consumption.

Table 3 provides some details about the distribution of consumption for children, women, and men according to their BMI status. It shows that even if the mean or median individual consumption of SSBs is low in France, consumers in the last decile consume at least twice the average consumption and consumers in the last centile consume at least four times the average consumption. For example, 5% of normal weight children consume more than 40 litres per year, 5% of obese men consume more than 128 litres per year and 5% of obese women consume more than 139 litres per year. Results clearly indicate that consumption increases with BMI, and this is true for every decile in the distribution.

The previous results are based on estimates of individual consumption. However, for households composed of a single adult we do observe the individual consumption. We present in the Appendix 6.5 a comparison of the estimated distribution of individual consumption of SSBs with the observed distribution of SSB consumption of single adult households.

4.2 Individual consumption of sugar

Figure 1(d) provides the distribution of the estimated individual consumption of sugar due to beverage consumption. Men and women consume more sugar than children, and men consume slightly less than women. Moreover, the Figure clearly shows that sugar intake increases with BMI status. This is a consequence of the higher consumption of beverages by overweight or obese individuals. Tests of means equality reveals that for children, men, and women the average sugar intake of overweight individuals is significantly greater than that of normal weight individuals and significantly lower than that of obese individuals (Table 15 in the Annex). Moreover, Dunn's test of stochastic dominance shows that for children and men, the distribution of the consumption of obese individuals statistically dominates that of overweight individuals and the distribution of the consumption of overweight individuals statistically dominates that of normal weight individuals (Table 16 in the Annex). For women, the respective distribution of the consumption of overweight and obese individuals dominates the distribution of the consumption of normal weight women.

The total quantity of sugar intake is about 4.7 kg per year on average for normal weight adults and is greater than about 9 kg per year for the last decile (Table 4). In the latter case, this intake corresponds to the calorie intake required for 18 days, that is, about 5% of the total calorie intake for a year.[19] By way of comparison, the WHO recommends limiting the consumption of sugar to a maximum of 10% of the calorie intake (WHO, 2015). For obese adults, an individual in the last decile consumes more than 11 kg sugar per year and an individual in the last centile consumes more than 18.5 kg sugar per year. For children, the intake is lower, from 1.6 to 2.1 kg/year on average depending on the class of BMI. As for adults, consumption of the last decile is much larger.

Table 5 provides information about the characteristics of consumers as a function of their sugar intake. Because their respective levels of consumption differ strongly, we provide separate results for children and adults. In Table 5, the column 'All' provides the proportion of the different types of consumers in our sample. For example, among children, 85.3% are of normal weight and among adults, 24.9% are normal weight men.[20] The other columns provide the proportion of consumers for different percentiles of consumption.

An important result is that overweight and obese individuals are over-represented in the last percentiles; that is, the proportion of those who are overweight and obese tends to increase with the percentile class. This is particularly the case for adults. For example, obese men and women represent 4.6% and 5.5% of consumers in the 50th percentile respectively, but 8.4% and 10.5% respectively in the last five percentiles ($>$ 95th percentile). In the case of children it is mainly in the last decile that obese and overweight individuals are over-represented. Whereas the share of overweight and obese children in the

[19] Based on 4 kcal/g sugar and a consumption of 2,000 kcal per day, which is a level frequently used as a guideline for the daily energy intake for adults.

[20] As compared to the whole population, we have a slightly lower rate of overweight and obese people. According to Institut Roche de l'Obésité (2012) in France 14.3% of men and 15.7% of women were obese. In our sample, we have 11.6% and 12.4% of obese men and women, respectively. For overweight individuals, we have a similar difference. The share of overweight men and women in France was 38.8% and 26.3% respectively, whereas in our sample it is 35.5% and 23.3%, respectively. Finally, note that in our sample women are slightly over-represented as they account for 53% of adults.

population is 10.6% and 4.1%, respectively, the corresponding share in the last five centiles is 12.0% and 6.6%, respectively. Finally, whatever the age and gender category, obese consumers are over-represented in the last quartile and under-represented in the first three quartiles of sugar intake.

4.3 Elasticities

Using household purchase data, we estimate the demand model, controlling for endogeneity (see Tables 17 and 18, as well as a technical note on the price endogeneity issue in the Appendix). In order to introduce some heterogeneity in consumer preferences for brands and categories, and in how households react to price changes, we estimate a similar demand model for six groups of households. We distinguish households by the obesity status of the panelist and by the presence of children. Demand results are provided in Appendix 6.4.

Tables 6 and 7 provide the aggregated elasticities at the product category level for each of the 6 households groups. Overall, own-price elasticities for product categories range from -1.4 to -2.7. There are some regularities among the different matrices. In most cases, the demand for nectar is the most elastic and the demand for diet products is the least elastic. Demand for cola is also less elastic than most of the other products. With respect to the presence of children or the obesity status of the panelist, there are differences in the elasticities. Among households with children, those with an overweight or obese status of the panelist have the less elastic demands (own-price elasticities range between -1.4 and -1.9). Conversely, among households without children, those with a normal obesity status have the less elastic demand. Our estimates of elasticities are consistent with the findings of Dharmasena and Capps (2012) and Zhen et al. (2014). A recent meta-analysis by Green et al. (2013) reports an average own-price elasticity for SSBs of -1.3. In our case, we get aggregate own-price elasticity for SSBs as a whole ranging from -1.09 to -1.96 for the different group of households.[21]

Tables 6 and 7 also provides aggregated cross-price elasticities. For a given category, the cross-price elasticities with the other categories of beverages are similar. The cross-price elasticity with the outside

[21] Own-price elasticities of SSBs are -1.45, -1.96 and -1.70 for normal weight, overweight and obese households without children, respectively and -1.26, -1.15 and -1.09, respectively, for households with children.

good is somewhat different. To interpret this, consider a category, assume that its price increases, then the demand decreases, which is compensated for by an increase in the demand for all other products including the outside good. Because the values of the cross-price elasticities are similar, this means that when consumers choose an alternative product (except the outside good), they choose the alternative in proportion to its initial market share.

4.4 Impact of taxation on consumption

We evaluate the impact of taxing SSBs on individual consumption. We consider an excise tax of €0.20/litre which on average represents a 20% increase in the price of regular SDs and 15% in the price of fruit juice and nectar. However because we consider the different products in each category, the percent change in the price of different products in a given category is not identical. For example, the percent change in the price of national brand products is lower than the percent change in the price of private label products, as the latter are offered at a lower price to consumers. In this scenario, SSBs, as well as fruit juice, are taxed. In other words, beverages containing sugar are taxed, whatever the origin of the sugar. [22]

The average decrease in annual sugar consumption due to taxation is estimated to be about 290 g for normal weight children, 380 g for obese children, 940 and 1260 g for normal weight and obese women, respectively, and 940 and 1225 g for normal weight and obese men, respectively (Table 8). We also find that for all three groups of consumers, the average reduction in sugar consumption increases significantly with the BMI status (results of the tests in Table 15 in the Appendix). In addition, Dunn's test of stochastic dominance shows that for children and men, the distribution of the change in sugar consumption of obese individuals statistically dominates that of overweight individuals and the distribution of the change in sugar consumption of overweight individuals statistically dominates that of normal weight individuals (Table 16 in the Appendix). For women, the distribution of the change in sugar consumption

[22] There is a debate about including fruit juice in a taxation scheme as fruit juice does not contain added sugar and might be considered as healthy option. However, in this analysis, we include fruit juice in the taxation scheme in order to limit substitutions between sugary products, whatever the origin of sugar. An alternative method of taxation is to design an excise tax on sugary beverages which is a function of the sugar content of the product. In our context, the tax on fruit juice, nectar, and regular colas would be higher than that on iced tea, other regular SDs and flavoured water.

of obese individuals dominates the respective distributions of consumption of normal weight women and overweight women.

Finally, we find that the decrease might be much greater for a fraction of the population. Thus, for 10% of the adult population, the decrease is greater than about 2 kg, equivalent to more than one teaspoon per day or 1% of the energy intake.

Table 9 provides information about the characteristics of consumers as a function of the decrease in sugar intake due to taxation. In Table 9, the column 'All' provides the proportion of different types of consumers in our sample. The other columns provide the proportion of consumers for different percentiles of change in sugar intake.

