Let them eat cake? The net welfare impacts of a fat tax

Let Them Eat Cake?
The Net Consumer Welfare Impact of a Fat Tax*

Valeria di Cosmo† Silvia Tiezzi‡

University of Turin University of Siena

April 2024

Abstract

When judging the distributional impact of a sin tax, what matters is not just how much low income people would pay but how much the tax would benefit or harm them overall. In this paper, we assess the consumer welfare impact of a fat tax net of its expected benefits computed as savings from lost weight. We use Italian data to estimate a censored Exact Affine Stone Index (EASI) incomplete demand system for food groups and simulate changes in purchases, calorie intake, consumers' welfare and the monetary value of the tax's short run health benefits. Our results suggest costs from taxation are larger than benefits at all income levels. As a fraction of income, the net impact would be slightly regressively distributed.

Keywords: sin taxes, health benefits, welfare costs, exact affine stone index demand system, demand elasticities, micronutrients intake.

JEL classification: O12, D12, I15.

---

*We thank Daniel Stoehlker and seminar participants at the University of Turin (I); at the 2022 Italian Economic Association Meeting, and at the 2023 Italian Health Economics Association (AIES) Annual Conference for helpful comments. Declaration of interest: None

†Department of Economics and Statistics "Cognetti de Martiis", University of Turin, Lungo Dora Siena, 100 10153 Turin, valeria.dicosmo@unito.it.

‡Department of Economics and Statistics, University of Siena, Piazza San Francesco, 7/8 53100 Siena, silvia.tiezzi@unisi.it.

1 Introduction

Sin taxes, also known as corrective taxes, are imposed in many countries on goods like cigarettes, alcohol, soft drinks, and junk food due to their overconsumption (Cawley et al., 2019; Wright et al., 2017). These taxes are intended to improve social welfare by reducing consumption and alleviating associated healthcare costs.

Taxing unhealthy foods remains a contentious policy due to concerns about effectiveness and negative distributional effects.

Effectiveness relates to the ability of the tax to shape behavior, i.e. to reduce consumption of unhealthy foods or nutrients to a socially optimal level when market failures or inefficient outcomes exist. To this end, whether to tax the product itself or the harmful nutrient responsible for health issues is an important design issue. For instance, in the case of sugar taxes, since the health damage from sugar-sweetened beverages is directly linked to their sugar content, the tax should target sugar directly. This approach aligns the tax with the source of costs, encouraging firms to reformulate ingredients and consumers to switch from high-sugar to low-sugar options.

In fact, sugar-sweetened drink taxes are often designed as a fixed amount per liter of soft drink. For example, Italy plans to implement a volumetric sugar tax on July 1, 2024, after several delays and debates. However, such a tax may not maximize health benefits as consumers lack incentives to switch to low-sugar options, and producers may not reformulate their drinks.

As to the distributional effects, sin taxes tend to be regressive, disproportionately affecting lower-income individuals who consume unhealthy goods at higher rates. However, if lower-income groups reduce their consumption the most following a price increase, sin taxes could potentially benefit disadvantaged social groups more. In such cases, if the welfare gains outweigh the costs, arguments against the regressivity of taxing unhealthy foods lose strength. Assessing the overall welfare and distributional impact of sin taxes requires weighing the costs and benefits, yet evidence on net welfare impacts is limited (Allcott et al., 2019b), with most literature focusing solely on costs or benefits, and benefits rarely measured in monetary terms.

This paper analyses the net welfare effects of a nutrient tax on saturated fat for Italian households. We consider both the monetary value of the tax's health benefits, represented by savings from weight loss, and its associated welfare costs. Our proposed nutrient tax targets foods high in saturated fats as our data highlights excessive saturated fat consumption.

The dataset we assemble for this research is unique in its scope and it is the result of a combination of several datasets. We need household spending information on the entire current consumption bundle to assess how Italian consumers reallocate such bundle following a price change and, therefore, to accurately measure behavioral responses after a price change. Individual product household scanner data are increasingly used for estimating consumer demand models (Dubois et al., 2022), especially for analyzing the demand for specific industrial products or groups of products. Our focus, however, lies in scrutinizing the entire array of consumption goods (including fresh foods and foods not only available through great distribution) to which consumers allocate their available income, as we want to study substitutions and complementarities between food groups in assessing the impact of food tax policies. Such an endeavour would prove very challenging in terms of household scanner data.

Instead, we collect nationally representative pooled cross-sections of Italian households consumption expenditures and associated prices indices. One advantage of household expenditure data lies in its inclusivity, covering purchases from all food retailers, including small retail stores, and transactions involving fresh produce without a barcode. These expenditures and prices, combined with data on food nutrients released by the European Institute of Oncology (IEO), are used to estimate a censored Exact Affine Stone Index (EASI) incomplete demand system (Lewbel & Pendakur, 2009) for 16 food groups that allows us to simulate changes in purchases, in consumer surplus (using the equivalent variation as a money metric measure of consumer surplus variation after a price change), and in weight outcomes after the introduction of a sin tax based on the saturated fat content of foods. We compute the 16×16 matrix of compensated price elasticities. These cross-price elasticities measure pure substitution (or complementarity) net of any income effect. This is an important piece of information for evaluating the effectiveness of our sin tax.

In designing our counterfactual tax simulations, we refrain from pre-determining which nutrient should be taxed, instead focusing on the "bad" nutrient that exceeds WHO guidelines the most in consumption. This approach tailors our simulated sin tax to the specific country, considering prevailing social and cultural norms. Our analysis reveals that a sugar tax might not be the most effective option for Italy if the aim is to address socially costly consumption. We find no evidence of widespread consumption exceeding official recommendations regarding added sugar. Saturated fat emerges as a more viable target for taxation, and we determine the taxation level required to reduce saturated fat consumption to WHO-recommended thresholds.

To compute the short-run tax benefits, we first transform changes in consumption due to the tax into changes in bodyweight (Hall et al., 2011). To transform bodyweight changes into actual monetary benefits, we match our expenditure-price-nutrient database with data from the Italian module of the European Health Interview Survey (EHIS) released by Eurostat, a representative survey on the health and expenditure of Europeans, and use a two-part model to estimate the impact of the weight variation, generated one year after implementation of the tax, on individual monthly health expenditure. This is our money metric of the short-run benefits from the fat tax.

We make the following contributions to the existing literature. First, unlike studies based on household-level purchase data[1], we use a sample of single-household data[2] to ensure a unique correspondence between the recorded expenditure on the different food categories, the costs and the health benefits from taxation. Data on households with more

[1]An exception is Dubois et al. (2020) who study purchase decisions made by individuals for immediate consumption on-the-go. In this study, purchases and consumption are closely aligned.

[2]According to ISTAT, Annuario Statistico Italiano 2019, at January 1^{st} 2019 one member households accounted for the largest share of Italian households: 33%.https://www.istat.it/it/archivio/236772

than one member would not allow to compute health benefits at the individual level.

Second, we assess the monetary value of lost weight associated with the tax by estimating individual health expenditures avoided by weight changes induced by the tax. Universal health coverage is provided by Italy's National Health Service (Servizio sanitario nazionale, or SSN), established in 1978. The SSN automatically covers all citizens and legal foreign residents. Public funding of Italy's SSN accounted for 74.2% of total health spending in 2018, with total expenditure standing at 8.8% of GDP (OECD, 2019). Primary and inpatient care are free at the point of use. Most preventive screenings are also provided free of charge. For medicines, prescribed procedures, and specialist visits patients make co-payments for each prescribed procedure up to a ceiling determined by law. We consider averted health expenditure above the SSN coverage as our proxy for the tax monetary benefits. To accurately measure the marginal effect of weight changes on healthcare costs we follow Cawley & Meyerhoefer (2012) and use a two-part model of medical expenditures (Jones, 2000).

Third, we assess both the costs and benefits of sin taxes in absolute terms and relative to income. Our findings challenge the hypothesis that sin taxes yield net benefits for lower-income individuals. Assuming full pass-through of the tax policy, we observe that a fat tax aimed at reducing saturated fat consumption by 30% results in a marginal net welfare cost for the average Italian consumer. Regarding distributional impacts, when considering net welfare effects relative to total expenditure, low-income individuals incur a proportionally larger net loss from the fat tax compared to high-income individuals, indicating regressive relative net impacts.

Interestingly, we also show that a slight rise in the current value-added tax (VAT) on specific groups of foods yields net welfare and distributional impacts very similar to those of the nutrient tax based on the saturated fat content of foods.

Our research contributes to streams of literature aimed at comprehending the effects of sin taxes. The first encompasses the empirical literature utilizing a demand system approach. Estimating a comprehensive demand system provides an optimal framework for computing theoretically grounded price elasticities fully capturing the behavioral responses of consumers after a price increase. This includes the reallocation of consumer spending across the entire consumption basket following a price change. The associated monetary metric, reflecting the variation in welfare after a price change, incorporates these behavioural reactions.

Chouinard et al. (2007); Zhen et al. (2014); Harkanen et al. (2014); Harding & Lovenheim (2017); Caro et al. (2020) and McCullough et al. (2020) all address the important aspect of substitution between food groups when assessing the impact of food and beverage tax policies by estimating a utility-theoretic demand system. In these studies, consumption is typically measured at the household level, encompassing both adults and children. Consequently, welfare changes induced by the tax are evaluated either at the household level or as per capita averages[3]. Given that health benefits are individual-specific, it's challenging to establish a unique correspondence between purchases, consumption, and the health benefits from taxation for each individual within the household. Consequently, studies in this literature focus on welfare costs and/or health benefits at the household level, with health benefits typically not expressed in monetary units.

We depart from these studies by utilizing single-household data, aiming to align purchases, consumption, welfare costs, and welfare benefits at the individual level. Additionally, we estimate health benefits in monetary terms.

Our study also intersects with the body of literature focusing on the money metric estimation of health benefits derived from sin taxes. Recent studies conducted in high-income countries such as Australia, Canada, and the USA have reported equal or greater health benefits in monetary terms for lower-income groups (Kao et al., 2020; Lal et al., 2017; Wilde et al., 2019). These studies calculate the monetary values of long-run health benefits as savings in healthcare expenditures based on the predicted reduction in mortality and morbidity from diseases associated with overconsumption of unhealthy nutrient

[3]It's worth noting an exception in Xiang et al. (2018), where welfare costs of a SSB tax for different household types, including single households, are estimated. However, the demand system employed in this study is highly aggregated and does not allow for substitutions between food groups.

targeted for taxation. These studies typically assume zero substitution between the group of goods targeted for taxation and other food or beverage groups. Furthermore, the elasticities employed to estimate changes in weight and BMI driving the predicted reduction in mortality or morbidity are often imported from external sources. Consequently, there could be some misalignment between consumer costs and benefits.

We diverge from this literature in two key ways. Firstly, we estimate both own- and cross-price elasticities for our sample of individuals, and the change in consumer surplus associated with the tax accounts for complementarities and substitutions among all food groups resulting from the tax's introduction. We measure pure substitutions (or complementarities) by computing compensated cross-price elasticities, net of any income effect. Secondly, for each individual in our sample, we approximate the short-run tax benefits by estimating the expected savings in out-of-pocket healthcare costs resulting from predicted weight changes one year after the tax's introduction.

Finally, our study connects to very few papers linking consumer costs and benefits to assess the full impact of sin taxes (Allcott et al., 2019a; Dubois et al., 2020). Allcott et al. (2019a) are the first to provide a tractable theoretical and empirical framework accounting for the three key elements for evaluating welfare benefits from a sugar tax in the US: correcting consumer bias, externalities, and revenue recycling through income transfers. Their results suggest positive and slightly regressive net gains. Dubois et al. (2020) use the estimates of internalities of Allcott et al. (2019a) and, under lump-sum redistribution, find that a sugar tax in the UK would be only mildly regressive. Instead, our results suggest short-run costs from taxation larger than short-run benefits at all income levels. As a fraction of income, the net impact is slightly regressively distributed.

The rest of the paper is structured as follows. Section 2 presents the data sources used in our analysis. Section 3 explains the demand model, estimation procedure, and derived elasticities. In Section 4, we investigate the welfare costs and distributional implications of our sin tax simulations. Section 5 discusses the monetary value of short-run health benefits and assesses the net consumer welfare and distributional impact of the simulated fat tax.

Section 6 concludes.

2 Data

2.1 Expenditures

We aim to include food expenditures from all food retailers, including small local shops which still contribute to total food spending in Italy as Italians purchase fresh produce not only from supermarkets but also from small shops, markets, bakeries, butchers, and other establishments (Cozzi, 2008).

To ensure such comprehensive coverage, we utilize 5 independent cross-sections of micro-data on food consumption expenditure from the Household Budget Survey (HBS), administered by the Italian National Institute of Statistics (ISTAT), spanning from January 2014 to December 2018[4]. The HBS data encompasses all food expenditures, irrespective of the retailer type, aligning seamlessly with the level of aggregation in the nutrients dataset, which is crucial for our analysis. Each annual cross-section contains monthly consumption expenditures from approximately 23,000 Italian households across about 480 municipalities[5]. Additionally, the survey provides detailed insights into household structure and sociodemographic characteristics, including regional location, household size, gender, age, education, and employment status of each member [6].

The HBS provides expenditure data at the household level. Given our aim to align welfare costs and the monetary value of health benefits from the tax at the individual level, it's essential to establish a clear link between each individual and recorded expenditures

[4]https://www.istat.it/it/archivio/180341

[5]ISTAT employs a weekly diary to gather expenditure data on frequently purchased items and conducts face-to-face interviews for data on significant and durable expenditures. Each month, two weeks are randomly selected, and households are equally divided into two groups, assigned to one of the two selected weeks. Expenditures are classified into roughly 280 elementary goods and services, with slight variations in the item list from year to year.

[6]All annual samples are drawn independently according to a two-stage design. Details on the sampling procedure used to collect data in the first year of this survey can be found in: ISTAT File Standard-Indagine sui Consumi delle Famiglie-Manuale d'uso, anno 2014. Downloadable at *http : //www.istat.it/it/archivio/*4021.

on each food category. We thus focus on households with only one member. However, recognizing that households with children may have different characteristics and responses to a fat tax compared to those with a single adult member, and for the sake of comparison, we also estimate our demand model and compute elasticities and the tax's welfare costs for households with two adults and a child aged fourteen or younger.

Our final sample of single adults includes a total of 12,369 individuals classified into 21 regions and three urban types (metropolitan areas, medium-size cities and small cities). The complementary sample of two adults and one child includes 4,772 households. The household food consumption module assembles data on expenditure for about 200 items based on a seven-day recall. We aggregate the food items into 16 food-at-home groups and one food away-from-home (food_afh) item for a total of 17 aggregates based on the typical composition of Italian meals and the nutritional characteristics of foods: alcoholic drinks; bread and pasta; cereals and rice; eggs and milk; fat and cheese; fish; food_afh; fruit; oil; drinks other than sweetened or alcoholic beverages; processed meat; poultry; red meat; sugar-sweetened beverages; snacks and sweets; vegetables; other. The latter category is used to define a composite numeraire good which, combined with residual food items, includes all non-food current consumption expenditures. We use this aggregate as a numeraire in our incomplete demand system (LaFrance & Hanemann, 1989; Hanemann & Morey, 1992). The HBS also provides data on household non-food expenditure, which we use to calculate household total current consumption expenditure (i.e. expenditure on food and non-food items) and budget shares on a monthly basis. For both household types, we report descriptive statistics for the average budget share of each food group and for log prices in Table B1, Appendix B[7].

[7]One limitation of this data is that we do not know the exact composition of food_afh, which accounts for 22% of total food consumption in our sample, and includes food and drink from bars, restaurants and on-the-go (e.g. purchases from vending machines and food stalls). We therefore cannot calculate bad nutrients consumption from these sources. We can, however, disaggregate the budget share of food_afh into its three largest categories: on-the-go (0.07); bars and pastry shops (0.006); restaurants and taverns (0.145). This shows that, although consumption on-the-go might be an important segment of food away from home especially among children and adolescents (Dubois et al., 2020) and high in bad nutrients, in our sample of adults consumption on-the-go and from bars and pastry shops does not cover more than 7.6% of total food and drink expenditure. We also show in Section 3.2 that although an increase in the

2.2 Prices

Since the HBS does not provide information on prices paid by consumers, we use monthly consumer price indices (equal to 100 in 2015) from January 2014 to December 2018, also supplied by ISTAT. These disaggregated price indices are the inputs used to build the overall Harmonised Consumer Price Index (HCPI) compiled by Eurostat to monitor inflation in Europe. We need to associate each expenditure category in the HBS with its own price index. Aggregation of the items in the HBS is constrained by the HCPI breakdown by type of good, according to COICOP (Classification of Individual Consumption by Purpose) developed by the United Nations Statistics Division to classify and analyze individual consumption expenditure incurred by households. To aggregate expenditure items in the HBS we conform to the COICOP using 5 digits. This provides a very granular disaggregation of prices that matches our selected HBS expenditure categories. One concern is the lack of cross-sectional variation and high collinearity of prices, which often occurs in estimations of highly disaggregated demand systems on pooled cross-sectional data. We address this concern by computing Stone-Lewbel prices (Lewbel, 1989) for the food groups in our demand system. With the assumption of constant expenditure shares within a group, the prices of individual goods within each food group are weighted with their expenditure shares in the food group. Since these shares vary for every household in our sample, the Stone-Lewbel procedure adds cross-sectional variability to our price indices [8].

2.3 Nutrients

The nutrients (calories, fats and sugar) content of our food groups is calculated by applying conversion factors from the 2015 edition of the Food Composition Database for Epidemiological Studies in Italy (Banca Dati di Composizione degli Alimenti per Studi Epidemiologici in Italia) released by the European Institute of Oncology (EIO)[9] which

price of fat and cheese, sweets and snacks, or sweetened beverages causes substitution towards food away from home, such substitutions are small in magnitude.

[8]Figure B.1 in Appendix B shows the monthly time series of price indices for our 16 food groups.

[9]http://www.bda-ieo.it/wordpress/en/

allows us to calculate nutrient values per kilogram of each food group[10]. Table 1 shows sugar, saturated fats and calories content per kg of final product. As expected, sugar is exceptionally high for sweets and snacks and sweetened beverages. Saturated fat is high in fat and cheese, oil, and processed meat.

Table 1: Nutrients (grams) and Kcal per kg by food group

Food category	N	Sugar	Saturated fats	Kcal
Vegetables	10,638	0.0	1.4	647
Fruit	10,638	30.6	18.0	1762
Pasta and Bread	10,638	0.0	5.2	2899
Cereals and Rice	10,638	0.0	4.8	2938
Eggs and Milk	10,638	12.5	34.9	1710
Fish	10,638	3.3	11.3	1228
Poultry	10,638	0.1	25.4	1675
Red Meat	10,638	0.3	29.5	1515
Processed Meat	10,638	2.5	71.4	2788
Fat and Cheese	10,638	0.0	160.5	3453
Oil	10,638	1.4	216.6	8660
Sweets and Snacks	10,638	113.7	30.6	3056
Sweetened beverages	10,638	37.6	0.1	414
Other drinks	10,638	2.8	9.8	860

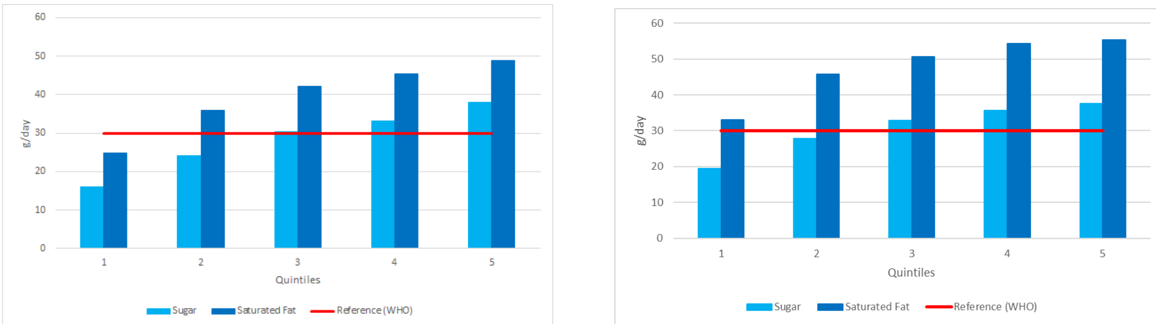
We take 30 g/day for sugar and saturated fat intakes as a reference value, as recommended by the WHO (WHO (2018), and WHO (2015)), and we compute overconsumption as the difference between the average sugar (excluding sugar from fruit) and saturated fats daily intake of the individuals in our sample and the reference value (Griffith et al. (2016)).

[10] The calorie, fat and sugar contents of food_afh are not calculated. Without information on the detailed food items of each meal purchased, the nutrient components for food_afh cannot be assessed. However, the estimated elasticities reveal small complementarity or substitution effects between food_afh and the other food groups. We therefore presume that ignoring nutrients in food_afh has little impact on our findings.

Figure 1: Sugar and saturated fats overconsumption by income quintile

(a) Single adults

(b) Two adults & one child



Notes: Average daily intake (grams/day) of sugar and saturated fats for single adults (a) and for equivalent adult (b) across quintiles of total expenditure. Horizontal lines represent the threshold recommended by the WHO (30 grams/day).

Our data shows only a slight excess consumption of sugar among single adults. Perhaps surprisingly, a similar trend is observed in households with two adults and one child, even though about 20% of Italian children were overweight before the Covid19 crisis (WHO, 2022). Results of recent research (Spinelli et al., 2023) show that the percentage of children in Italy who consumed unhealthy foods such as sugary drinks more than three days a week was lower than the global average and that children's dietary habits are strongly influenced by multiple interacting factors related to shared environments and norms, attitudes, beliefs and behaviours of people living in the same place or with a similar background (Crudu et al., 2021). These findings cast doubts on the appropriateness of food taxes to address child obesity in Italy implying that a more comprehensive approach considering complex interacting factors may be necessary[11].

In particular, Figure 1 shows that there is no overconsumption of sugar for individuals below the fourth quintile of the income distribution (Figure 1 (a)) and below the third quintile for equivalent adult in households with two adults and one child (Figure 1 (b)). The sugar consumption sample average (28 g/day) is below the threshold recommended by the WHO (30 g/day) for single adults (Figure 1 (a)) and slightly above the WHO

[11] Fletcher et al. (2010) investigate the potential for soft drink taxes in the US to combat rising levels of child and adolescent obesity through a reduction in consumption. Their results suggest that soft drink taxation leads to a moderate reduction in soft drink consumption by children and adolescents, but that reduction is completely offset by increases in consumption of other high-calorie drinks.

threshold (30.67 g/day) in households with two adults and one child (Figure 1 (b)). This suggests that the volumetric sugar tax to be implemented in Italy starting July 1 2024 would certainly raise cash, but is likely to be of little use for reducing the amount of sugar per can or for protecting health.

By contrast, the average intake of saturated fats is 42 g/day for single adults (Figure 1 (a)) and 47.8 g/day per adult equivalent in households with two adults and one child (Figure 1 (b)), exceeding the 30 g/day threshold recommended by the WHO by about 33%. Excess consumption increases with income, with individuals in quintiles above the first displaying larger consumption excesses. This data suggests that in Italy the usual picture of lower income individuals showing higher consumption of unhealthy nutrients than higher income people is reversed. Very similar patterns hold for other household types. Figure A1 in Appendix A shows the distribution of consumption of harmful nutrients in households with two adults and two adults and two children. Figures 1 and A1 seem to suggest that families in higher quintiles of the income distribution consume more unhealthy nutrients than those in lower quintiles even though we cannot assess the intra-household allocation of consumption.

2.4 Health expenditures and health related variables

Data on individual body weight and other health related variables in our sample comes from the 2015 Italian module of the European Health Interview Survey (EHIS), a survey on the health of the population of EU member states conducted every four years. To this, we add variables measuring the relative risk (RR) of developing type 2 diabetes for overweight and obese populations and the relative risk (RR) of developing cardiovascular diseases (CVD) (Abdullah et al., 2010; Bogers et al., 2007; Park et al., 2017). To match health expenditure from HBS data with individual body weight and the other health related variables, we apply the matching method developed by Rubin (1986) and Moriarity & Scheuren (2003). The two-step matching procedure is detailed in Appendix E.

3 The Demand Model

We estimate an incomplete Exact Affine Stone Index (EASI) implicit Marshallian demand system (Lewbel & Pendakur, 2009) including 16 food groups and a composite numéraire that incorporates all other consumption goods and services plus a residual food category[12]. The demand system's estimated parameters can be used to provide exact measures of changes in welfare, unlike those of conditional demand models (LaFrance & Hanemann, 1989; Hanemann & Morey, 1992). Conditional demand systems underestimate the degree of substitution among expenditure groups after a price change (Zhen et al., 2014), because weak separability between food expenditure and that of all other consumption implies that only substitutions among food groups are taken into account. An incomplete demand system, on the other hand, produces unconditional predictions of demand responses to a simulated price change.

One potential problem in estimating a demand system with household level data is the existence of zero observations due to infrequent purchase of highly disaggregated food categories. We adopt Shonkwiler & Yen (1999) two-step estimation procedure to address this issue. After modifying the EASI incomplete demand system to account for censoring, the implicit Marshallian budget shares equations to be estimated are:

$$w^j = \Phi(v^{'}\lambda^j)\left[\sum_{r=1}^{R} b_r^j (y)^r + \sum_{t=1}^{T} g_t^j z_t + \sum_{k=1}^{J} a^{jk} lnp^k\right] + \tau^j \phi(v^{'}\lambda^j) + \varepsilon^j \tag{1}$$

$$(y)^r = \left(lnx - \sum_{j=1}^{J} w^j lnp^j + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a^{jk} lnp^j lnp^k\right) \tag{2}$$

where w^j is the budget share of commodity j; J is the number of goods with the J^{th} good

[12]The EASI demand system has several additional benefits with respect to the popular Quadratic Almost Ideal Demand (QAID) system (Banks et al., 1997). First, it makes it possible to specify and test for Engel curves that are more flexible than quadratic ones. This is an important characteristic when estimating a highly disaggregated demand system such as the present, and it may have an impact on price coefficient estimates. Second, the EASI error term can be interpreted as unobserved consumer heterogeneity that is seldom explained by observed demographic and price changes alone. These unobserved preference heterogeneity parameters show up both in the budget-share and cost functions, and are therefore relevant factors for predicting demand and assessing welfare variations after a price change.

being the composite numéraire; y is real household income; R is the highest order of the polynomial in y to be determined empirically; p^k is the price index of the k^{th} good; T is the number of exogenous demand shifters; z_t is the t^{th} demand shifter; b_r^j, g_t^j and a^{jk} are parameters to be estimated; and ε^j is the error term. Denoting the vector of predictors of positive consumption and the vector of their associated parameters by v and λ for equation j, $\Phi(v'\lambda^j)$ and $\phi(v'\lambda^j)$ are the normal cumulative distribution and probability density functions, respectively, related to the first-stage probit equations introduced to correct the bias in the coefficients of the EASI model caused by censoring. Finally, x in (2) is nominal total consumption expenditure.

To ensure integrability of the demand equations we impose the theoretical restrictions of homogeneity: $\sum_{k=1}^{K} a^{jk} = 0$ for all $j = 1, ..., J$; symmetry: $a^{jk} = a^{kj}$; and adding up. Adding up requires that the sum of the J coefficients associated with the constant of each share equation (denoted z_0) is equal to one: $\sum_{j=1}^{J} g_0^1 = 1$; and that the sum of the J coefficients associated with any other variable in the budget shares equations is equal to zero: $\sum_{j=1}^{J} a^{jk} = 0, \quad k = 1, ..., J$; $\sum_{j=1}^{J} b_r^j = 0, \quad r = 1, ..., R$; $\sum_{j=1}^{J} g_t^j = 0, \quad t = 1, ..., T$.

The EASI demand system is nonlinear and endogenous. Nonlinearity arises from the fact that b_r multiplies a power of y. Endogeneity is due to the budget-shares $w^j, j = 1, ..., J$ being on both sides of the system of equations. Estimation is further complicated by the presence of censoring. However, like the QAID, the EASI demand system can be approximated using linear-in-the-parameters equations. The approximated model replaces y with $\tilde{y} = lnx - \sum_{j=1}^{J} w^j lnp^j$, where $\tilde{y}$ is the log nominal expenditures deflated by the Stone price-index[13]. To correct for endogeneity due to the introduction of budget shares into log real total expenditure we create an instrument for y constructed as $logx$ deflated by a modified Stone price index where $\bar{w}^j$, the sample-average budget share for food group j, replaces w^j (Lewbel & Pendakur, 2009): $\hat{y} = lnx - \sum_{j=1}^{J} \bar{w}^j lnp^j$.