The key result is that overweight and obese individuals are over-represented in the last percentiles; that is, the proportion of those who are overweight and obese tends to increase with the percentile class. This is particularly the case for adults. For example, obese men and women, who represent 5.4% and 6.6% of the adult population, represent 7.3% and 9.5% of consumers in the 75 - 90th percentile, respectively, and 8.4% and 10.0%, respectively in the last five percentiles (> 95th percentile). Overweight and obese men and women, who represent 41% of the adult population, represent 59% of the last five percentiles, that is, adults who reduce their sugar intake by more than 2.5 kg/year. It means that individuals with a high obesity status, and therefore are likely to have larger internalities, are more impacted by the tax than people with a 'normal' BMI.

In the case of children, we get similar results but to a lower extent. It is mainly in the last decile (last two columns) that obese and overweight individuals are over-represented. Whereas the share of overweight and obese children in the population is 10.6% and 4.1%, the corresponding share in the last five centiles is 11.8% and 7.1%, respectively.

4.5 A comparison with the standard way to estimate per capita consumption

We compare the impact of the tax on sugar consumption (from beverages) when using the method we propose to estimate individual consumption and the classical method consisting of estimating per-capita consumption by dividing household consumption by the number of individuals in the household. We apply the general method described in Section 3.3. to estimate the impact of the tax. Differences in the results then come from the method used to estimate per-capita consumption.

Table 10 provides the difference in the estimated impact of the tax when using the two alternative methods to estimate per-capita consumption. On average, compared to the standard method, our method provides a lower response to the tax for children's consumption and a higher response to the tax for adults.[23] Thus, in the standard method, in a given household, the consumption of children and adults are identical. On the contrary, with our disaggregation method, we do not impose equal consumption among individuals in the household; as a result, we get children's consumption lower than adult consumption. Interestingly, for adults, the difference in estimates is larger for overweight and obese adults than for normal weight individuals. This is because our method provides estimates of consumption that increase with the BMI status of the individual.[24]

5 Discussion

This paper contributes to the literature on SSB taxation by addressing the issue of heterogeneous consumption and the heterogeneous impact of taxation policies. The main originality of our approach is the way in which we deal with this heterogeneity. While most papers estimate individual consumption by dividing observed household consumption by the size of the household, in this paper we recover individual consumption by taking into account the many characteristics of both individuals and households. In particular, we distinguish individuals by their gender, age category and BMI status. Then using standard methods we estimate demand elasticities, and analyse the impact of taxing SSBs on individual consumption.

Thanks to the disaggregation method, we document the heterogeneity of consumption among consumers. As is the case in quantile-based studies (e.g., Gustavsen and Rickertsen, 2013; Etilé and Sharma, 2015), we find a high level of heterogeneity of consumption. Thus, even if on average the SSB consump-

[23]Note that the average difference computed over all individuals is lower than 1g/year, meaning that the differences relates to the heterogeneity of consumption by type of individuals and not from a systematic bias.

[24]It also explains why the difference in the estimates of the impact of the tax decreases with the BMI status for children.

tion in France remains low (45 l/year) compared to the average consumption in some other countries, some consumers experience a very high level of consumption. However, a more original result is that both for adults and children, SSB consumption increases with BMI status. In particular, we show that whatever the age and gender category, obese consumers are over-represented in the last quartile of sugar intake (due to beverage consumption) and under-represented in the first three quartiles of sugar intake. This is an important result in the debate about the impact of SSB taxation on consumers. Thus, as shown in Table 8, obese and overweight individuals are more impacted by the tax than those of normal weight.

The average decrease in sugar intake due to taxation is estimated to be about 940 g for normal weight men and women. The average decrease is 1.22 kg for overweight women and 1.12 kg for overweight men. We also find that the decrease may be much greater for a fraction of the population. Thus, for 10% of the adult population, the decrease is larger than 1.95 kg which corresponds to a 1% decrease in energy intake. The heterogeneity of response according to the BMI status is a key point in the analysis of the impact of a tax. Thus, a tax policy would be more effective if consumers who incur the largest internalities (obese individuals) are the most impacted by the tax (Griffith et al., 2017). Our results suggest that it is partly the case. Overweight and obese men and women, who represent 41% of the adult population, represent 59% of the last five percentiles of reduction in sugar intake; that is, adults who reduce their sugar intake by more than 2.5 kg/year. This is a new result, as the literature analyses the sensitivity to price of consumers according to their level of consumption, but does not identify who the heavy consumers are (Finkelstein et al., 2013; Gustavsen and Rickertsen, 2013; Etilé and Sharma, 2015). On the contrary, in our study, we find that obese and overweight individuals are over-represented in the upper deciles of a change in sugar intake in response to the tax.

To provide an order of magnitude of the impact of the change in energy intake due to taxation on health, we used published results from the DIETRON model (Scarborough et al., 2012), an epidemiological model linking a change in nutriment intake to the number of deaths avoided. Based on the average change in energy intake due to the taxation, the impact on the number of deaths avoided is estimated at 300 per year; that is, about 0.8% of the total number of deaths taken into account in the DIETRON model (see details in the Appendix). It shows that the potential impact of a SSB taxation on health should not be under-estimated. This is a (very) rough estimate of the health impact of the tax assuming that the tax only affects non-alcoholic beverage consumption and has no impact on the remaining food and drink consumption: an issue that is debatable. Thus, Finkelstein et al. (2013) find that SSB taxation might generate some substitutions with the consumption of non-beverage items; substitutions that lower the net energy impact of SSB taxation, and, as a consequence, would lower the estimated health impact.

Our approach has some clear limitations. First, the disaggregation method we propose takes into account the individual heterogeneity of consumption due to different individual characteristics (age, gender, BMI), as well as some household heterogeneity. However, there is still some unexplained heterogeneity within the household. Second, in the context of our model, we cannot address the issue of the regressivity of the tax, as we do not have information on the food expenditure share of households. Thus, taxing food consumption is generally considered as regressive (e.g., Allais et al., 2010). However, the health impacts of a policy may in some cases be progressive; that is, it might have a greater impact for low-income consumers than high-income consumers. Finally, though it is not yet evaluated, a tax might act as a signalling device (Cornelsen and Smith, 2018). This effect is ignored in this setting as information issues are not integrated in this framework. However, if confirmed, this mechanism would reinforce the incentive to promote SSB taxation.

References

Allais, O., Bertail, P., and Nichèle, V. (2010). The effects of a fat tax on French households' purchases: a nutritional approach. *American Journal of Agricultural Economics*, 92(1):228–245.

Allais, O. and Tressou, J. (2009). Using decomposed household food acquisitions as inputs of a kinetic dietary exposure model. *Statistical Modelling*, 9(1):27–50.

Allcott, H., Mullainathan, S., and Taubinsky, D. (2014). Energy policy with externalities and internalities. *Journal of Public Economics*, 112:72–88.

Berardi, N., Sevestre, P., Tepaut, M., and Vigneron, A. (2016). The impact of a soda tax on prices: evidence from french micro data. *Applied Economics*, 48(41):3976–3994.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63:841–890.

Bonnet, C., Dubois, P., and Orozco, V. (2014). Household food consumption, individual calories intake and obesity in France. *Empirical Economics, Physica Verlag*, 46(3):1143–1166.

Caby, D. (2016). Obesité : quelles consequences pour l'économie et comment les limiter ? Technical report, Trésor-Eco, 179.

Capacci, S., Allais, O., Bonnet, C., and Mazzocchi, M. (2019). The impact of the French soda tax on prices and purchases. an ex post evaluation. *PLOS ONE*, 14(10):1–22.

Chesher, A. (1998). Individual demands from household aggregates: time and age variation in the composition of diet. *Journal of Applied Econometrics*, 13(5):505–524.

Colchero, M. A., Popkin, B. M., Rivera, J. A., and Ng, S. W. (2015). Beverage purchases from stores in Mexico under the excise tax on sugar sweetened beverages: observational study. *British Medical Journal*, 352:h6704.

Cornelsen, L. and Smith, D. (2018). Viewpoint: Soda taxes. four questions economists need to address. *Food Policy*, 74:138–142.

Dharmasena, S. and Capps, O. J. (2012). Intended and unintended consequences of a proposed national tax on sugar-sweetened beverages to combat the us obesity problem. *Health Economics*, 21:669–694.

Dubois, P., Griffith, R., and O'Connell, M. (2020). How well targeted are soda taxes? *American Economic Review*, 110(11):3661–3704.

Etilé, F. and Sharma, A. (2015). Do high consumers of sugar-sweetened beverages respond differently to price changes? A finite mixture IV-Tobit approach. *Health Economics*, 24(9):1147–1163.

Falbe, J., Thompson, H. R., Becker, C. M., Rojas, N., McCulloch, C. E., and Madsen, K. A. (2016). Impact of the Berkeley excise tax on sugar-sweetened beverage consumption. *American Journal of Public Health*, 106(10):1865ÂŰ1871.

Finkelstein, E., DiBonaventura, M., Burgess, S., and Hale, B. (2010). The costs of obesity in the workplace. *Journal of Occupational and Environmental Medicine*, 52(10):971–976.

Finkelstein, E. A., Ruhm, C., and Kosa, K. (2005). Economic causes and consequences of obesity. *Annual Review of Public Health*, 26(239-257).

Finkelstein, E. A., Zhen, C., Bilger, M., Nonnemaker, J., Farooqui, A. M., and Todd, J. E. (2013). Implications of a sugar-sweetened beverage (SSB) tax when substitutions to non-beverage items are considered. *Journal of Health Economics*, 32(1):219 – 239.

Global Food Research Program (2020). Sugary drink taxes around the world. Technical report, http://globalfoodresearchprogram.web.unc.edu/files/2020/08/SugaryDrink_tax_maps_2020_August_REV.pdf.

Green, R., Cornelsen, L., Dangour, A. D., Turner, R., Shankar, B., Mazzocchi, M., and Smith, R. D. (2013). The effect of rising food prices on food consumption: systematic review with meta-regression. *British Medical Journal*, 346.

Griffith, R., O'Connell, M., and Smith, K. (2017). Corrective taxation and internalities from food consumption. *CESifo Economic Studies*, pages 1–14.

Gustavsen, G. and Rickertsen, K. (2013). Adjusting vat rates to promote healthier diets in Norway: A censored quantile regression approach. *Food Policy*, 42:88–95.

Gustavsen, G. W. and Rickertsen, K. (2011). The effects of taxes on purchases of sugar-sweetened carbonated soft drinks: a quantile regression approach. *Applied Economics*, 43(6):707 – 716.

Han, E. and Powell, L. M. (2013). Consumption patterns of sugar sweetened beverages in the united states. *Journal of the Academy of Nutrition and Dietetics*, 113(1):43–53.

Institut Roche de l'Obésité (2012). Enquête épidémiologique nationale sur le surpoids et l'obésité. Technical report, Institut Roche de l'obésité, 60 pages.

Malik, V., Shulze, M., and Hu, F. (2006). Intake of sugar-sweetened beverages and weight gain: a systematic review. *American Journal of Clinical Nutrition*, 84:274–288.

McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5):447–470.

NCD Risk Factor Collaboration (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet*, 387, Issue 10026:1377 – 1396.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69:307–342.

Petrin, A. and Train, K. (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research*, 47(1):3–13.

Revelt, D. and Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics & Statistics*, 80(4):647 – 657.

Scarborough, P., Allender, S., Clarke, D., Wickramasinghe, K., and Rayner, M. (2012). Modelling the health impact of environmentally sustainable dietary scenarios in the UK. *European Journal of Clinical Nutrition*, 66(6):710–715.

Silver, L., Ng, S., Ryan-Ibarra, S., Taillie, L., Induni, M., Miles, D., and et al. (2017). Changes in prices, sales, consumer spending, and beverage consumption one year after a tax on sugar-sweetened beverages in Berkeley, California, US: A before-and-after study. *PLoS Medicine*, 14(4):e1002283.

Vasdekis, V. and Trichopoulou, A. (2000). Nonparametric estimation of individual food availability along with bootstrap confidence intervals in household budget surveys. *Statistics and Probability Letters*, 46:337–345.

Wada, R., Han, E., and Powell, L. M. (2015). Associations between soda prices and intake: Evidence from 24-h dietary recall data. *Food Policy*, 55:54 – 60.

WHO (2015). *Guideline: Sugars intake for adults and children.* Geneva: World Health Organization, 49 pp.

Zhen, C., Finkelstein, E. A., Nonnemaker, J. M., Karns, S. A., and Todd, J. E. (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.

Figure 1: **Estimated individual consumption of beverages and sugar in beverages**

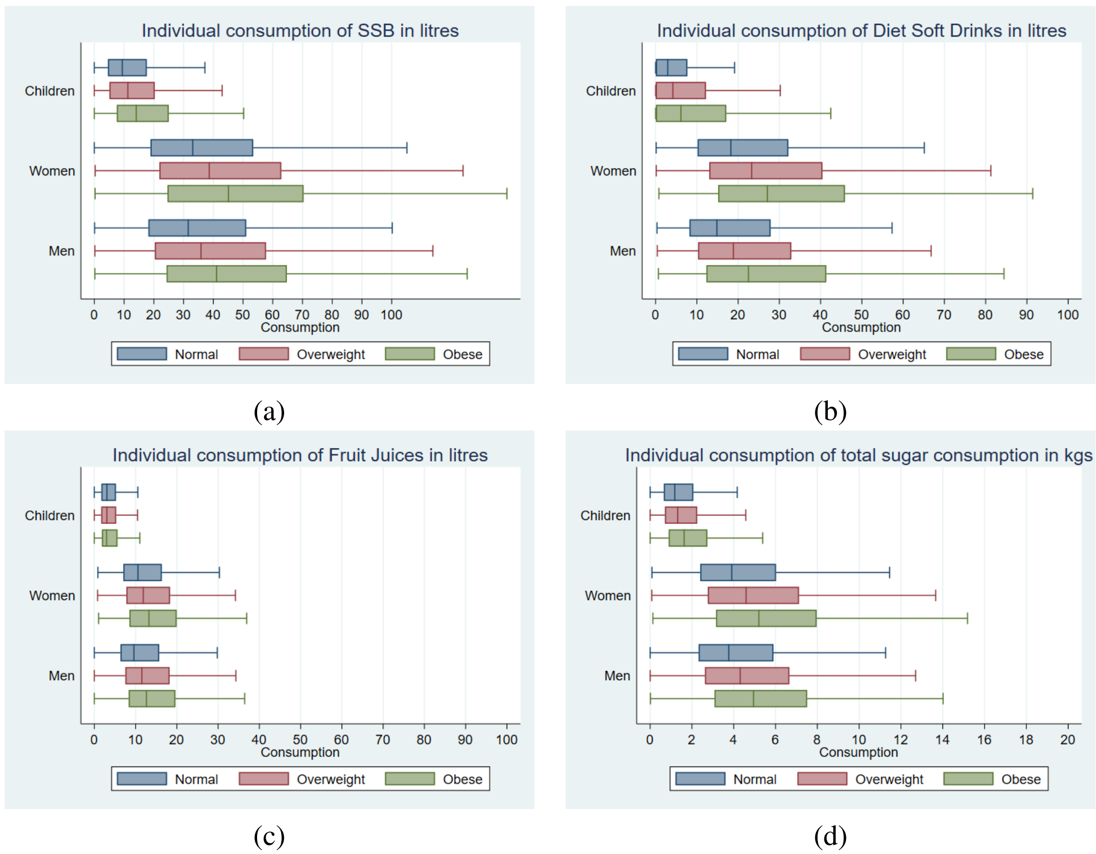


Table 1: **At-home consumption for non-alcoholic beverages in France, 2011**

Presence of children	All	No children			With children		
BMI status of shopper	All	Normal	Overweight	Obese	Normal	Overweight	Obese
Number of households	20.323	7,073	3,820	1,830	4,710	1,882	1,008
Proportion of non-consumer(%)							
Bottled water	10	12	10	9	10	10	10
Regular soft drinks	13	17	16	14	9	7	6
Cola	36	41	41	40	28	28	26
Iced tea	72	77	76	75	65	65	63
Other soft drinks	25	31	29	25	17	15	13
Flavoured water	94	94	94	94	93	95	94
Diet soft drinks	42	47	43	36	40	34	32
Cola	58	64	61	52	56	49	48
Iced tea	95	96	96	96	94	94	93
Other soft drinks	80	83	80	78	78	73	72
Flavoured water	76	78	76	72	74	72	71
Fruit juice and Nectar	8	9	9	11	5	6	6
Fruit juice	10	10	10	14	7	7	7
Nectar	50	54	55	56	42	42	43
SSB (Regular + Nectar)	10	12	12	10	6	5	4
Average individual consumption (l/y or kg/y)*							
Bottled water	62	72	79	86	37	38	40
Regular soft drinks	35	40	43	45	25	25	27
Cola	23	27	29	30	15	15	17
Iced tea	15	19	19	20	10	10	10
Other soft drinks	14	17	19	20	9	9	10
Flavoured water	13	19	17	19	7	7	8
Diet soft drinks	27	31	34	40	17	18	19
Cola	21	24	27	31	12	13	15
Iced tea	13	17	19	20	9	8	7
Other soft drinks	14	19	19	20	9	9	9
Flavoured water	14	18	19	19	8	9	8
Fruit juice and Nectar	21	24	25	27	14	14	15
Fruit juice	14	17	18	19	9	9	10
Nectar	13	15	16	17	8	9	9
SSB (Regular + Nectar)	41	46	49	51	29	29	32
Added sugar (in beverages)	3.5	3.9	4.2	4.4	2.5	2.5	2.7
Total sugar (in beverages)	4.8	5.4	5.7	6.1	3.4	3.5	3.7

* Average consumption is computed over individuals who consume a positive quantity of a product.