In addition to a constant, we specify the vector of demand shifters z_k to include the fol-

[13]Lewbel & Pendakur (2009) show that the linearized version of the model estimated by OLS performs almost as well as fully-efficient nonlinear estimation.

lowing variables: a dummy for gender (1= male); the level of education in 5 classes (1= no formal education, 2= primary school, 3= lower middle school, 4= high school diploma, 5= undergraduate or postgraduate degree); marital status in 6 classes (1= single, 2=married, 3= married but not co-habiting, 4= legally separated, 5= divorced, 6 = widowed); employment status in 7 classes (1= employed, 2= in search of first employment, 3=unemployed, 4= student, 5= housewife, 7= other employment position, 8= retired); age in 9 classes (1= between 18 and 24 years, 2= between 25 and 29 years, 3= between 30 and 34 years, 4 = between 35 and 39 years, 5= between 40 and 44 years, 6=between 45 and 49 years, 7 = between 50 and 54 years, 8 = between 55 and 59 years, 9= between 60 and 64 years); three Census regions (metropolitan area, medium size city, small town). Descriptive statistics for these demand shifters are shown in Table B2 in Appendix B[14].

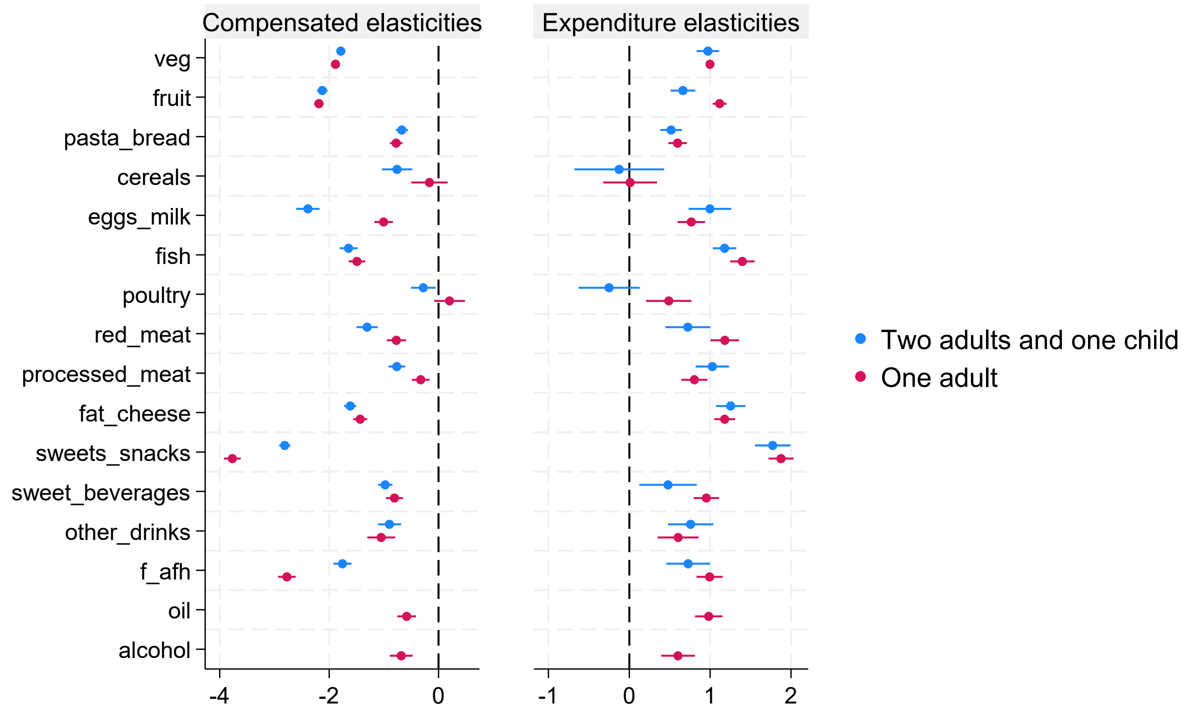
3.1 Elasticities

Behavioral reactions after a price change are measured by own and cross-price elasticities. The latter, in particular, highlight substitutions and complementarities among food groups, i.e. changes in the purchased quantities of other food groups after a price change[15]. Our structural model estimates lead to a 16×16 matrix of 256 estimated price elasticities. Figure 2 focuses on the main diagonal of the matrix of compensated elasticities (left) and on expenditure elasticities (right) for single adults (red dots) and for households with two adults and one child (blue dots). All estimated own-price elasticities have the expected negative sign and most are statistically significant at 1% in both samples. Between households, the largest differences in compensated elasticities are for eggs and milk, sweets and snacks, and food_afh with families with children being more reactive to a change in the price of eggs and milk and less reactive to a change in the price of sweets and snacks and food_afh compared to single adults.

[14]We estimated the model twice: once for the sample of single adults and once for the sample of two adults and one child. Estimated coefficients for both models are available from the authors upon request.

[15]Appendix D shows the equations for the Marshallian price elasticities of quantities, the Marshallian expenditure elasticities, and the Hicksian price elasticities of quantities derived from the EASI demand model. Standard errors of the elasticities are bootstrapped by 200 replications.

Figure 2: Own-price compensated and expenditure elasticities



Notes: Compensated own-price elasticities (left) for each food group for single households (red) and households with two adults and one child (blue). Expenditure elasticities (right) for each food group for single households (red) and households with two adults and one child (blue). For the latter, elasticities for oil and alcohol are not reported as the Lewbel procedure did not converge due to the small number of observations. Bootstrapped standard errors with 200 replications.

Among the food groups, sweets and snacks show the largest own-price elasticity (-3.770 for single adults), implying that a 1% increase in their price would decrease the quantity purchased by about 3.8%. The purchased quantity of fat and cheese is also elastic to its price (-1.432 for single adults). The own-price elasticity for sweetened beverages is -0.805 for single adults, which falls in the range -0.8 to -0.10 of the literature (Finkelstein et al., 2010)[16]. Most expenditure elasticities are positive and significant at 1% and most food groups are necessities with an expenditure elasticity of less than one. Fish, fat and cheese, sweets and snacks are luxuries in both samples with an expenditure elasticity greater than one.

Since we estimate welfare costs and benefits for single adults, we next focus on substi-

[16]Studies such as Allcott et al. (2019a), focusing only on soft drinks, find higher own-price elasticities (-1.37).

tution and complementarities among food groups for single adults only. Table 2 shows the full set of compensated and expenditure (last row) elasticities at the sample mean[17]. Since compensated elasticities are utility-constant, cross-price elasticities measure pure substitutions after a price change net of any income effect. The cells of each row show the price elasticity of the food group of the row due to a change in price of the food group of the column. For example, the third entry in the first column (0.228) is the percentage change in the demand for pasta and bread following a 1% increase in the price of vegetables. Positive and significant cross-price elasticities indicate substitutions, while negative and significant ones indicate complementarities. We are particularly interested in complementarities and substitutions with fat and cheese, sweets and snacks and sweetened beverages. Increasing the price of sweets and snacks causes substitutions with vegetables (0.202), alcohol (0.199), and food_afh (0.181), and complementarities with eggs and milk (-0.139), cereals (-0.411), sweetened beverages (-0.462), and other drinks (-0.182). Increasing the price of sweetened beverages causes substitution with fruit, fish, food_afh (0.096, 0.019, 0.023) and complementarities with oil, sweets and snacks, and other drinks (-0.160, -0.112, -0.160). Raising the price of fat and cheese causes substitution with bread and pasta (0.091) and complementarity with fruit (-0.064)[18]. Table D2 in Appendix D shows, for single adults only, compensated own-price and expenditure elasticities at low and high total expenditure, our proxy for income. As expected, low-income individuals react more than high-income persons after a price change in most food groups. However, higher-income individuals, whose consumption of fat and cheese is larger, exhibit greater own-price elasticity for these items (-1.633 compared to -1.088 for low-income individuals). This challenges the narrative that poorer individuals consume more unhealthy nutrients and are more responsive to price increases.

[17]Table D.1 in Appendix D shows the full set of uncompensated (Marshallian) elasticity point estimates. The full set of standard errors is available from the authors.

[18]One concern is the substitution of food categories higher in fats and sugar (fat and cheese, sweets and snacks, and sweetened beverages) with food_afh, because we cannot assess the bad nutrients content of the latter. Inspection of Table 2 shows a small substitution of fat and cheese, sweets and snacks, and sweetened beverages with food_afh (0.032, 0.181, and 0.023, respectively).

Table 2: Compensated price elasticities and expenditure elasticities (single adults) - sample means.

	Vegetables	Fruit	Pasta and bread	Cereals and Rice	Eggs and Milk	Fish	Poultry	Red meat	Processed meat	Fat and Cheese	Oil	Sweets and snacks	Sweetened beverages	Other drinks	Alcohol	Fafh
Vegetables	**-1.883*** **	-0.002	0.133***	0.000	0.092***	-0.131***	0.005	0.087***	0.055**	0.007	-0.030*	0.202***	-0.016*	0.020	-0.070***	0.117***
Fruit	-0.003	**-2.186*** **	0.032	-0.104***	-0.002	-0.028	0.010	0.060	0.120***	-0.064*	-0.016	0.044	0.096***	0.033	0.157***	0.348***
Pasta and Bread	0.228***	0.037	**-0.773*** **	-0.034	-0.004	-0.069	-0.098*	0.011	-0.239***	0.091**	-0.004	-0.045	0.007	-0.013	-0.044	0.162***
Cereals and Rice	-0.002	-0.558***	-0.155	**-0.165**	0.114	-0.076	-0.087	-0.256	0.207	-0.017	-0.190*	-0.411**	-0.066	-0.009	0.388*	0.396**
Eggs and Milk	0.208***	-0.003	-0.005	0.033	**-1.003*** **	-0.131**	0.054	-0.226***	-0.290***	0.066	-0.037	-0.139**	0.006	0.022	0.039	0.185**
Fish	-0.236***	-0.034	-0.072	-0.017	-0.104**	**-1.491*** **	-0.006	-0.036	0.066	0.020	-0.042	0.074	0.049*	-0.045	-0.072	0.161*
Poultry	0.013	0.019	-0.164*	-0.031	0.068	-0.009	**0.202**	-0.609***	-0.086	-0.139	-0.245***	0.161*	-0.050	0.027	-0.101	-0.060
Red Meat	0.147***	0.068	0.011	-0.055	-0.168***	-0.034	-0.359***	**-0.770*** **	-0.162**	-0.068	0.052	-0.008	-0.018	-0.103*	-0.162**	-0.029
Processed meat	0.101**	0.149***	-0.257***	0.048	-0.235***	0.068	-0.055	-0.177**	**-0.324*** **	-0.098	-0.054	-0.062	-0.011	-0.233***	-0.090	0.045
Fat and Cheese	0.013	-0.079*	0.098**	-0.004	0.053	0.021	-0.089	-0.074	-0.098	**-1.432*** **	0.041	-0.044	-0.011	-0.069	-0.036	0.134*
Oil	-0.152*	-0.056	-0.010	-0.121*	-0.081	-0.118	-0.430***	0.154	-0.149	0.113	**-0.582*** **	-0.056	-0.160***	0.164*	-0.050	0.208*
Sweets and snacks	0.275***	0.041	-0.036	-0.071**	-0.083**	0.056	0.077*	-0.006	-0.046	-0.033	-0.015	**-3.770*** **	-0.112***	-0.107**	0.122**	0.555***
Sweetened beverages	-0.090*	0.364***	0.024	-0.047	0.016	0.153*	-0.099	-0.061	-0.034	-0.034	-0.179***	-0.462***	**-0.805*** **	-0.388***	0.045	0.293***
Other drinks	0.046	0.052	-0.017	-0.003	0.023	-0.057	0.022	-0.141*	-0.293***	-0.086	0.076*	-0.182**	-0.160***	**-1.048*** **	-0.053	0.192*
Alcohol	-0.154***	0.235***	-0.057	0.109*	0.038	-0.088	-0.079	-0.213**	-0.108	-0.044	-0.022	0.199**	0.018	-0.050	**-0.680*** **	-0.377***
Fafh	0.052***	0.104***	0.042***	0.022**	0.036**	0.040*	-0.009	-0.008	0.011	0.032*	0.018*	0.181***	0.023***	0.037*	-0.076***	**-2.772*** **
Expenditures	**0.998*** **	**1.117*** **	**0.597*** **	**0.010**	**0.767*** **	**1.398*** **	**0.488*** **	**1.182*** **	**0.805*** **	**1.180*** **	**0.982*** **	**1.877*** **	**0.953*** **	**0.604*** **	**0.602*** **	**0.994*** **

$* = p < 0.10$; $** = p < 0.05$; $*** = p < 0.01$. Standard errors were bootstrapped with 200 replications.

4 Counterfactuals

In the main counterfactual experiment we use our demand estimates for single adults to simulate the introduction of a specific (s) tax (τ) proportional to the saturated fat content of a food group. Let η^j denote the saturated fat content of one kg of food group j. We assume that the post-tax price of commodity j, $p^j_{1,s}$, is related to pre-tax price, p^j_0, according to:

$$p^j_{1,s} = p^j_0 + \tau\eta^j \tag{3}$$

As explained in Section 2 we detect in our sample an average excess consumption of saturated fat of about 30%. We therefore select the rate of tax that results in a 30% decrease in saturated fat purchased assuming a 100% pass-through of taxes to prices[19]. For each commodity (i.e. food group) $j, j = 1, ..., J$, the specific tax on saturated fat is:

$$\tau\eta^j = \frac{-0.30}{\epsilon^j} p^j_0 \eta^j \tag{4}$$

where ϵ^j is the own-price compensated elasticity of quantity for commodity j.

We also separately simulate an easy-to-implement and administer increase in the existing Value Added Tax (VAT) on the food groups richest in saturated fat: fat and cheese, processed meat and sweets and snacks[20]. The results of this additional counterfactual experiment are shown in Appendix F. Here we focus on the effects of the specific fat tax.

Table 3 shows the vector of percentage price variation after the introduction of the specific fat tax.

[19]Griffith et al. (2019) review the pass-through of soft-drink taxes to prices finding that a 100% pass-through is the most common finding. Dubois et al. (2020), study the on-the-go segment of the UK market and add to the previous evidence suggesting a soda tax pass-through close to 100%

[20]Current VAT on food products in Italy is 4% for necessities (vegetables, fruit, bread and pasta, fat and cheese and oil) and at 10% for non-necessities (cereals and rice, meat, fish, sweet and snacks, sweetened beverages and other beverages).