Table 2: **Prices paid for non-alcoholic beverages in France, 2011**

Presence of children	No children			With children		
BMI status of shopper	Normal	Overweight	Obese	Normal	Overweight	Obese
Prices (€/l)*						
Bottled water	0.30 (0.05)	0.30 (0.06)	0.31 (0.06)	0.27 (0.05)	0.27 (0.06)	0.27 (0.05)
Regular soft drinks						
Cola	0.75 (0.27)	0.77 (0.28)	0.75 (0.32)	0.69 (0.25)	0.67 (0.26)	0.69 (0.27)
Iced tea	0.86 (0.32)	0.91 (0.37)	0.86 (0.34)	0.82 (0.30)	0.78 (0.39)	0.85 (0.39)
Other soft drinks	1.61 (1.21)	1.57 (1.14)	1.59 (1.16)	1.55 (1.20)	1.50 (1.17)	1.52 (1.50)
Flavoured water	0.90 (0.10)	0.92 (0.12)	0.99 (0.19)	0.95 (0.12)	1.00 (0.19)	0.94 (0.24)
Diet soft drinks						
Cola	0.68 (0.31)	0.68 (0.33)	0.70 (0.36)	0.66 (0.30)	0.67 (0.26)	0.66 (0.32)
Iced tea	0.87 (0.31)	0.95 (0.37)	0.97 (0.30)	0.87 (0.34)	0.81 (0.35)	0.81 (0.37)
Other soft drinks	1.57 (1.60)	1.66 (1.99)	1.38 (1.64)	1.71 (2.05)	1.34 (1.61)	1.59 (1.98))
Flavoured water	0.92 (0.33)	0.92 (0.37)	0.93 (0.39)	0.91 (0.31)	0.88 (0.31)	0.91 (0.39)
Fruit juice and Nectar						
Fruit juice	2.27 (1.55)	2.12 (1.39)	2.05 (1.40)	2.07 (1.47)	2.06 (4.45)	1.78 (1.15)
Nectar	1.38 (0.52)	1.42 (0.54)	1.37 (0.54)	1.28 (0.52)	1.25 (0.55)	1.19 (0.59)

* (): standard deviation across brands, retailers and periods.

Table 3: **Distribution of estimated individual consumption of SSBs for different groups of consumers (litres/year)**

	Mean	Median	75 ptile	90 ptile	95 ptile	99 ptile
Children						
Normal weight	13	9	17	29	40	67
Overweight	15	11	20	33	44	79
Obese	19	14	25	39	52	100
Women						
Normal weight	42	33	53	82	109	176
Overweight	47	38	62	93	118	179
Obese	54	45	70	103	139	206
Men						
Normal weight	40	31	51	78	102	180
Overweight	45	35	57	87	113	189
Obese	50	41	64	98	128	186

Average, median, and centiles computed over consumers who actually consume a positive quantity of the beverages.

Table 4: **Distribution of estimated individual intake of total sugar for different groups of consumers (kg/year)**

	Mean	Median	75 ptile	90 ptile	95 ptile	99 ptile
Children						
Normal weight	1.59	1.18	2.06	3.28	4.35	7.35
Overweight	1.75	1.31	2.25	3.60	4.57	8.76
Obese	2.10	1.62	2.73	4.26	5.51	10.51
Women						
Normal weight	4.77	3.90	6.02	9.04	11.60	18.47
Overweight	5.42	4.59	7.12	10.17	12.57	18.69
Obese	6.13	5.20	7.96	11.26	14.58	21.67
Men						
Normal weight	4.66	3.76	5.90	8.77	11.33	19.17
Overweight	5.21	4.30	6.66	9.72	12.47	20.35
Obese	5.88	4.94	7.51	11.16	13.81	19.53

Average, median, and centiles computed over consumers who actually consume a positive quantity of the beverages.

Table 5: **Proportion (%) of different types of consumers for different percentiles of intake of total sugar**

	All	< 50 ptile	50 - 75 ptile	75 - 90 ptile	90 - 95 ptile	> 95 ptile
Children						
Sugar intake (kg)		< 0.85	0.85 - 1.73	1.73 - 2.94	2.94 - 3.95	> 3.95
Normal weight	85.3	85.6	86.3	83.9	82.5	81.4
Overweight	10.6	10.6	10.1	10.8	11.8	12.0
Obese	4.1	3.6	3.7	5.2	5.8	6.6
Adults						
Sugar intake (kg)		< 3.20	3.20 - 5.64	5.64 - 8.72	8.72 - 11.22	> 11.22
Normal weight men	24.9	27.0	24.6	21.8	19.5	19.7
Overweight men	16.7	15.6	17.2	18.4	17.3	18.1
Obese men	5.4	4.6	5.3	7.1	6.8	8.4
Normal weight women	34.1	36.2	34.3	30.1	28.8	28.9
Overweight women	12.3	11.1	12.5	14.0	16.9	14.4
Obese women	6.6	5.5	5.6	8.5	10.5	10.5

Average, median, and centiles computed over consumers who actually consume a positive quantity of the beverages.

Table 6: **Aggregated elasticities for households without children**

(a) Obesity status of the panelist: normal

	ColaR	FJ	Nectars	IceTeaR	OtherSDR	FlavWatR	Diet	OG
ColaR	-1.6903	0.1651	0.1664	0.1750	0.1695	0.1790	0.1750	0.1368
FJ	0.8336	-1.8983	0.8352	0.8200	0.8343	0.8246	0.8309	0.5741
Nectars	0.1060	0.1089	-2.1259	0.1049	0.1052	0.1079	0.1071	0.0916
IceTeaR	0.0364	0.0341	0.0342	-1.5780	0.0352	0.0376	0.0365	0.0423
OtherSDR	0.2722	0.2706	0.2660	0.2700	-1.9525	0.2794	0.2715	0.1769
FlavWatR	0.0068	0.0062	0.0057	0.0069	0.0064	-1.7624	0.0070	0.0114
Diet	0.1614	0.1524	0.1503	0.1607	0.1550	0.1686	-1.5782	0.1392

(b) Obesity status of the panelist: overweight

	ColaR	FJ	Nectars	IceTeaR	OtherSDR	FlavWatR	Diet	OG
ColaR	-1.9602	0.1845	0.1851	0.1931	0.1869	0.2010	0.1934	0.1473
FJ	0.7419	-2.3481	0.7732	0.7329	0.7480	0.7327	0.7334	0.6577
Nectars	0.1202	0.1230	-2.6068	0.1186	0.1193	0.1185	0.1187	0.2109
IceTeaR	0.0402	0.0383	0.0384	-1.9889	0.0391	0.0426	0.0406	0.0482
OtherSDR	0.4847	0.4908	0.4812	0.4830	-2.7494	0.5098	0.4863	0.6470
FlavWatR	0.0096	0.0094	0.0094	0.0099	0.0096	-2.1432	0.0105	0.0125
Diet	0.2113	0.2032	0.2042	0.2129	0.2078	0.2273	-1.7817	0.1987

(c) Obesity status of the panelist: obese

	ColaR	FJ	Nectars	IceTeaR	OtherSDR	FlavWatR	Diet	OG
ColaR	-1.8624	0.1531	0.1550	0.1602	0.1532	0.1650	0.1592	0.3412
FJ	0.4896	-2.1469	0.4986	0.4802	0.4811	0.4874	0.4885	0.3864
Nectars	0.0917	0.0957	-2.4004	0.0910	0.0901	0.0931	0.0926	0.0749
IceTeaR	0.0360	0.0346	0.0354	-1.6486	0.0348	0.0373	0.0359	0.0299
OtherSDR	0.9685	0.9751	0.9844	0.9813	-2.5506	1.0089	0.9911	0.5743
FlavWatR	0.0091	0.0086	0.0085	0.0091	0.0086	-2.0724	0.0090	0.0103
Diet	0.2113	0.2036	0.2065	0.2126	0.2059	0.2225	-1.5771	0.0924

Table 7: **Aggregated elasticities for households with children**

(a) Obesity status of the panelist: normal

	ColaR	FJ	Nectars	IceTeaR	OtherSDR	FlavWatR	Diet	OG
ColaR	-1.9602	0.1845	0.1851	0.1931	0.1869	0.2010	0.1934	0.1473
FJ	0.7419	-2.3481	0.7732	0.7329	0.7480	0.7327	0.7334	0.6577
Nectars	0.1202	0.1230	-2.6068	0.1186	0.1193	0.1185	0.1187	0.2109
IceTeaR	0.0402	0.0383	0.0384	-1.9889	0.0391	0.0426	0.0406	0.0482
OtherSDR	0.4847	0.4908	0.4812	0.4830	-2.7494	0.5098	0.4863	0.6470
FlavWatR	0.0096	0.0094	0.0094	0.0099	0.0096	-2.1432	0.0105	0.0125
Diet	0.2113	0.2032	0.2042	0.2129	0.2078	0.2273	-1.7817	0.1987

(b) Obesity status of the panelist: overweight

	ColaR	FJ	Nectars	IceTeaR	OtherSDR	FlavWatR	Diet	OG
ColaR	-1.4265	0.1687	0.1680	0.1740	0.1705	0.1776	0.1743	0.1340
FJ	0.6171	-1.7265	0.6262	0.6081	0.6079	0.6153	0.6058	0.5328
Nectars	0.1167	0.1184	-1.9296	0.1144	0.1149	0.1156	0.1139	0.1247
IceTeaR	0.0376	0.0357	0.0345	-1.3208	0.0365	0.0385	0.0378	0.0411
OtherSDR	0.3515	0.3482	0.3412	0.3474	-1.6392	0.3679	0.3492	0.2545
FlavWatR	0.0075	0.0071	0.0063	0.0074	0.0070	-1.9201	0.0075	0.0142
Diet	0.2015	0.1978	0.1939	0.2022	0.1979	0.2096	-1.3647	0.1882

(c) Obesity status of the panelist: obese

	ColaR	FJ	Nectars	IceTeaR	OtherSDR	FlavWatR	Diet	OG
ColaR	-1.3787	0.1739	0.1718	0.1797	0.1741	0.1827	0.1769	0.0392
FJ	0.5700	-1.6514	0.5755	0.5668	0.5620	0.5754	0.5669	0.2788
Nectars	0.1049	0.1071	-1.8769	0.1040	0.1022	0.1048	0.1035	0.1210
IceTeaR	0.0407	0.0397	0.0385	-1.4031	0.0390	0.0424	0.0401	0.0281
OtherSDR	0.3839	0.3825	0.3654	0.3792	-1.5141	0.4037	0.3734	0.3924
FlavWatR	0.0037	0.0038	0.0035	0.0042	0.0038	-1.7176	0.0037	0.0074
Diet	0.1878	0.1834	0.1825	0.1901	0.1831	0.1931	-1.4190	0.1817

Table 8: **Distribution of variation in estimated individual consumption of total sugar (kg/year)**

		Average	Median	75 ptile	90 ptile	95 ptile	99 ptile
Children	**Obesity status**						
	Normal weight	-0.288	-0.215	-0.362	-0.575	-0.767	-1.370
	Overweight	-0.320	-0.246	-0.405	-0.632	-0.837	-1.434
	Obese	-0.381	-0.302	-0.481	-0.752	-0.928	-2.204
Women	**Obesity status**						
	Normal weight	-0.942	-0.761	-1.174	-1.721	-2.208	-3.524
	Overweight	-1.224	-1.006	-1.575	-2.311	-2.840	-4.196
	Obese	-1.260	-1.050	-1.591	-2.269	-2.883	-4.700
Men	**Obesity status**						
	Normal weight	-0.939	-0.750	-1.144	-1.735	-2.236	-3.897
	Overweight	-1.118	-0.891	-1.385	-2.075	-2.706	-4.633
	Obese	-1.225	-1.018	-1.532	-2.221	-2.807	-4.462

Average, median, and centiles computed over consumers who consume at least one sugary beverage.

Table 9: **Proportion (%) of different types of consumers for different percentiles of decrease in sugar intake due to taxation**

	All	< 50 ptile	50 - 75 ptile	75 - 90 ptile	90 - 95 ptile	> 95 ptile
Children						
Decrease in sugar intake (kg)		< 0.22	0.22 - 0.37	0.37 - 0.59	0.59 - 0.78	> 0.78
Normal weight	85.3	86.2	85.3	83.1	81.2	81.1
Overweight	10.6	10.2	10.7	10.9	13.7	11.8
Obese	4.1	3.6	4.0	6.1	5.1	7.1
Adults						
Sugar intake (kg)		< 0.84	0.84 - 1.30	1.30 - 1.95	1.95 - 2.50	> 2.50
Normal weight men	24.9	27.1	23.6	19.8	19.0	17.7
Overweight men	16.7	15.6	17.6	18.5	19.4	20.6
Obese men	5.4	4.5	6.4	7.3	7.4	8.4
Normal weight women	34.1	37.0	32.1	28.8	24.7	23.4
Overweight women	12.3	10.4	13.2	16.2	18.9	20.0
Obese women	6.6	5.4	7.1	9.5	10.6	10.0

Average, median, and centiles computed over consumers who actually consume a positive quantity of the beverages.

Table 10: **Difference in the estimated impact on total sugar consumption of the tax: average per capita household consumption versus our method to estimate individual consumption; average difference and standard error (g/year)**

	Children	Men	Women
Obesity status			
Normal weight	-121 (217)*	14 (175)*	21 (165)*
Overweight	-100 (221)*	61 (204)*	58 (208)*
Obese	-37 (207)*	98 (246)*	99 (211)*

* means that the averages of per capita household consumption and estimated individual consumption are significantly different from 0 at 1%.

6 Appendix

6.1 Descriptive statistics

Table 11: **Descriptive statistics of household characteristics**

	All households	Single households
Income		
Modest	15%	20%
Middle	72%	72%
High	13%	8%
Education		
Low	29%	28%
Middle	27%	24%
High	44%	48%
Socio-professional category		
Farmers	1%	0%
Artisans, merchants and employers	4%	1%
Managers and engineers	12%	11%
Intermediate (foremen, etc.)	18%	19%
Employees	23%	25%
Workers	19%	6%
Retired	21%	31%
Unemployed	4%	7%
Region		
Paris	19%	23%
East	9%	8%
North	10%	8%
West	20%	19%
Centre West	22%	22%
South	21%	20%
Rural	26%	16%
Nb of women	1.37 (0.88)	0.58 (0.49)
Nb of children		
< 10	0.50 (0.84)	-
[11;17]	0.25 (0.58))	-
BMI of the family's head	24.86 (4.94)	25.20 (4.99))

6.2 Specification of the disaggregation model

To select which household characteristics will enter the model of disaggregation for each beverage category, we first run simple regressions of per-capita household consumption using household characteristics as explanatory variables (Table 12). In this table, there are two groups of variables. From the first group of variables (income, education, socio-professional category (SPC), region, rural area), we select those that are significant for each beverage category which are then integrated into the disaggregation model. Variables from the second group (number of females, number of children $<$10, number of teenagers, BMI of the family's head) are control variables for the regressions.[25] Among the first group of variables, that is, variables that will be used as household characteristics in the disaggregation model, the variable 'Rural' is not significant in the case of two product categories, namely iced tea and flavoured water. We thus exclude it in the disaggregation model for the two product categories. With respect to the region of residence, we exclude it from the disaggregation model for nectar as none of the region fixed effects has a significant impact. With respect to the SPC, we include the variable in the model for nectar as several SPCs seem to significantly affect nectar consumption. Finally, income and education variables are likely to affect consumption, and in some cases, both variables affect per-capita household consumption. Because the two variables are significantly correlated, in order to determine which variable to include, we run three alternative disaggregation models for each beverage category. These models use the following explanatory variables: the variables that were previously selected alone (model 1); the set of variables included in model 1 plus the income variable (model 2); and the set of variables included in model 1 plus the education variable (model 3).[26] We then select the best model on the basis of the mean squared error (Table 13 in the Annex). The level of education is selected in the models for flavoured water and for diet SDs. For all other beverages, the income variable is selected. The list of variables included in the disaggregation model for each beverage category is summarized in Table 14.