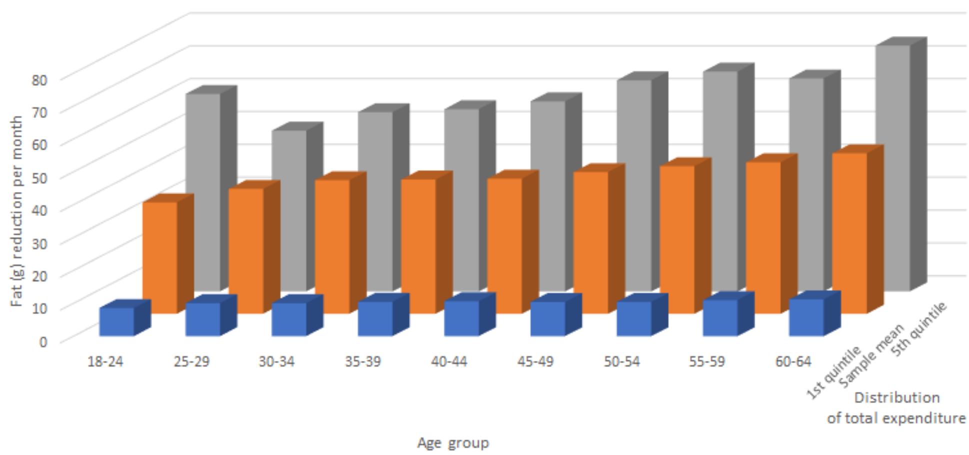
Table 3: Percentage price variation under specific fat tax

Food groups	**Price variation**
Vegetables	0.021
Fruit	0.245
Pasta and Bread	0.201
Cereals and Rice	0.000
Eggs and Milk	1.036
Fish	0.224
Poultry	0.000
Red Meat	1.124
Processed Meat	6.409
Fat and cheese	3.328
Oil	11.070
Sweets and Snacks	0.241
Sweetened beverages	0.003
Other drinks	0.278
Alcohol	0.000
Food_afh	0.000

The effectiveness of a fat tax can be evaluated by how much consumers decrease their fat consumption after the tax. Figure 3 shows the variation in saturated fat consumption (grams) per month after the tax across the distributions of age and total monthly expenditure (our proxy for income). The age groups are 1=18-24, 2=25-29, 3=30-34, 4=35-39, 5=40-44, 6=45-49, 7=50-54, 8=55-59 and 9=60-64 years. The fat tax achieves relatively large reductions in fat consumption among individuals with average (orange columns) and high (grey columns) total expenditure but is not successful at targeting individuals in the lowest quintile (blue columns) of the expenditure distribution. High-income individuals are the most likely to be fat consumers (and are therefore affected by the tax), and they show the largest reductions in saturated fat consumption[21]. Across the age distribution, young consumers are equally likely to be affected by the policy as adults.

[21]In the context of a sugar tax in Catalonia, Fichera et al. (2021) also find that the sugar tax has a stronger impact on wealthier consumers.

Figure 3: Reduction in saturated fat consumption (grams)



Notes: This figure shows the variation in grams of saturated fat consumption per month after the introduction of the fat tax across the distribution of age in nine classes and of total expenditure, our proxy for disposable income.

4.1 Consumer-Welfare Costs and Redistribution

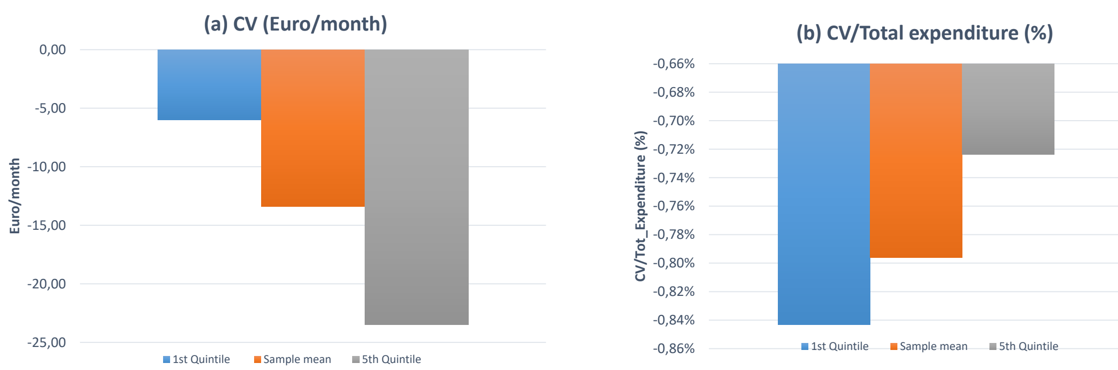
We use our demand estimates to compute the compensating variation (CV), a money metric measure of welfare change after a price change, defined as the minimum sum of money necessary to fully compensate a consumer after the price change. If w_0 is the baseline level of the welfare before any price change, CV is the sum of money necessary to render an individual indifferent to the change in tax policy: $CV = c(w_0, \mathbf{p}_1) - c(w_0, \mathbf{p}_0)$ where $c(w_0, \mathbf{p}_0)$ is the minimum cost of achieving w_0 at prices $\mathbf{p}_0$, and $c(w_0, \mathbf{p}_1)$ is the minimum cost of attaining utility w_0 at the price vector $\mathbf{p}_1$. To calculate the CV, we use the True Cost of Living (TCOL) index (Deaton & Muellbauer, 1980), the ratio of the cost of achieving a given level of economic welfare after a price change to the cost of achieving the same level of economic welfare before the price change: $TCOL = \frac{c(w_0,\mathbf{p}_1)}{c(w_0,\mathbf{p}_0)}$. The CV and the TCOL are clearly related to each other: $CV = c(w_0, \mathbf{p}_0) \times (TCOL - 1)$.

The EASI log change in the TCOL index (Lewbel & Pendakur, 2009) is calculated as:

$$ln\left(\frac{x_1}{x_0}\right) = (\mathbf{p}_1 - \mathbf{p}_0)'w_0 + 0.5(\mathbf{p}_1 - \mathbf{p}_0)'\Gamma(\mathbf{p}_1 - \mathbf{p}_0) \tag{5}$$

where x_1 is the post-tax income necessary to maintain utility at the pre-tax level; $\mathbf{p}_1$ is the $J \times 1$ vector of new log prices after the tax is imposed, and Γ is a $J \times J$ matrix of parameters whose element Γ_{ij} equals a^{jk} in equation 1. Equation 5 captures two effects of the fat tax on welfare. The first term on the right-hand-side is the Stone price effect that ignores any changes in budget shares of the taxed goods. The second term measures the effect of changing budget shares as a consequence of substitution. The total effect will be smaller than the Stone price effect if budget shares of the taxed goods decrease in response to the tax. Figure 4 illustrates the consumer-welfare effects of the specific fat tax for single adults. The welfare loss increases with income. At mean income, CV is 13.40 €per month. Relative to income, proxied by total monthly consumption expenditure, the welfare loss is regressively distributed[22].

Figure 4: Compensating Variation (CV)



Notes: figure (a) shows welfare costs in €/month across the distribution of total expenditure, our proxy for disposable income. Welfare costs are measured by the compensating variation (€/month) after a price change. Figure (b) shows the distribution of welfare costs as a fraction of total expenditure.

[22]Results for two adults and one child are not reported to save space. They do not markedly differ from those depicted in Figure 4. The compensating variation exhibits less variability across total expenditure compared to single adults, yet both absolute and relative CV patterns remain similar. At the mean income, the absolute CV stands at 11.53€per month and 18.49€per month for the fifth quintile. As a percentage of total expenditure, the CV shows less variability compared to the same metric calculated for singles, with values of 0.47%, 0.43%, and 0.38% for the first, mean, and fifth quintiles respectively.

5 Monetary value of weight loss

One potential consequence of excess saturated fat consumption is weight gain. We proxy the value of the short run tax benefits with the value of health benefits associated with weight loss (Hall et al., 2011; Lin et al., 2011; Harkanen et al., 2014; Xiang et al., 2018). To do this, we first calculate the impact of the fat tax on food consumption by multiplying the matrix of uncompensated price elasticities (Table D1 in Appendix D) by a vector containing the percentage changes in consumer prices. Table 4 (left) shows these relative demand changes, computed as $\frac{(q_1^j - q_0^j)}{q_0^j} = \epsilon_j \times \frac{(p_1^j - p_0^j)}{p_0^j}$, for each food group j. We then compute the variation in harmful nutrient intake after the introduction of the tax. This is shown in Figure 3 as a function of age and across the distribution of total expenditure (our proxy for disposable income). As explained before, the tax achieves relatively large fat reductions among those individuals with an average and high level of total expenditure, but it is not successful at targeting individuals in the lowest quintile of the expenditure distribution. We therefore expect health benefits to be progressively distributed, i.e. larger at higher incomes.

To calculate individual weight change in response to reduction in fat consumption, we adopt the approximate rule of thumb proposed by Hall et al. (2011) for an average overweight adult, based on dynamic simulation models predicting individual weight changes resulting from energy balance interventions: every 100 kJ/day change in energy intake will lead to a bodyweight change of about 1 kg (or 10 kcal/day per pound of weight change) with half the weight change achieved in about 1 year and 95% in about 3 years[23]. Table 4 (right) shows the average reduction in body weight (in kg) and the average change in energy intake (kJ/day) one year after the introduction of the fat tax. We obtain an average body-weight loss of 1.72 kg one year after the introduction of the tax.

To translate body-weight variation into monetary benefits we use a two-part model (2PM) of monthly health expenditures at the individual level (Jones, 2000), as adopted by

[23] We also computed the effect of changes in energy intake on body weight using the approach proposed by Dall et al. (2009) and applied in Harkanen et al. (2014). We obtained slightly larger bodyweight changes. The results are available from the authors upon request.

Table 4: Changes in quantities purchased (left); changes in body weight (kg) and daily energy intake (kJ) one year after imposition of the tax (right)

Variable	**Change**
Vegetables	-0.037
Fruit	-0.441
Pasta and Bread	-0.137
Cereals and Rice	0
Eggs and Milk	-1.714
Fish	-0.439
Poultry	0
Red Meat	-1.776
Processed Meat	-7.313
Fat and Cheese	-5.251
Oil	-15.222
Sweets and Snacks	-0.830
Sweetened beverages	-0.004
Other drinks	-0.476
Alcohol	0
Food_afh	0

Change in body weight (kg)

	Mean	**Min**	**Max**
1st quintile	-0.89	-5.01	0.17
Sample mean	-1.72	-8.52	0.00
5th quintile	-2.11	-9.65	0.00

Change in daily kJ

	Mean	**Min**	**Max**
1st quintile	-179.045	-1001.83	34.55121
Sample mean	-342.487	-1704.23	0
5th quintile	-427.682	-1943.13	0

Cawley & Meyerhoefer (2012). The first part estimates the probability of positive health expenditure, while the second part estimates health expenditure, if any.

Monthly health expenditure at the individual level is included in the HBS data. Expenditures included are for general practitioners, specialist examinations, dentists and dental services, nurses and other paramedical services, clinical analysis, diagnostic tests, hospitalization in clinics and hospitals, expenditures on prescription and non-prescription drugs and sanitary articles such as medicines, plasters, syringes, first aid kits, bandages and the like, vitamins, minerals and homoeopathic products. Health expenditure is not distributed evenly across respondents. In particular, for the first quintile of the expenditure distribution health expenditure is only 20% less than that of the fifth quintile. Although there is a national healthcare system in Italy that provides free medical care by general practitioners and accessible costs for medical specialists, high-income classes may prefer to pay specialists directly to avoid long waiting lists. As a result, health spending for the first quintile of the income distribution is smaller than for the fifth quintile.

Since the HBS data does not include information on weight, BMI or the health status of households, we match HBS data with the 2015 Italian module of the European Health Interview Survey (EHIS). Table E.2 in Appendix E shows descriptive statistics for the variables resulting from the matching and used in our empirical analysis.

Our base regression specification for estimating the marginal impact of weight on health expenditures is:

$$he_i = \alpha + \boldsymbol{\beta}' \mathbf{X}_i + \varepsilon_i \tag{6}$$

where he_i denotes monthly health expenditures (in Euro) by household i; α is the constant term and $\mathbf{X}_i$ denotes a vector of explanatory variables including age, gender, education level, employment position, marital status, macro-region, income quintile, weight (kg) and height (cm) of each individual; ε_i is the idiosyncratic error term. Table 5 lists regression results for the sample resulting from the matching. The cells indicate marginal effects (reflecting both parts of the two-part model) and standard errors of the marginal effects at the sample mean for the first quintile and for the fifth quintile of the expenditure distribution. Weighing an additional kilogram does not raise health expenditure for individuals in the first quintile of the expenditure distribution. Instead, one additional Kg of weight increases health expenditures by almost 4€ per month on average, and by 6€ per month for individuals in the fifth quintile of the expenditure distribution. Conversely, losing one kilogram decreases monthly health expenditure by the same amounts.

Table 5: Marginal effects of weight on monthly health expenditures

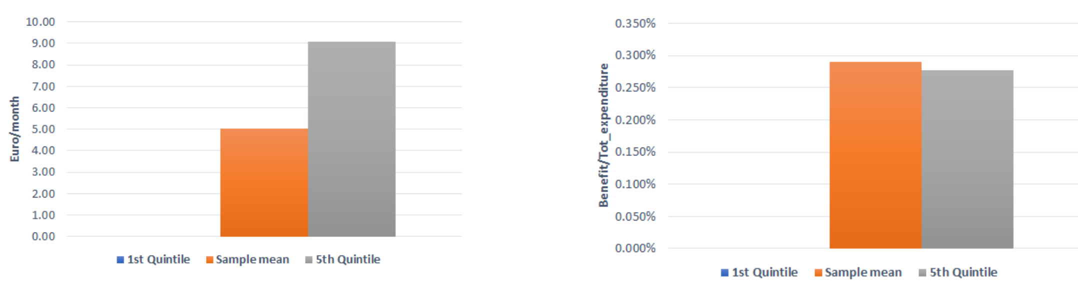
Health Expenditures (Dep. Var.)	Sample mean		1st quintile		5th quintile	
	Probit	GLM	Probit	GLM	Probit	GLM
Weight (kg)	0.0157**	3.979***	0.0209***	-0.325	0.0115	5.965***
	(0.00758)	(1.398)	(0.00564)	(0.616)	(0.00878)	(1.924)
Height (cm)	0.0113*	5.098***	0.00217	2.086**	0.00723	8.186***
	(0.00669)	(1.424)	(0.00844)	(0.876)	(0.00961)	(2.574)
Gender (1=male)	0.791***	153.0***	0.735***	18.40	0.729***	250.6***
	(0.178)	(34.92)	(0.140)	(18.46)	(0.241)	(52.54)
Age	0.0815***	12.43***	0.0562**	8.583***	0.0851***	17.81***
	(0.00876)	(2.287)	(0.0274)	(3.006)	(0.0170)	(5.428)
NorthEast	-0.00255	10.63	0.180	49.20**	-0.0437	-3.196
	(0.0479)	(10.48)	(0.193)	(21.82)	(0.0783)	(23.54)
Centre	-0.0368	-7.103	0.121	-0.999	-0.0832	-24.17
	(0.0456)	(10.83)	(0.157)	(18.51)	(0.0837)	(25.64)
South	0.279***	-29.58***	0.494***	-7.732	-0.0134	-64.31**
	(0.0447)	(11.34)	(0.131)	(14.85)	(0.0962)	(31.57)
Islands	0.0964	-19.78	0.211	-33.94	-0.157	-97.27*
	(0.0773)	(21.53)	(0.185)	(21.54)	(0.176)	(57.55)
2nd Quintile	0.347***	21.18				
	(0.0708)	(20.20)				
3rd Quintile	0.704***	45.78**				
	(0.0828)	(18.77)				
4th Quintile	1.067***	109.8***				
	(0.0774)	(19.13)				
5th Quintile	1.380***	231.0***				
	(0.0643)	(19.51)				
Education	-0.0656***	-11.45*	-0.152*	1.231	-0.0431	-21.88
	(0.0253)	(6.538)	(0.0853)	(8.835)	(0.0485)	(15.34)
Marital status	-0.0298**	-11.10***	-0.0118	-10.94*	-0.0305	-12.36
	(0.0144)	(3.667)	(0.0445)	(5.590)	(0.0247)	(7.596)
Employment position	-0.0332	-12.92	-0.0571	-7.305	0.0822	27.07
	(0.0452)	(10.15)	(0.0946)	(10.03)	(0.0916)	(27.66)
Constant	-5.069***	-1,335***	-3.467**	-323.8*	-2.762	-1,937***
	(1.804)	(369.0)	(1.470)	(166.0)	(2.306)	(557.2)
Obs	8,513	8,513	887	887	2,585	2,585

$* = p < 0.10$; $** = p < 0.05$; $*** = p < 0.01$. Standard errors in parenthesis.

As expected, individuals in the first quintile do not benefit from losing weight, as their health expenditure is significantly lower than individuals in the highest quintile. Interestingly, in the first-step probit regression, the weight coefficient is positive and significant at 5% for the first quintile, with an implied elasticity close to 0.7. So gaining an extra kilogram increases the probability of positive health expenditure for individuals in the lowest quintile, even if they do not benefit from a one-kg reduction when their health expenditure is already positive. To obtain the monetary value of health benefits we multiply the vector of marginal effects in Table 5 by the vector of weight variations resulting from the tax (right hand side of Table 4). Benefits (measured as reduction in monthly health expenditure across different total expenditure groups) are shown in Figure 5 in €/month (Figure 5 (a)) and as a fraction of total consumption expenditure (Figure 5 (b)). No benefits emerge for low-income individuals. High-income individuals benefit more than individuals at the sample mean of the expenditure distribution. Relative to income, we do not detect remarkable differences between individuals at the sample mean of the expenditure distribution and individuals in the highest quintile of the expenditure distribution.

Figure 5: Health benefits

(a) Health benefits (€/month) **(b)** Health benefits/total expenditure (%)



Notes: figure (a) shows health benefits in €/month across the distribution of total expenditure, our proxy for disposable income. Health benefits are calculated as savings in health expenditures (€/month) due to weight lost after the tax. Figure (b) shows health benefits as a fraction of total expenditure.

5.1 Net Welfare Impacts

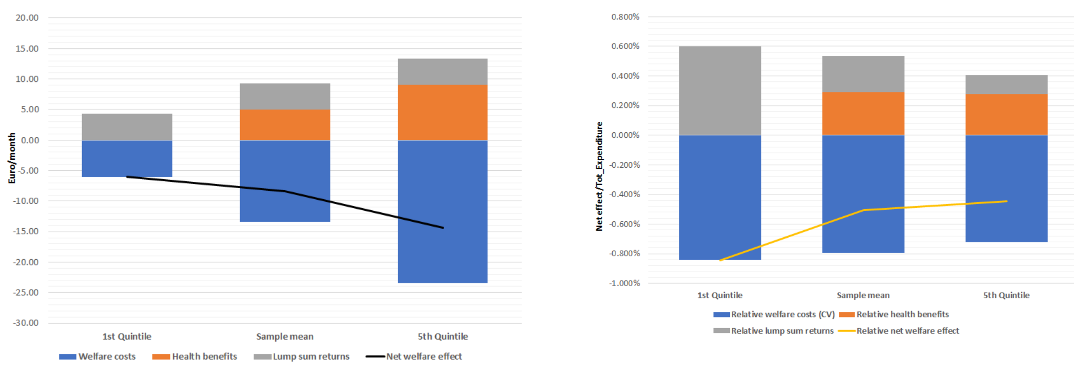
We combine the results in Section 4 with the empirical estimates of the monetary value of weight loss to compute the net welfare impacts from the simulated tax.

We decompose the welfare effects into three distinct components. They are plotted in Figure 6 across the distribution of total expenditure. "Redistributed Revenues" are public revenues from the fat tax equally redistributed as lump-sum transfers. "Welfare benefit" is the money-metric welfare benefit due to weight loss after the tax. "Welfare cost" is the compensating variation, i.e. the amount of money that makes the choice between an increase in their income or the introduction of the tax indifferent to consumers. "Net Welfare Impact" is the difference between welfare costs and benefits.

6 shows that welfare costs are higher than benefits for all groups. In addition, net impacts result in small and progressively distributed losses. This is different from the results of Allcott et al. (2019a), who found, in the context of a sugar tax, small and regressive net benefits. The lump sum returns only marginally offset the welfare costs of the tax.

Figure 6: Net Welfare Impacts

(a) Benefits and Costs (€/month) **(b)** Relative Benefits and Costs



Notes: figure (a) decomposes welfare changes resulting from the fat tax across the distribution of total expenditure. "Welfare costs" are measured by the compensating variation (€/month). "Health benefits" are calculated as savings in health expenditures (€/month) due to weight lost after the introduction of the tax. "Lump sum return" is public revenues (€/month) from the fat tax redistributed equally across the distribution of total expenditure. "Net welfare effect" is the difference between "Welfare costs" and "Health benefits". Figure (b) decomposes costs, benefits and net impacts relative to total expenditure.

The right hand side of Figure 6 shows costs and benefits as a fraction of total expenditure, our proxy for income. Relative to total expenditure, the fat tax generates small and regressively distributed net welfare losses, in line with Allcott et al. (2019a). In order to check whether an easy-to-implement ad valorem tax would lead to different results, we also simulate the introduction of an alternative *ad valorem* tax reducing consumption of saturated fat by 30% and resulting in an increase in the price of food categories high in saturated fats: processed meat, snacks and sweets and fat and cheese. Results are shown in Appendix F. Again, benefits are lower than costs for all groups. Compared to the specific tax, ad valorem taxation implies slightly smaller benefits for individuals in the highest quintile of the expenditure distribution but the net welfare effects from the two tax policies are of similar magnitude.

6 Summary and Conclusion

Modern economies often rely on excise taxation to mitigate socially costly consumption habits. While it is commonplace to scrutinize the potential welfare costs of implementing new tax policies, there is a notable absence in evaluating net welfare impacts, considering potential benefits in addition to costs. However, such an assessment is crucial in evaluating changes in welfare, as recognizing benefits may bolster the social and political acceptability of new tax policies. Our study addresses this gap in the literature by evaluating the net welfare impacts of taxes imposed on unhealthy foods. Specifically, we examine the potential effects of a fat tax in Italy, analyzing both its short-run benefits — proxied by the monetary value of weight loss — and its costs.

First, we assess the suitability of a nutrient tax based on sugar or saturated fat for Italian consumers. Given the WHO threshold of 30g/day maximum consumption of a harmful nutrient, we find that sugar consumption barely reaches this threshold for both singles and families, whereas saturated fat consumption significantly exceeds it. Thus, a fat tax emerges as more suitable for Italy.

Second, we estimate the costs associated with taxation in terms of compensating variation. Our results indicate that high-income individuals are likely to reduce fat consumption more significantly in response to a fat tax, leading to substantial direct consumer surplus loss for this group. Recognizing that families with children may differ substantially from our selected sample of single adults, we estimate the demand system and related elasticities for both single adults and for households with two adults and one child.

Our findings reveal small differences in own and cross-price elasticities between single adults and families with one child, suggesting similar patterns in the welfare costs associated with simulated taxation.

Finally, to ensure consistency in expenditure, consumption, and associated costs and benefits of taxation we restrict our analysis to single adults to calculate the benefits associated with the proposed policy, proxied by the value of direct effects on weight reduction.

Contrary to expectations, our analysis reveals progressive net losses, particularly for high-income individuals.

One limitation is our exclusive focus on weight gain due to excessive saturated fat consumption, without considering broader health impacts or potential public healthcare cost savings. Additionally, our analysis overlooks child obesity, a significant concern in Italy, although recent research suggests broader policy approaches may be more appropriate in addressing this issue (Crudu et al., 2021). Despite the limitations, we hope our study may contribute to shifting the discussion of sin taxes from mere welfare-cost calculation to a more comprehensive assessment, aiding policymakers in making informed decisions.

References

Abdullah, A., Peeters, A., de Courten, M. & Stoelwinder, J. (2010). The magnitude of association between overweight and obesity and the risk of diabetes: A meta-analysis of prospective cohort studies. *Diabetes Research and Clinical Practice*, *89*, 309–319.

Allcott, H., Lockwood, B. B. & Taubinsky, D. (2019a). Regressive sin taxes, with an applicatin to the optimal soda tax. *The Quarterly Journal of Economics*, 1557–1626.

Allcott, H., Lockwood, B. B. & Taubinsky, D. (2019b). Should we tax sugar-sweetened beverages? an overview of theory and evidence. *Journal of Economic Perspectives*, *33*(3), 202–227.

Alpman, A. (2016). Implementing rubin's alternative multiple-imputation method for statistical matching in stata. *The Stata Journal*, *16*(3), 717-739. Retrieved from `https://doi.org/10.1177/1536867X1601600311` doi: 10.1177/1536867X1601600311

Banks, J., Blundell, R. & Lewbel, A. (1997). Quadratic engel curves and consumer demand. *The Review of Economics and Statistics*, *79*, 527–539.

Bogers, R., Bemeals, W., Hoogenveen, R., Boshuizen, H., Woodward, M., Knekt, P. & et al. (2007). Association of overweight with increased risk of coronary heart disease partly independent of blood pressure and cholesterol levels. *Arch. Intern. Med.*, *167*, 1720–1728.

Briggs, A., Mytton, O., Kehlbacher, A., Tiffin, R., Rayner, M. & Scarborough, P. (2013). Overall and income specific effect on prevalence of overweight and obesity of 20% sugar sweetened drink tax in uk: econometric and comparative risk assessment modelling study. *BMJ*, *347*(31), f6189.

Brownell, K. D., Farley, T., Willett, W., Popkin, B. M., Chaloupka, F. J., Thompson, J. W. & Ludwig, D. S. (2009). The public health and economic benefits of taxing sugar-sweetened beverages. *New England Journal of Medicine*, *361*(16), 1599–1605.

Caro, J. C., Valizadeh, P., Correa, A., Silva, A. & Ng, S. W. (2020). Combined fiscal policies to promote healthier diets: Effects on purchases and consumer welfare. *PLoS One*, *15*(1), 1–23.

Cawley, J., Frisvold, J., Hill, A. & Jones, D. (2019). The impact of the philadelphia beverage tax on purchases and consumption by adults and children. *Journal of Health Economics*, *102225*(67).

Cawley, J. & Meyerhoefer, C. (2012). The medical care costs of obesity: An instrumental variables approach. *Journal of Health Economics*, *31*(1), 219-230.

Chen, X., Christensen, T. & Tamer, E. (2018). Monte Carlo confidence sets for identified sets. *Econometrica*.

Chouinard, H. H., Davis, D., LaFrance, J. & Perloff, J. (2007). Fat taxes: Big money for small change. *Forum for Health Economics and Policy*, *10*.

Cohen, M. L. (1991). Statistical matching and microsimulation models. In C. F. Citro & E. A. Hanushek (Eds.), *Improving information for social policy decisions – the uses of microsimulation modeling: Volume II, technical papers* (pp. 62–88). National Academy Press.

Cozzi, S. (2008). *La distribuzione commerciale in italia : caratteristiche strutturali e tendenze evolutive* (Collana Contributi ISTAT). ISTAT.

Cremer, H., Goulao, C. & Roeder, K. (2016). Earmarking and the political support of fat taxes. *Journal of Health Economics*, *50*, 258–267.

Crudu, F., Neri, L. & Tiezzi, S. (2021). Family ties and child obesity in italy. *Economics and Human Biology*, 100951.

Dall, T., Fulconi, V., Zhang, Y., Reimers, K., Packard, P. & Astwood, J. (2009). Potential health benefits and medical cost savings from calorie, sodium, and saturated fat reductions in the american diet. *American Journal of Health Promotion*, *23*, 412-422.

Deaton, A. & Muellbauer, J. (1980). *Economics and consumer behavior.* Cambridge University Press.

D'Orazio, M., Zio, M. D. & Scanu, M. (2006). *Statistical Matching: Theory and Practice (Wiley Series in Survey Methodology).* John Wiley & Sons.

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

Dubois, P., Griffith, R. & O'Connell, M. (2022). The use of scanner data for economics research. *Annual Review of Economics*, *14*, 723–745.

Fichera, E., Mora, T., Lopez-Valcarcel, B. & Roche, D. (2021). How do consumers respond to "sin taxes"? new evidence from a tax on sugary drinks. *Social Science and Medicine*, *274*, 113799.

Finkelstein, E. A., Zhen, C., Nonnemaker, J. & Todd, J. E. (2010, 12). Impact of targeted beverage taxes on higher- and lower-income households. *Archives of Internal Medicine*, *170*(22), 2028-2034.