[25]According to the results, the second group of variables are in most cases strongly significant, justifying their presence as control variables in the regression.

[26]We do not test a fourth model with both income and education, as we have constraints on the number of variables that will be finally included in the disaggregation model.

Table 12: **Regression of per-capita household consumption on household characteristics.**

	Cola	**Fruit J**	**Nectar**	**Iced Tea**	**Other SD**	**Flav W**	**Diet**
Income							
Middle	-1.10 (0.62)*	-0.28 (0.24)	0.73 (0.39)*	-0.28 (0.70)	1.02 (0.34)***	0.95 (1.06)	1.48 (0.79)*
High	-0.86 (0.86)	-0.61 (0.34)*	1.24 (0.52)**	0.08 (0.86)	1.89 (0.45)***	2.43 (1.40)*	1.02 (1.09)
Education							
Middle	-0.27 (0.48)	-0.10 (0.21)	-0.67 (0.31)**	0.01 (0.49)	-0.46 (0.27)*	-0.72 (0.85)	-0.68 (0.67)
High	-0.84 (0.47)*	-0.23 (0.20)	-0.66 (0.30)**	0.41 (0.49)	-0.72 (0.26)***	-0.77 (0.82)	-0.63 (0.65)
Socio-professional category							
Artisans	-1.60 (2.03)	-1.63 (0.92)*	-2.99 (1.41)**	1.43 (2.19)	-0.24 (1.15)	2.97 (4.20)	-0.70 (3.11)
Managers	-1.32 (1.88)	-0.71 (0.85)	-2.47 (1.32)*	-1.14 (2.05)	0.26 (1.06)	5.47 (3.96)	1.76 (2.91)
Intermediate	-0.73 (1.83)	-0.54 (0.83)	-1.71 (1.30)	0.54 (2.00)	0.76 (1.03)	5.77 (3.90)	3.40 (2.86)
Employees	0.54 (1.82)	-0.40 (0.83)	-1.40 (1.29)	1.01 (1.98)	1.13 (1.03)	4.87 (3.88)	1.51 (2.85)
Workers	0.04 (1.81)	-0.36 (0.83)	-1.59 (1.29)	1.29 (1.97)	0.85 (1.03)	5.35 (3.87)	1.79 (2.85)
Retired	-3.09 (1.84)*	-0.58 (0.83)	-2.33 (1.30)*	-0.17 (2.01)	0.67 (1.04)	6.70 (3.91)*	0.11 (2.87)
Unemployed	-2.03 (2.09)	-1.07 (0.93)	-0.19 (1.45)	3.62 (2.24)	0.17 (1.18)	5.96 (4.35)	-0.68 (3.19)
Region							
East	3.37 (0.73)***	1.15 (0.31)***	0.66 (0.47)	1.14 (0.73)	1.43 (0.41)***	2.25 (1.26)*	2.43 (0.97)***
North	3.07 (0.70)***	1.46 (0.30)***	0.72 (0.45)	3.14 (0.72)***	2.04 (0.39)***	2.33 (1.17)**	4.63 (0.93)***
West	-0.03 (0.61)	1.43 (0.26)***	0.50 (0.38)	0.37 (0.68)	-0.10 (0.34)	-0.17 (1.11)	-1.90 (0.83)**
Centrer West	0.23 (0.59)	0.83 (0.25)***	-0.02(0.366)	0.80 (0.62)	0.48 (0.33)	0.11 (1.03)	-1.68 (0.79)**
South	-1.10 (0.60)*	0.32 (0.25)	-0.17 (0.37)	-0.08 (0.64)	-0.46 (0.33)	1.62 (1.05)	-2.39 (0.79)***
Rural	0.98 (0.42)**	1.05 (0.18)***	0.58 (0.27)**	0.05 (0.44)	0.95 (0.23)***	0.80 (0.74)	1.54 (0.59)***
Nb of female	-1.08 (0.30)***	-0.78 (0.13)***	-0.70 (0.19)***	-0.59 (0.30)**	-0.44 (0.17)***	-0.33 (0.50)	-0.58 (0.41)
Nb of children							
< 10	1.61 (0.31)***	1.49 (0.14)***	1.50 (0.20)***	1.35 (0.31)***	1.78 (0.18)***	1.13 (0.54)**	2.17 (0.44)***
[11;17]	1.93 (0.39)***	2.51 (0.18)***	2.68 (0.25)***	1.87 (0.38)***	2.43 (0.22)***	1.95 (0.69)***	3.46 (0.55)***
BMI of the family's head	0.22 (0.03)***	0.13 (0.01)***	0.11 (0.02)***	0.09 (0.03)**	0.12 (0.02)***	0.04 (0.06)	0.39 (0.05)***
Constant	37.66 (2.20)***	22.74 (0.98)***	23.26 (1.52)***	27.03 (2.39)***	23.29 (1.25)***	23.87 (4.44)***	36.78 (3.32)***

The income of reference is 'low', the education level of reference is 'low', the SPC of reference is 'farmers'; the region of reference is Paris.

Table 13: **Mean squared error of disaggregation models.**

	Model 1	**Model 2**	**Model 3**
Cola	2212	**2173**	2177
FJ	548	**536**	538
Nectar	770	**754**	761
Iced Tea	1037	**1021**	1022
Other SD	612	**603**	604
FW	1360	1330	**1301**
Diet	3128	3028	**2991**

Model 1 does not consider income and education variables.

Model 2 considers only income variable.

Model 3 considers only education variable.

Table 14: **Household variables included in the disaggregation model**

	Cola	FJ	Nectar	Iced tea	Other SD	FW	Diet
Income	x	x	x	x	x		
Education						x	x
SPC			x				
Region	x	x		x	x	x	x
Rural	x	x	x		x		x

6.3 Result on the estimated individual consumption

Table 15: **p value of the means test for equality of mean consumption**

	Regular SSBs	Diet soft drinks	Fruit juices	Added sugar intake	Total sugar intake	Variation of sugar intake (due to a tax)
Children						
Normal-Overweight	0.00	0.00	0.26	0.00	0.00	0.01
Overweight-Obese	0.00	0.02	0.12	0.00	0.00	0.01
Women						
Normal-Overweight	0.00	0.00	0.00	0.00	0.00	0.00
Overweight-Obese	0.00	0.00	0.01	0.00	0.00	0.00
Men						
Normal-Overweight	0.00	0.00	0.00	0.00	0.00	0.00
Overweight-Obese	0.00	0.00	0.00	0.00	0.00	0.00
Men versus Women	0.00	0.00	0.51	0.00	0.02	0.01

H0: both samples come from a population with the same mean; $p < 0.05$ rejects H0.

Table 16: **P value of the two-sample Kolmogorov-Smirnov test for equality of distribution functions**

	Regular SSBs	Diet soft drinks	Fruit Juice	Added sugar intake	Total sugar intake	Variation of sugar intake (due to a tax)
Children						
Normal-Overweight	0.00*	0.00	0.74	0.00	0.00*	0.02*
Overweight-Obese	0.00*	0.01	0.53	0.00	0.00*	0.00*
Women						
Normal-Overweight	0.00*	0.00	0.00	0.00	0.00*	0.00*
Overweight-Obese	0.00*	0.00	0.28	0.00	0.00	0.24
Men						
Normal-Overweight	0.00*	0.00	0.00	0.00	0.00*	0.00*
Overweight-Obese	0.00*	0.00	0.00	0.00	0.00*	0.00*

H0: both samples come from a population with the same distribution; $p < 0.05$ rejects H0.
* means that the second distribution statistically dominates the first one by using the Dunn's pairwise comparison test of distribution (p-value is always 0.00). We perform the test for the consumption of regular SDs, for the total sugar intake, and for the variation in total sugar intake. We check that the compared cumulative distributions never cross. This is always the case except for the comparison of the variation of total sugar intake for overweight versus obese women. In this latter case, we cannot apply the Dunn's pairwise comparison test.