Fletcher, J. M., Frisvold, D. E. & Tefft, N. (2010). The effects of soft drink taxes on child and adolescent consumption and weight outcomes. *Journal of Public Economics*, *94*(11–12), 967–974.

Griffith, R., Luhrmann, M., O'Connell, M. & Smith, K. (2016). Using taxation to reduce sugar consumption. *IFS Briefing Note*, *BN180*.

Griffith, R., O'Connell, M., Smith, K. & Stroud, R. (2019). The evidence on the effects of soft drink taxes. *IFS Briefing Note*, *BN255*.

Griffith, R., O'Connell, M., Smith, K. & Stroud, R. (2020). What's on the menu? policies to reduce young people's sugar consumption. *Fiscal Studies*, *41*(1), 165-197.

Gruber, J. & Koszegi, B. (2004). Tax incidence when individuals are time-inconsistent: The case of cigarette excise taxes. *Journal of Public Economics*, *9-10*(88), 1959-1987.

Grummon, A. H., Lockwood, B. B., Taubinsky, D. & Allcott, H. (2019). Designing better sugary drink taxes: tax the sugar not the liquid. *Science*, *365*(6457), 989–990.

Hall, K. D., Sacks, G., Chandramohan, D., Chow, C. C., Wang, C., Gortmaker, S. L. & Swinburn, B. A. (2011). Quantification of the effect of energy imbalance on bodyweight. *The Lancet*, *378*(9793), 826-837.

Hanemann, V. & Morey, E. (1992). Separability, partial demand systems, and consumer's surplus measures. *Journal of Environmental Economics and Management*, *22*(3), 241–258.

Harding, M. & Lovenheim, M. (2017). The effect of prices on nutrition: Comparing the impact of product and nutrient-specific taxes. *Journal of Health Economics*, *53*, 53–71.

Harkanen, T., Kotakorpi, K., Pietinen, P., Pirttila, J., Reinivuo, H. & Suoniemi, I. (2014). The welfare effects of health-based food tax policy. *Food Policy*, *49*, 196–206.

Heien, D. & Wesseils, C. R. (1990). Demand systems estimation with microdata: a censored regression approach. *Journal of Business and Economic Statistics*, *8*, 365–371.

Hines, J. (2007). Taxing consumption and other sins. *Journal of Economic Perspectives*, *49-68*(21), 1.

Hirukawa, M., Liu, D. & Prokhorov, A. (2021). msreg: A command for consistent estimation of linear regression models using matched data. *The Stata Journal*, *21*(1), 123–140.

Hirukawa, M. & Prokhorov, A. (2018). Consistent estimation of linear regression models using matched data. *Journal of Econometrics*, *203*(1), 344–358.

Irz, X. (2017). *Demand for food and nutrients and its climate impact: A micro-econometric analysis of economic and socio-demographic drivers* (Natural Resources and Bioeconomy Studies No. 28). Natural Resources Institute Finland, Helsinki.

Jones, A. (2000). Health econometrics. In A. Culyer & J. Newhouse (Eds.), *Handbook of health economics, vol. 1a* (p. 265-336).

Kao, K.-E., Jones, A., Ohinmaa, A. & Paulden, M. (2020). The health and financial impacts of a sugary drink tax across different income groups in canada. *Economics and Human Biology*, *38*, 100869.

LaFrance, J. & Hanemann, W. (1989). The dual structure of incomplete demand systems. *American Journal of Agricultural Economics*, *71*(2), 262–274.

Lal, A., Mantilla-Herrera, A., Veerman, L., Backholer, K., Sacks, G. & Moodie, M. (2017). Modelled health benefits of a sugar-sweetened beverage tax across different socioeconomic groups in australia: a cost-effectiveness and equity analysis. *PLoSMed*, *14*(June 6), e1002326.

Leicester, A. & Windmeijer, F. (2004). *The fat tax: Economic incentives to reduce obesity* (Briefing Note No. 49). Institute for Fiscal Studies.

Lewbel, A. (1989). Identification and estimation of equivalence scales under weak separability. *The Review of Economic Studies*, *56*, 311–316.

Lewbel, A. & Pendakur, K. (2009). Tricks with hicks: The easi demand system. *The American Economic Review*, *99*(3), 827–863.

Lin, B.-H., Smith, J.-Y., Travis A. Lee & Hall, K. D. (2011). Measuring weight outcomes for obesity intervention strategies: The case of a sugar-sweetened beverage tax. *Economics and Human Biology*, *9*(4), 329-341.

McCullough, E., Shin, S., Arsenault, J. & Zhen, C. (2020). *Harnessing food demand systems for improved nutrition in tanzania.* (unpublished)

Miao, Z., Begin, J. & Jensen, H. (2013). Accounting for product substitution in the analysis of food taxes targeting obesity. *Health Economics*, *22*, 1318–1343.

Moriarity, C. & Scheuren, F. (2003). A note on rubin's statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics*, *21*(1), 65–73. Retrieved 2022-05-26, from `http://www.jstor.org/stable/1392351`

Muller, L., Lacroix, A., Lusk, J. & Ruffieux, B. (2017). Distributional impacts of fat taxes and thin subsidies. *The Economic Journal*, *127*, 2066–2092.

Nelson, E., Fitzgerald, J. & Tefft, N. (2019). The distributional impact of a green payment policy for organic fruit. *PLoS One*, *14*(2), 1–25.

Nghiem, N., Jacobi, L., Ramírez-Hassan, A., Waterlander, W. & Blakely, T. (2018). *Thomas bayes goes to the virtual supermarket: Combining prior food price elasticities and experimental data in a large demand system of dietary choices.* (Unpublished manuscript)

O'Donoghue, T. & Rabin, M. (2006). Optimal sin taxes. *Journal of Public Economics*, *10-11*, 1825–1849.

OECD. (2019). *Oecd health statistics, 2019,* (Tech. Rep.). Retrieved from `http://www.oecd.org/els/health-systems/health-data.htm`

Park, D., Lee, J.-H. & Han, S. (2017). The magnitude of association between overweight and obesity and the risk of diabetes: A meta-analysis of prospective cohort studies. *Medicine*, *96*, 1–7.

Pendakur, K. (2009). Easi made easier. In D. Slottje (Ed.), *Quantifying consumer preferences (contributions to economic analysis, vol. 288)* (chap. 7). Emerald Group Publishing Limited.

Program, G. F. R. (2019). *Sugary drink taxes around the world* (Tech. Rep.). Chapel Hill, NC: Global Food Research Program. http://globalfoodresearchprogram.web.unc.edu/multicountry- initiative/resources.

Radner, D., Allen, R., Gonzalez, M., Jabine, T. & Muller, H. (1980). Report on exact and statistical matching techniques. *Statistical Policy Working Paper 5. U.S. Department of of Commerce, Office of Federal Statistical Policy and Standards.*

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J. & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology*, *27*(1), 85–96.

Rodgers, W. L. (1984). An evaluation of statistical matching. *Journal of Business & Economic Statistics*, *2*(1), 91–102.

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics*, *4*(1), 87–94. Retrieved 2022-05-26, from `http://www.jstor.org/stable/1391390`

Shonkwiler, J. S. & Yen, S. (1999). Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics*, *81*, 972–982.

Spinelli, A., Censi, L., Mandolini, D., Ciardullo, S., Salvatore, M., Mazzarella, G. & Nardone, P. (2023). Inequalities in childhood nutrition, physical activity, sedentary behaviour and obesity in italy. *Nutrients*, *15*(18), 3893–.

WHO. (2015). *Sugars intake for adults and children* (Tech. Rep.).

WHO. (2018). *Saturated fatty acid and trans-fatty acid intake for adults and children* (Tech. Rep.).

WHO. (2022). *European childhood obesity surveillance initiative (cosi) report on the fifth round of data collection, 2018–2020* (Tech. Rep.). WHO Regional Office for Europe, Copenhagen.

Wilde, P., Huang, Y., Sy, S., Abrahams-Gessel, S., Jardim, T. & Paarlberg, R. (2019). Cost effectiveness of a us national sugar-sweetened beverage tax with a multistakeholder approach: who pays and who benefits. *American Journal of Public Health*, *109*(February 2), 276–284.

Wooldridge, J. W. (2010). *Econometric analysis of cross section and panel data (second edition)*. The MIT Press.

WorldBank. (2020). *Taxes on sugar-sweetened beverages: International evidence and experiences.* (Tech. Rep.).

Wright, A., Smith, K., Hellowell, M. P., Griffith, R. & O'Connell, M. (2017). Policy lessons from health taxes: a systematic review of empirical studies. *BMC Public Health*, *583*(17).

Xiang, D., Zhan, L. & Bordignon, M. (2018). A reconsideration of the sugar sweetened beverage tax in a household production model. *Cesifo Working Papers*, *7087*.

Zhen, C., Finkelstein, E. A., Nonnemaker, J., Karns, S. & 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.

Zhen, C., Muth, M., Okrent, A., Karns, S., Brown, D. & Siegel, P. (2019). Do differences in reported expenditures between household scanner data and expenditure surveys matter in health policy research? *Health Economics*, *28*(6), 782-800.

Online appendix of "Let Them Eat Cake. The Net Consumer Welfare Impact of a Fat Tax"

Valeria Di Cosmo* Silvia Tiezzi†

University of Turin University of Siena

9th May 2024

Abstract

This supplement contains plots of excess consumption of sugar and saturated fat for households of different size (section A); descriptive statistics (section B), Engel curve plots for the 16 food groups (section C), additional estimated elasticities (section D), the details of the statistical matching procedure (section E), and an additional counterfactual experiment considering a Value Added Tax (VAT) on fat and cheese, processed meat and sweets and snacks (section F).

Keywords: sin taxes, welfare benefits, welfare costs, exact affine stone index demand system, demand elasticities, micronutrients intake.

JEL classification: O12, D12, I15.

*Department of Economics and Statistics "Cognetti de Martiis", University of Turin, Lungo Dora Siena, 100 10153 Turin, valeria.dicosmo@unito.it.

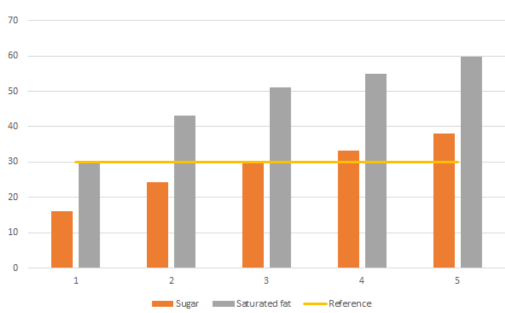
†Department of Economics and Statistics, University of Siena, Piazza San Francesco 7/8, 53100, Siena, silvia.tiezzi@unisi.it.

A Excess consumption in other household types.

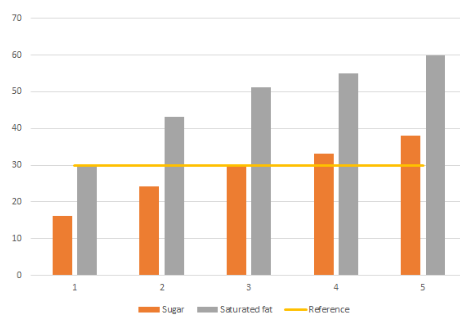
In this appendix, we use data on food consumption expenditure in the Household Budget Survey (ISTAT) combined with nutrients data from the Composition Database for Epidemiological Studies in Italy (EIO). Figure A.1 documents excess consumption of added sugar and saturated fats in households with two adults and two adults and two children across income quintiles. For each household type, excess consumption of sugar and saturated fats increases along the distribution of income, consistently with what we observe for single adults and two adults and one child.

Figure A1: Consumption of sugar and saturated fats.

(a) Two adults



(b) Two adults, two children



B Descriptive Statistics

Table B1: Summary Statistics

Expenditure Shares

	One adult					Two adults & one child				
	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
Alcohol	12,369	0.010	0.018	0	0.223	4,772	0.007	0.011	0	0.216
Bread & Pasta	12,369	0.013	0.013	0	0.204	4,772	0.014	0.011	0	0.144
Cereals & Rice	12,369	0.003	0.005	0	0.092	4,772	0.003	0.004	0	0.068
Eggs & Milk	12,369	0.010	0.010	0	0.133	4,772	0.014	0.012	0	0.146
Fat & Cheese	12,369	0.012	0.012	0	0.176	4,772	0.014	0.010	0	0.177
Fish	12,369	0.013	0.018	0	0.278	4,772	0.015	0.017	0	0.147
Food_afh	12,369	0.051	0.064	0	0.671	4,772	0.046	0.050	0	0.401
Fruit	12,369	0.015	0.014	0	0.183	4,772	0.015	0.012	0	0.135
Oil	12,369	0.005	0.009	0	0.243	4,772	0.004	0.008	0	0.213
Other	12,369	0.816	0.105	0.240	1.000	4,772	0.806	0.093	0.333	1.000
Otherdrinks	12,369	0.010	0.010	0	0.309	4,772	0.010	0.009	0	0.100
Processed Meat	12,369	0.012	0.014	0	0.159	4,772	0.014	0.011	0	0.098
Poultry	12,369	0.007	0.010	0	0.145	4,772	0.008	0.010	0	0.125
Red meat	12,369	0.014	0.018	0	0.203	4,772	0.016	0.017	0	0.172
Sweet drinks	12,369	0.004	0.006	0	0.102	4,772	0.005	0.005	0	0.068
Sweets & Snacks	12,369	0.017	0.015	0	0.180	4,772	0.021	0.014	0	0.140
Vegetables	12,369	0.023	0.021	0	0.267	4,772	0.022	0.017	0	0.143

Log Lewbel prices

	One adult					Two adults & one child				
	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
Alcohol	12,369	-0.658	0.271	-1.609	0.889	4,772				
Bread & Pasta	12,369	-1.730	0.168	-2.280	-1.426	4,772	-1.114	0.190	-1.757	-0.675
Cereals & Rice	12,369	-3.437	0.272	-4.256	-2.773	4,772	-2.843	0.283	-3.660	-2.117
Eggs & Milk	12,369	-1.670	0.188	-2.867	1.240	4,772	-0.953	0.192	-2.038	-0.551
Fat & Cheese	12,369	-0.818	0.177	-1.686	-0.276	4,772	-0.418	0.197	-1.148	0.183
Fish	12,369	-1.190	0.227	-1.991	-0.609	4,772	-0.519	0.216	-1.372	0.144
Food_afh	12,369	0.191	0.229	-0.797	0.826	4,772	0.439	0.296	-0.669	1.276
Fruit	12,369	-0.400	0.243	-1.589	0.112	4,772	-0.098	0.263	-1.115	0.629
Oil	12,369	2.658	0.309	-3.636	-1.568	4,772				
Other	12,369	3.075	0.269	1.737	4.114	4,772	3.213	0.237	2.147	3.912
Otherdrinks	12,369	-1.996	0.192	-2.639	-1.596	4,772	-1.468	0.195	-1.985	-1.032
Processed Meat	12,369	-1.472	0.169	-2.037	-0.967	4,772	-0.883	0.166	-1.428	-0.504
Poultry	12,369	0.010	0.013	-0.004	0.038	4,772	-1.389	0.172	-2.000	-0.697
Red meat	12,369	-1.162	0.186	-1.980	-0.688	4,772	-0.728	0.179	-1.274	-0.288
Sweet drinks	12,369	-2.779	0.253	-3.494	-1.957	4,772	-1.983	0.257	-2.688	-1.389
Sweets & Snacks	12,369	0.710	0.315	-1.026	1.391	4,772	0.870	0.315	-0.328	1.886
Vegetables	12,369	1.623	0.534	-0.757	3.048	4,772	1.547	0.497	-0.409	3.171

Notes: For Alcohol and Oil prices for households with three members the Lewbel procedure did not converge due to the small number of observations.

Table B2: Summary Statistics cont'ed

Control Variables										
	One adult					**Two adults & one child**				
	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
Total monthly expenditure	12,369	1869,209	1090.125	110	9697.53	4,772	2894.956	1421.251	326.6	9670.06
Gender	12,369	1.465	0.499	1	2	4,772	1.233	0.423	1	2
Education	12,369	3.831	0.828	1	5	4,772	3.925	0.744	1	5
Marital status	12,369	2.315	1.786	1	6	4,772	1.908	0.622	1	6
Employment position	12,369	2.011	1.971	1	8	4,656	1.924	1.004	1	4
Age	12,369	6.008	2.209	1	9	4,772	7.800	1.546	4	15
Metropolitan area	12,369	0.154	0.361	0	1	4,772	0.129	0.336	0	1
Medium city	12,369	0.294	0.456	0	1	4,772	0.272	0.445	0	1
Small city	12,369	0.551	0.497	0	1	4,772	0.599	0.490	0	1

Table B3: Share of food expenditures by the education level of the reference person.

One adult

Mean expenditure share	no education	primary school	lower middle school	high school	undergraduate or postgrad. degree
Alcohol	0.013	0.01	0.012	0.010	0.009
Bread & pasta	0.022	0.014	0.011	0.008	0.006
Cereals & rice	0.008	0.020	0.016	0.013	0.010
Eggs & Milk	0.017	0.014	0.012	0.01	0.008
Fat & Cheese	0.015	0.017	0.014	0.012	0.01
Fish	0.015	0.015	0.013	0.012	0.011
Food_afh	0.024	0.021	0.041	0.05	0.064
Fruit	0.018	0.020	0.017	0.015	0.012
Oil	0.006	0.007	0.005	0.004	0.003
Other	0.740	0.754	0.788	0.822	0.859
Otherdrinks	0.014	0.013	0.011	0.01	0.008
Processed meat	0.014	0.017	0.015	0.004	0.009
Poultry	0.016	0.012	0.009	0.012	0.005
Red meat	0.023	0.022	0.017	0.007	0.01
Sweet drinks	0.003	0.005	0.004	0.014	0.003
Sweets & snacks	0.003	0.018	0.018	0.004	0.015
Vegetables	0.003	0.034	0.026	0.022	0.018
Obs.	83	526	3,387	5,781	2,592

Two adults & one child

Mean expenditure share	no education	primary school	lower middle school	high school	undergraduate or postgrad. degree
Alcohol	0.013	0.007	0.008	0.008	0.007
Bread & pasta	0.019	0.025	0.017	0.014	0.011
Cereals & rice	0.003	0.005	0.003	0.003	0.002
Eggs & Milk	0.009	0.019	0.016	0.013	0.021
Fat & Cheese	0.009	0.015	0.014	0.014	0.012
Fish	0.010	0.016	0.015	0.015	0.014
Food_afh	0.039	0.014	0.035	0.046	0.060
Fruit	0.013	0.021	0.016	0.015	0.014
Oil	0.002	0.008	0.005	0.004	0.003
Other	0.805	0.736	0.780	0.807	0.839
Otherdrinks	0.004	0.007	0.011	0.009	0.008
Processed meat	0.011	0.017	0.016	0.014	0.010
Poultry	0.005	0.015	0.010	0.008	0.006
Red meat	0.019	0.025	0.019	0.016	0.012
Sweet drinks	0.006	0.007	0.006	0.005	0.004
Sweets & snacks	0.016	0.026	0.022	0.021	0.018
Vegetables	0.018	0.031	0.025	0.022	0.019
Obs.	6	71	1,264	2,366	1,065

Table B4: Share of food expenditures by geographic area.

	One adult				Two adults & one child			
Expenditure share	North	Centre	South	Islands	North	Centre	South	Islands
Alcohol	0.011	0.009	0.011	0.009	0.008	0.007	0.007	0.007
Bread & pasta	0.011	0.014	0.016	0.017	0.013	0.014	0.017	0.019
Cereals & rice	0.002	0.003	0.004	0.003	0.003	0.003	0.003	0.003
Eggs & Milk	0.009	0.010	0.012	0.011	0.012	0.013	0.017	0.017
Fat & Cheese	0.012	0.012	0.014	0.011	0.013	0.012	0.016	0.013
Fish	0.009	0.013	0.016	0.019	0.012	0.016	0.019	0.019
Food_afh	0.061	0.046	0.038	0.044	0.053	0.049	0.034	0.031
Fruit	0.013	0.016	0.017	0.017	0.014	0.015	0.017	0.016
Oil	0.004	0.004	0.005	0.006	0.004	0.004	0.005	0.006
Other	0.840	0.820	0.779	0.793	0.827	0.813	0.770	0.774
Otherdrinks	0.008	0.009	0.012	0.014	0.008	0.010	0.011	0.013
Processed meat	0.011	0.013	0.015	0.012	0.013	0.014	0.016	0.015
Poultry	0.006	0.007	0.010	0.08	0.007	0.009	0.010	0.009
Red meat	0.011	0.016	0.018	0.018	0.014	0.016	0.020	0.020
Sweet drinks	0.004	0.009	0.004	0.005	0.005	0.005	0.005	0.007
Sweets & snacks	0.016	0.015	0.019	0.016	0.020	0.019	0.023	0.022
Vegetables	0.020	0.024	0.027	0.025	0.020	0.022	0.025	0.023
Obs.	5,954	2,434	3,148	833	2,294	918	1,187	373

Table B5: Share of food expenditures by gender of the reference person

	One adult		**Two adults & one child**	
Expenditure share	Female	Male	Female	Male
Alcohol	0.006	0.014	0.007	0.008
Bread & pasta	0.013	0.013	0.014	0.015
Cereals & rice	0.003	0.003	0.003	0.003
Eggs & Milk	0.011	0.009	0.013	0.014
Fat & Cheese	0.013	0.012	0.013	0.014
Fish	0.013	0.012	0.015	0.015
Food_afh	0.036	0.064	0.048	0.045
Fruit	0.016	0.014	0.014	0.015
Oil	0.005	0.004	0.004	0.004
Other	0.820	0.815	0.816	0.803
Otherdrinks	0.010	0.009	0.009	0.010
Processed meat	0.012	0.013	0.013	0.014
Poultry	0.007	0.007	0.007	0.009
Red meat	0.014	0.015	0.015	0.017
Sweet drinks	0.004	0.004	0.005	0.005
Sweets & snacks	0.018	0.016	0.020	0.021
Vegetables	0.025	0.021	0.022	0.022
Obs	5756	6613	1,112	3,660

Table B6: Share of food expenditures across the distribution of total expenditure

One adult

Expenditure share	1^{st} quintile	Central quintiles	5^{th} quintile
Alcohol	0.009	0.010	0.011
Bread & pasta	0.021	0.013	0.007
Cereals & rice	0.004	0.003	0.002
Eggs & Milk	0.015	0.010	0.006
Fat & Cheese	0.015	0.013	0.009
Fish	0.013	0.013	0.011
Food_afh	0.025	0.054	0.068
Fruit	0.018	0.016	0.012
Oil	0.005	0.005	0.004
Other	0.780	0.815	0.863
Otherdrinks	0.012	0.010	0.008
Processed meat	0.015	0.013	0.009
Poultry	0.012	0.008	0.005
Red meat	0.015	0.014	0.010
Sweet drinks	0.004	0.004	0.003
Sweets & snacks	0.018	0.017	0.014
Vegetables	0.029	0.023	0.017
Obs	2474	7421	2474

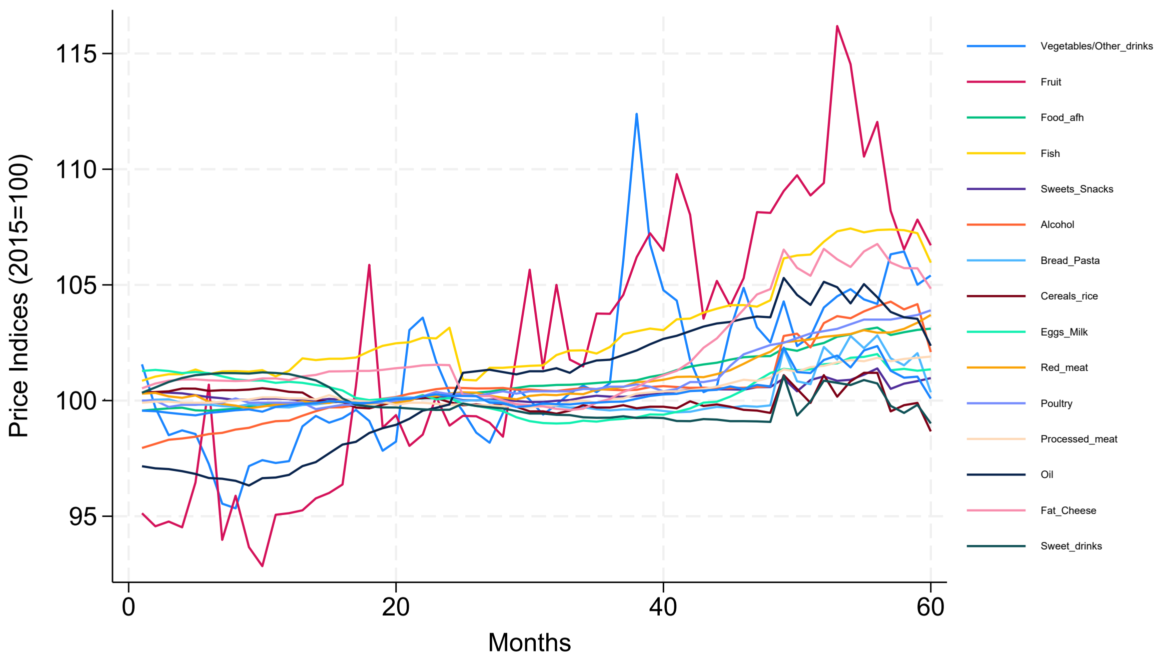
Two adults and one child

Expenditure share	1^{st} quintile	Central quintiles	5^{th} quintile
Alcohol	0.006	0.007	0.008
Bread & pasta	0.021	0.015	0.010
Cereals & rice	0.004	0.003	0.002
Eggs & Milk	0.019	0.014	0.010
Fat & Cheese	0.016	0.015	0.011
Fish	0.014	0.016	0.014
Food_afh	0.020	0.044	0.061
Fruit	0.018	0.016	0.012
Oil	0.005	0.004	0.004
Other	0.767	0.790	0.841
Otherdrinks	0.011	0.010	0.008
Processed meat	0.016	0.016	0.011
Poultry	0.013	0.009	0.006
Red meat	0.020	0.018	0.013
Sweet drinks	0.006	0.005	0.003
Sweets & snacks	0.022	0.023	0.018
Vegetables	0.026	0.024	0.018
Obs	957	1904	1908

1^{st} quintile: between 110 and 1022 Euro/month; Central quintiles: between 1123 and 2542 Euro/month; 4^{th}: between 2543 and 9697 Euro/month.

Figure B1 plots the time series of consumer monthly price indices (2015=100) from January 2014 (0 on the horizontal axis) to December 2018 (60 on the horizontal axis) supplied by ISTAT (). We address the lack of cross-sectional variability in the price indices by computing Lewbel prices.

Figure B1: Monthly price indices (Jan 2014=0 to Dec 2018=60)



Source: ISTAT, Indice nazionale dei prezzi al consumo per l'intera collettività

C Engel curves

We estimate our demand system using seemingly unrelated regression methods. Figures C1 and C2 plot the Engel curves for the 16 food aggregates. Inspection of these Figures suggests that the Engel curve shapes cannot be adequately represented by a linear or quadratic function. To determine the degree of the income polynomials, we add a degree at a time starting from $L = 2$ and test the joint significance of the b_L coefficients by minimum distance (Wooldridge, 2010). Under the null hypothesis that the L^{th} degree of polynomial is excludable, the test statistic is asymptotically distributed as $\chi^2_{(J-1)}$. At $L = 5$ the test statistic still rejects the null hypothesis. We therefore decided that a fifth polynomial in y was sufficient to capture the curvature of the Engel curves.