6.4 Demand model

Given the assumptions about the indirect utility function and assuming that ε_{ijt} is independently and identically distributed as an extreme value type I distribution, the market share of product j in month t is given by (Nevo, 2001):

$$s_{jt} = \int_{A_{jt}} \left(\frac{\exp(\alpha_i p_{jr(i)t} + \mu_{b(j)} + \sum_{c=1}^{C} \gamma_{c(j)})}{1 + \sum_{k=1}^{J_{it}} \exp(\alpha_i p_{kr(i)t} + \mu_{b(k)} + \sum_{c=1}^{C} \gamma_{c(k)})} \right) dP_\nu(\nu), \tag{5}$$

where A_{jt} is the set of households i that have the highest utility for product j in month t in the retailer visited by the household i, J_{it} the number of products available at month t for the household i, and a household is defined by the vector $(\nu_i, \varepsilon_{i0t}, ..., \varepsilon_{iJt})$. We assume that P_ν follows a cumulative normal distribution with mean α and standard deviation σ.

The random coefficient logit model generates a flexible pattern of substitutions between products driven by the different consumer price disutilities α_i. Thus, the own- and cross-price elasticities of the market share s_{jt} are written as:

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial s_{jit}}{\partial p_{kr(i)t}} \frac{p_{kr(i)t}}{s_{jit}} = \begin{cases} -\frac{1}{N} \sum_{i=1}^{N} \frac{p_{jr(i)t}}{s_{jit}} \int \alpha_i s_{ijt}(1 - s_{ijt})\, dP_\nu(\nu_i) & \text{if } j = k \\ \frac{1}{N} \sum_{i=1}^{N} \frac{p_{kr(i)t}}{s_{jit}} \int \alpha_i s_{ijt} s_{ikt}\, dP_\nu(\nu_i) & \text{otherwise.} \end{cases} \tag{6}$$

The above model is estimated using data purchases for 60 different products. As a consequence, the own- and cross-price elasticities are defined at the product level. We deduce the own- and cross-price elasticities at the category level $\eta_{cc't}$ as the variation of the market share of the category c, when the prices of all products belonging to the category c' increase by 1% at month t,

$$\eta_{cc't} = \frac{\partial s_{ct}}{\partial p_{c't}} \frac{p_{c't}}{s_{ct}} = \sum_{j \in c'} \eta_{cjt} \tag{7}$$

with

$$\eta_{cjt} = \frac{\partial s_{ct}}{\partial p_{jt}} \frac{p_{jt}}{s_{ct}} = \sum_{k \in c} \frac{\partial s_{kt}}{\partial p_{jt}} \frac{p_{jt}}{s_{kt}} \frac{s_{kt}}{s_{ct}} = \sum_{k \in c} \eta_{kjt} \frac{s_{kt}}{s_{ct}}$$

where η_{cjt} represents the percentage variation of the market share of category c when the price of product j increases by 1% at month t.

6.4.1 Identification

The estimation of the demand parameters relies on the assumption that the prices p_{jt} are independent of the error term ε_{ijt}. However, assuming $\varepsilon_{ijt} = \xi_{jt} + e_{ijt}$ where ξ_{jt} is a product-specific error term varying across periods and e_{ijt} is an individual-specific error term, the independence assumption cannot hold if unobserved factors included in ξ_{jt} (and hence in ε_{ijt}) are correlated with prices. For example, unobserved promotions, displays, and advertising of some products are likely to affect both prices and demand. To solve the problem where omitted product characteristics may be correlated with price, we use a control function approach as in Petrin and Train (2010). We regress prices on instrumental variables (W_{jt}) and the exogenous variables of the demand equation (product fixed effects):

$$p_{jt} = W_{jt}\gamma + \theta_j + \eta_{jt},$$

where η_{jt} is an error term that captures the remaining unobserved variation in prices. The estimated error term $\widehat{\eta}_{jt}$ of the price equation includes some omitted variables such as advertising variations and promotions that could explain price variations across products and time periods. Introducing this term in the mean utility of households δ_{jt} allows us to capture unobserved product characteristics varying across time.[27] Prices are now uncorrelated with the new product-specific error term varying across periods ($\zeta_{jt} = \xi_{jt} - \lambda\widehat{\eta}_{jt}$). We write:

$$\delta_{jt} = \mu_j - \alpha p_{jt} + \lambda\widehat{\eta}_{jt} + \zeta_{jt},$$

where λ is the estimated parameter associated with the estimated error term of the first stage.

In practice, we use input price indexes of water and sugar as it is unlikely that input prices are correlated with unobserved determinants of demand for soft drinks.[28]

Table 17: **Results of the price equation.**

	Without children			With children		
	Normal	Overweight	Obese	Normal	Overweight	Obese
Number of	0.02***	0.02***	0.03***	0.02***	0.01***	0.02***
competing products	(0.00)	(0.00)	(0.00)	0.02 (0.00)	0.01 (0.00)	(0.00)
Sugar price	0.00***	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Diet	1.61***	1.41***	1.44***	1.57***	1.92***	1.35***
	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.11)
Category fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Brand fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
F test IV (p value)	374(0.00)	169(0.00)	187(0.00)	183(0.00)	250(0.00)	61(0.00)
R^2	0.88	0.85	0.85	0.84	0.85	0.77
Number of observations	3,985	3,771	3,600	3,949	3,531	3,286

Standard errors are in parenthesis.

*** means significant at 1%.

6.4.2 Results of the demand model

Results of the six demand models are similar.[29] The estimated coefficients are of the same order of magnitude and of the same signs, with very few exceptions for some fixed effects. The following comments thus apply to any of the six demand models. First, the coefficient of the error term is positive and significant, which means that the unobserved part explaining prices is positively correlated with the choice of

[27] We have $\delta_{jt} = \mu_j - \alpha_j p_{jt} + \xi_{jt}$ with $\alpha_j = \alpha_{SD(j)} + \alpha_{D(j)} + \alpha_{FW(j)}$.

[28] These indexes are from the French National Institute for Statistics and Economic Studies.

[29] The subsamples of households without/with children and for the different obesity status (normal, overweight, obese) contains 295,702, 169,389, 107,179, 258,065, 104,352, and 70,538 observations respectively. For computational reasons, we randomly draw 100,000 observations in each of the six subsamples, except for the last one for which we keep the 70,538 observations. We used the simulated maximum likelihood method as in Revelt and Train (1998).

the product, thus justifying the need to control for endogeneity. On average, the price has a significant and negative impact on utility. Moreover, because the standard deviation is small relative to the average coefficient, price has a negative impact on utility for almost all consumers. Consumers prefer fruit juice as the fixed effect is the highest for all groups of households, and prefer regular products as the diet category fixed effect is negative. Brand fixed effects also play a role.[30] For example, even if the cola fixed effect is negative, consumers have a preference for some specific colas as brand fixed effects are significant, positive, and overcompensate for the negative value of the cola fixed effect for some specific brands.

[30] Brand fixed effects are not reported for confidentiality reasons.

Table 18: **Results of the random coefficients logit model**

	Without children			With children		
	Normal	Overweight	Obese	Normal	Overweight	Obese
Price						
Mean	-2.04	-2.37	-2.16	-1.97	-2.00	-1.96
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Std	0.41	0.37	0.34	0.26	-0.33	0.28
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Error term	1.45	1.73	1.64	1.52	1.44	1.49
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Diet	-1.20	-0.94	-0.77	-1.12	-0.74	-0.73
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
FW	-1.44	-1.22	-1.19	-1.41	-1.45	-1.30
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Cola	-1.90	-1.74	-1.69	-1.71	-1.51	-1.49
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Iced tea	-2.25	-1.82	-1.89	-1.83	-1.76	-1.64
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Other SD	-0.20	0.73	0.41	0.89	0.60	0.94
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
FJ	2.57	2.95	2.42	2.44	2.30	1.94
	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Nectars	-1.12	0.38	0.01	0.02	0.03	-0.16
	(0.00)	(0.00)***	(0.00)***	(0.00)***	(0.00)***	(0.00)***
Brand fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Log-likelihood	-263,204	-263,203	-270,911	-277,824	-286,321	-201,332
Number of observations	100,000	100,000	100,000	100,000	100,000	70,538

Standard errors are in parenthesis.

*** means significant at 1%.

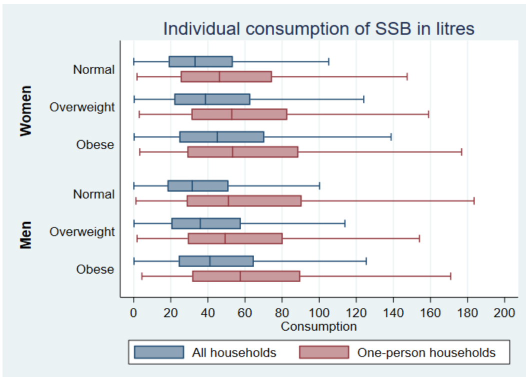
6.5 Comparison of estimated consumption of SSBs and observed consumption of single adult households

Figure 2(a) presents a comparison of the estimated distribution of individual consumption of SSBs with the observed distribution of SSB consumption of one-person household (which in the following is assumed to be an adult). As above, we distinguish consumers according to their gender and class of BMI.

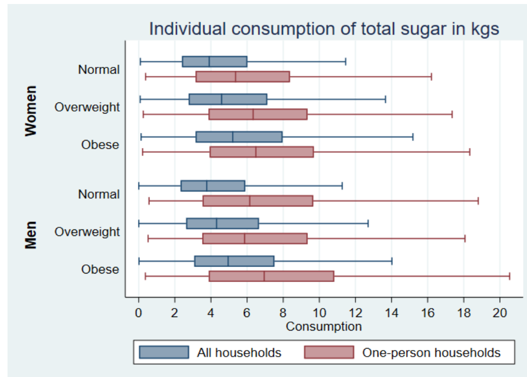
Data from one-person households reveal a large heterogeneity in the level of consumption for a given class of consumers. We also observe that the consumption of obese people is greater than that of normal weight individuals. For overweight individuals, results are less clear-cut. For SSBs, overweight women consume more than normal weight women, while results are the contrary for men. As compared to the estimated distributions, the main difference lies in the level of consumption. Thus, the observed consumption for one-person households is higher than the estimated consumption for adults of identical gender and class of BMI. There are various possible explanations for this difference. First, socio-demographic characteristics are different (Table 11). For example, one-person households tend to be more urban, which has a positive impact on the consumption of beverages (Table 12). Second, the average consumption per individual in a household decreases with the size of the household. Finally, it is important to bear in mind that for any household the sum of the estimated consumption for each individual is exactly equal to the observed consumption of the household. Then, a systematic under (or over) estimate of individual consumption is not possible. As a consequence, the observed differences in consumption levels are more likely to be related to the household characteristics, both observed and unobserved, rather than to a systematic bias in the estimated distribution. Finally, observed data is more heterogenous than the estimated one.

Figure 2(b) presents a comparison of the estimated distribution of individual sugar intake (from beverages) with the observed distribution of sugar intake of one-person households. Results are very similar to those discussed previously in the case of SSB consumption. That is, there is mainly a large heterogeneity of sugar intake for a given class of consumers, a sugar intake larger for obese than normal weight individuals, and a difference in the level between observed intake and estimated intake; a difference that

Figure 2: **Comparison between observed consumption for one-person households and estimated individual consumption of SSBs and sugar in beverages**



(a)



(b)

is likely related to the household characteristics.

6.6 Estimating the health impact

To assess the health impact of changes in sugar intake, we use the DIETRON model which evaluates the impact on mortality of changes in diets. As explained by Scarborough et al. (2012): 'the DIETRON model uses age- and sex-specific estimates of relative risk drawn from meta-analyzes of trials, cohort studies and case-control studies, to estimate the impact on chronic disease mortality of counterfactual population dietary scenarios'. The DIETRON model considers various nutrients (total energy, fibres, total fat, monounsaturated fatty acids, polyunsaturated fatty acids, saturated fatty acids, dietary cholesterol, salt (g/day) as well as fruit and vegetables). In our case, only the energy intake is modified. In DIETRON, a change in energy intake has an indirect impact on mortality. Thus, a change in energy intake impacts the BMI and a change in the BMI has an impact on mortality.

Formally, for a disease d and type t individuals, the number of deaths avoided for disease d and type t (DA_{dt}), is related to a change in the BMI. We have:

$$DA_{dt} = RR_{dt}^{\frac{\Delta(BMI)_t}{Unit_d}} ND_{dt}$$

where RR_{dt} is the relative risk for disease d and type t, $\Delta(BMI)_t$ is the change in BMI for type t, $Unit_d$ is the unit of the change in BMI for disease d, and ND_{dt} is the number of deaths from disease d for type t. In practice, individuals are defined by their gender (male; female), their class of age (25-59; 60-74), and their class of BMI (< 25; ≥ 25). We thus have 8 types of individuals. As we consider premature deaths, we limit the age to 74. Obviously, the total number of deaths avoided DA is the sum for all diseases and all types.

The change in BMI is related to the change in energy intake, we have:

$$\Delta(BMI)_t = \frac{\Delta(BW)_t}{H_t^2} = \frac{1}{H_t^2}\frac{k_t\Delta(EI)_t}{PAL}$$

where $\Delta(BW)_t$ is the change in body weight, H_t is the average height of individuals of type t, $\Delta(EI)_t$ is the average change in energy intake of individuals of type t, PAL is the physical activity level (we assume that PAL remains constant and does not depend on type t), and k_t is a parameter for type t.

When $\Delta(EI)_t$ is in MJ/day and $\Delta(BW)_t$ in kg, in DIETRON k_t= 17.7 for men and 20.7 for women. To provide an order of magnitude, a 10 kcal per day decrease in calorie intake with a PAL=1.6 (which corresponds to a moderate level of physical activity) translates to a decrease in body weight by 0.46 kg and 0.54 kg at steady state, for men and women respectively.

The number of premature deaths (age at death between 25 and 74 years) from the different diseases that are considered in the DIETRON model, and the relative risks associated to a change in body weight are provided in Table 19. To compute the impact of the tax on body weight we use the estimated average change in energy intake due to the tax (assuming only beverage consumption is impacted by the tax) for the different categories of people (Table 20). We also use the average height of the French adult population, that is 1.74 m for men and 1.62 m for women. Finally, Table 21 provides the share of the population for each type.

Table 19: **Number of premature deaths (age 25-74) in France (in 2014) and relative risks (RR) associated to a change in BMI**

	Number of deaths		Unit of change of BMI	Relative risk (RR)		
	Men	Women	(kg/m^2)	All BMI	BMI ≤ 25	BMI > 25
Oesophagus cancer	1881	399	1	1.11		
Pancreas cancer	3041	1975	5	1.14		
Colorectum cancer	3908	2538	1	1.03		
Breast cancer(age < 60)	21	2772	2	0.94		
Breast cancer(age ≥ 60)	55	3315	2	1.03		
Endometrial cancer	0	1669	5	1.52		
Kidney cancer	3523	874	5	1.31		
CHD men	6837	-	5		1.27	1.42
CHD women	-	1654	5		1.01	1.35
Stroke	3313	2017	5		0.92	1.39

Source (mortality data): https://www.cepidc.inserm.fr/

Source (Unit of change of BMI and RR): Scarborough et al. (2012)

Table 20: **Variation of sugar intake and energy intake for the different type of population**

Gender	Class of BMI	Variation of intake	
		Sugar (g/year)	Energy (MJ/day)
Men	< 25	939	0,0431
Men	≥ 25	1144	0,0525
Women	< 25	942	0,0432
Women	≥ 25	1237	0,0567

Change in sugar intake taken from Table 8

Table 21: **Share of the French adult population (25-74 years)**

Category			Share (%)
Gender	Age	BMI	
Men	25-59	< 25	17,3
Women	25-59	< 25	22,6
Men	25-59	≥ 25	19,3
Women	25-59	≥ 25	15,0
Men	60-74	< 25	4,1
Women	60-74	< 25	6,4
Men	60-74	≥ 25	8,1
Women	60-74	≥ 25	7,1

Own calculation using data from OBEPI (2012) for obesity rate and INSEE for the population