Figure C1: Kernel estimation of expenditure shares on log total expenditure

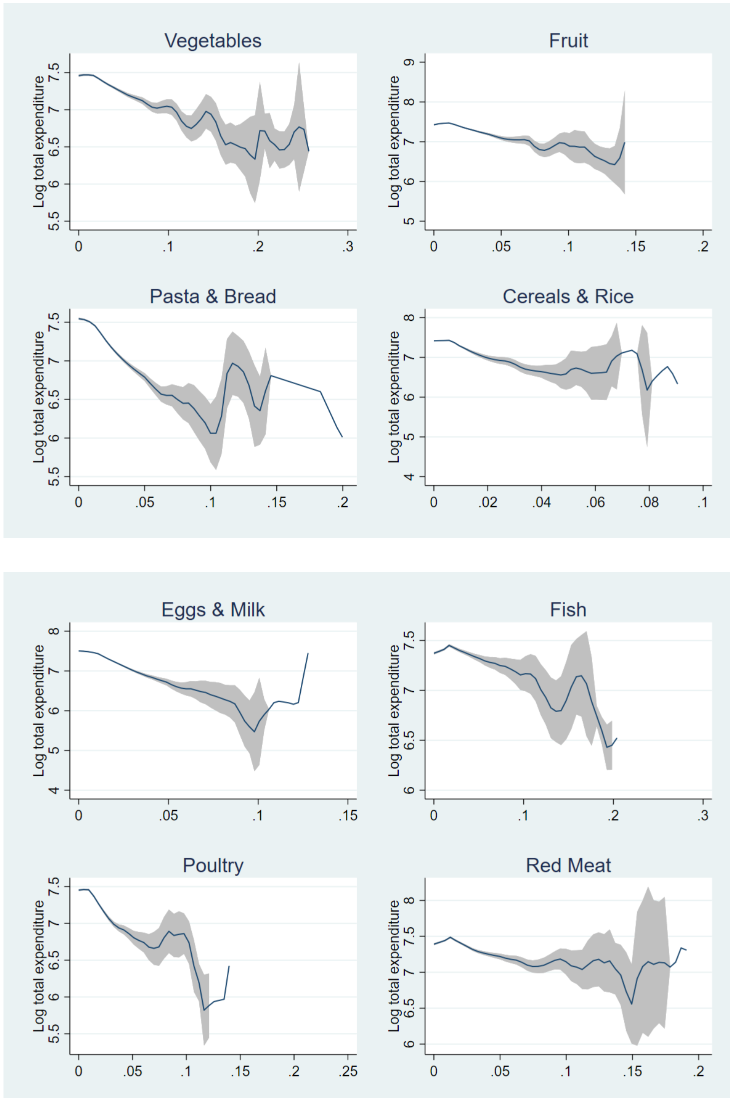


Figure C2: Kernel estimation of expenditure shares on log total expenditure

D Elasticities

Marshallian price elasticities of quantities, expenditure elasticities, and Hicksian price elasticities of quantities derived from the EASI demand system are computed as (Irz, 2017):

$$\frac{\partial lnq^i}{\partial lnp^j} = \frac{a^{ij}}{w^i} + \bar{w}^j - \delta_{ij} - w^j \left[\sum_{r=1}^{R} b_r^i r \left(\hat{y} \right)^{r-1} + \frac{1}{w^i} + 1 \right] \tag{1}$$

$$\frac{\partial lnq^i}{\partial lnx} = \left[\sum_{r=1}^{R} b_r^i r \left(\hat{y} \right)^{r-1} \right] \frac{1}{w^i} + 1 \tag{2}$$

$$\left. \frac{\partial lnq^i}{\partial lnp^j} \right|_{\bar{u}} = \frac{a^{ij}}{\bar{w}^i} - \delta_{ij} + \bar{w}^j \tag{3}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise[1]. Standard errors of elasticities are bootstrapped with 200 replications.

[1] When estimated at the sample mean, Marshallian price elasticity of quantities are computed as $\frac{\partial lnq^i}{\partial lnp^j} = \frac{a^{ij}}{\bar{w}^i} - \delta_{ij} - \frac{\bar{w}^j}{\bar{w}^i} \left[\sum_{r=1}^{R} b_r^i r \left(\hat{\bar{y}} \right)^{r-1} \right]$

Table D1: Uncompensated price elasticities (single adults) - sample means.

	Vegetables	Fruit	Pasta & bread	Cereals & Rice	Eggs & Milk	Fish	Poultry	Red meat	Processed meat	Fat & Cheese	Oil	Sweets & snacks	Sweetened beverages	Other drinks	Alcohol	Fafh
Vegetables	**-1.906*** **	-0.017	0.120***	-0.003	0.082***	-0.143***	-0.003	0.073**	0.043*	-0.005	-0.035**	0.185***	-0.020**	0.010	-0.080***	0.066*
Fruit	-0.028	**-2.203*** **	0.018	-0.108***	-0.013	-0.042	0.001	0.045	0.106***	-0.078**	-0.022	0.026	0.091***	0.022	0.146***	0.291***
Pasta & bread	0.215***	0.028	**-0.781*** **	-0.035	-0.010	-0.076	-0.103*	0.003	-0.246***	0.084*	-0.006	-0.055	0.005	-0.019	-0.050	0.132**
Cereals & and Rice	-0.003	-0.558***	-0.156	**-0.165**	0.114	-0.076	-0.087	-0.256	0.207	-0.017	-0.190	-0.411***	-0.066	-0.009	0.388**	0.396*
Eggs & milk	0.191***	-0.014	-0.016	0.031	**-1.011*** **	-0.141**	0.048	-0.236***	-0.299***	0.056	-0.040	-0.151***	0.003	0.015	0.031	0.146*
Fish	-0.267***	-0.056	-0.091*	-0.021	-0.118**	**-1.509*** **	-0.017	-0.055	0.049	0.003	-0.048	0.051	0.043	-0.058	-0.086	0.090
Poultry	0.002	0.011	-0.170*	-0.033	0.063	-0.015	**0.198**	-0.616***	-0.092	-0.145	-0.247***	0.153*	-0.052	0.022	-0.106	-0.085
Red meat	0.120**	0.050	-0.004	-0.058	-0.180***	-0.049	-0.369***	**-0.786*** **	-0.176**	-0.083	0.046	-0.028	-0.023	-0.114*	-0.174***	-0.089
Processed meat	0.083*	0.137***	-0.267***	0.046	-0.244***	0.058	-0.062	-0.187**	**-0.334*** **	-0.108*	-0.058	-0.076	-0.014	-0.241***	-0.098	0.004
Fat & cheese	-0.013	-0.097**	0.083*	-0.007	0.041	0.006	-0.098*	-0.090	-0.113*	**-1.447*** **	0.036	-0.064	-0.016	-0.080	-0.048	0.074
Oil	-0.174**	-0.071	-0.023	-0.123	-0.091	-0.130	-0.438***	0.141	-0.161	0.100	**-0.587*** **	-0.072	-0.164***	0.155	-0.060	0.158
Sweets & snacks	0.232***	0.012	-0.061*	-0.076***	-0.102***	0.032	0.062	-0.032	-0.069	-0.056	-0.024	**-3.801*** **	-0.119***	-0.125**	0.103**	0.459***
Sweetened beverages	-0.112**	0.350***	0.012	-0.050	0.006	0.141*	-0.106	-0.074	-0.046	-0.045	-0.184***	-0.478***	**-0.809*** **	-0.397***	0.035	0.244**
Otherdrinks	0.032	0.043	-0.025	-0.004	0.017	-0.065	0.017	-0.149*	-0.300***	-0.094	0.073	-0.192**	-0.162***	**-1.054*** **	-0.059	0.161*
Alcohol	-0.168***	0.226***	-0.065	0.107**	0.032	-0.096	-0.083	-0.221**	-0.116	-0.051	-0.025	0.189**	0.015	-0.056	**-0.686*** **	-0.408***
Food_afh	0.029*	0.089***	0.029*	0.019*	0.026*	0.027	-0.017	-0.021	-0.001	0.020	0.014	0.165***	0.019**	0.027	-0.086***	**-2.822*** **

$* = p < 0.10$; $** = p < 0.05$; $*** = p < 0.01$. Standard errors bootstrapped with 200 replications.

Table D2: Compensated own price and expenditure elasticities at different levels of total expenditure (single adults)

	1^{st} quintile		5^{th} quintile	
	Price	Expenditure	Price	Expenditure
Vegetables	-2.566***	1.178***	-1.651***	1.638***
Fruit	-2.260***	1.709***	-1.819***	1.348***
Pasta & Bread	-0.866***	0.838***	-0.768***	0.722***
Cereals & Rice	1.214*	-0.062	-0.804***	1.618***
Eggs & Milk	-0.822***	0.737*	-1.294***	1.052***
Fish	-1.719***	1.862***	-1.649***	1.476***
Poultry	0.128	0.733	-0.704***	0.717**
Red Meat	-1.019***	2.027***	-1.076***	0.915**
Processed Meat	-0.301	0.892	-0.689***	0.926***
Fat & Cheese	-0.938***	1.231***	-1.561***	1.219***
Oil	-0.324	1.103	-1.004***	0.63
Sweets & Snacks	-4.792***	2.128***	-3.004***	2.591***
Sweetened beverages	-1.172***	1.912**	-0.407***	0.316
Other drinks	-0.781**	0.695	-1.240***	0.315
Alcohol	0.15	-0.173	-1.468***	-0.448
Food_afh	-2.569***	-0.179	-2.361***	-0.348
Obs	2474	2474	2474	2474

1^{st} quintile: between 110 and 1022 Euro/month; 5^{th} quintile: between 2543 and 9697 Euro/month. $* = p < 0.10$; $** = p < 0.05$; $*** = p < 0.01$. Standard errors bootstrapped with 200 replications.

Table D3: Compensated own price elasticities by geographic area (single adults)

	North	Centre	South	Islands
Alcohol	-1.386***	-0.541*	0.255	-1.003*
Bread & pasta	-0.932***	-0.548***	-0.766***	-0.587*
Cereals & rice	-0.794***	0.299	0.239	-0.931
Eggs & Milk	-1.044***	-0.855***	-0.906***	-1.805***
Fat & Cheese	-1.621***	-0.993***	-1.297***	-1.545***
Fish	-1.451***	-1.674***	-1.487***	-1.352***
Food_afh	-2.974***	-2.524***	-2.471***	-2.789**
Fruit	-2.215***	-2.044***	-2.149***	-2.543***
Oil	-0.888***	-0.574**	-0.456**	0.486
Otherdrinks	-1.346***	-0.481*	-0.995***	-1.197*
Processed meat	-0.758***	-0.746***	0.252	0.185
Poultry	-0.306	0.362	0.575*	0.844
Red meat	-0.949***	-0.780***	-0.596***	-0.373
Sweetened beverages	-0.974***	-0.604***	-0.731***	-0.739***
Sweets & snacks	-3.944***	-3.551***	-3.716***	-3.484***
Vegetables	-1.887***	-1.885***	-2.081***	-2.272***

$* = p < 0.10$; $** = p < 0.05$; $*** = p < 0.01$. Standard errors bootstrapped with 200 replications.

Table D4: Expenditure elasticities by geographic area (single adults)

	North	Centre	South	Islands
Alcohol	0.362**	0.730*	0.259	1.640**
Bread & pasta	0.728***	0.641***	0.376**	0.755**
Cereals & rice	0.325	-0.468	0.183	2.513*
Eggs & Milk	0.688***	0.575***	0.771***	1.930***
Fat & Cheese	1.146***	1.012***	1.104***	1.282***
Fish	1.309***	1.228***	1.327***	2.342***
Food_afh	0.761***	0.952***	1.012***	1.205**
Fruit	1.172***	1.237***	0.979***	1.579***
Oil	1.074*	0.294	1.206**	0.953
Otherdrinks	1.015***	0.677*	0.497	0.643
Processed meat	0.893***	0.690***	0.526**	0.992*
Poultry	0.635**	0.369	0.463	0.098
Red meat	1.215***	1.050***	1.150***	1.26
Sweetened beverages	1.227**	0.543	0.990***	0.903
Sweets & snacks	1.859***	1.839***	1.972***	2.581***
Vegetables	1.155***	0.858***	0.979***	1.123****

$* = p < 0.10$; $** = p < 0.05$; $*** = p < 0.01$. Standard errors bootstrapped with 200 replications.

E Statistical Matching

We follow Alpman (2016)'s two-step procedure to implement Rubin (1986) statistical matching between two datasets. In particular, if dataset 1 contains the variable weight, dataset 2 contains the variable health expenditures, and 1 and 2 contain a set of common variables, X, statistical matching allows the creation of a new dataset containing health expenditures, weight and X for all respondents. Health expenditures are included in the Household Budget Survey (HBS), and the weight of each individual is included in the European Health Interview Survey (EHIS) for 2015. Variables shared by the two datasets are: number of family members, age, gender, income quintile, geographic location, education level and employment status of the respondent[2].

The purpose of the matching is to obtain a new dataset that includes health expenditures, individual weight and a set of control variables. We use the dataset resulting from the matching to estimate equation 6 in our paper. The first step of the procedure generates the predicted weight and health expenditure values for each observation of the incomplete original dataset as a function of the assumed partial correlation between weight and health expenditures, conditional on the control variables. In the second step, each unit in the EHIS for which health expenditures is missing is matched with the corresponding unit in the HBS with the closest predicted value of health expenditures calculated in step 1, conditional on the set of control variables. Similarly, each unit in the HBS for which weight is missing is matched with the corresponding unit in the EHIS database that has the closest predicted value of weight as calculated in step 1, conditional on the control variables. We allow the partial correlation, ρ, between health expenditures and weight, conditional on the variables, to vary between 0.1 and 1. We run our regressions considering multiple imputations of health expenditures and weight using all values of ρ between 0.1 and 1. As suggested by Alpman (2016), multiple imputation reduces the risk of downward bias

[2] The absence of continuous variables in the set of common variables prevented us from using the more recent matching method and estimation procedure proposed by Hirukawa & Prokhorov (2018); Hirukawa et al. (2021).

in the estimated standard errors[3]. Consistently with our main empirical analysis, we consider single households aged less than 65 years. Summary statistics of both the initial and matched datasets are shown below.

Table E1: Summary statistics, original datasets (EHIS and HBS)

Variable	Obs	Mean	Std. dev.	Min	Max
HBS					
Health expenditures	12,419	117.22	255.90	0	5710.94
Gender (1=male)	12,419	1.47	0.50	1	2
Age	12,419	9.00	2.21	4	12
Geographical location	12,419	2.62	1.31	1	5
Income quintile	12,419	3.26	1.39	1	5
Employment position	11,607	1.91	1.02	1	4
Education	12,419	3.83	0.83	1	5
Marital status	12,419	2.08	1.39	1	4
Employment status (1=employed)	12,419	3.26	1.39	1	5
EHIS					
Weight (kg)	1,498	71.94	13.93	40	127
Height (cm)	1,500	170.97	9.12	140	195
Gender (1=male)	1,507	1.41	0.49	1	2
Age	1,507	8.66	2.16	4	13
Geographical location	1,507	2.50	1.26	1	5
Employment position	1,507	1.77	0.70	1	3
Education	1,507	3.93	0.78	1	5
Marital status	1,507	1.87	1.24	1	4
Income quintile	1,507	3.77	1.15	1	5

[3] In Stata, we used `mi impute` and `mi estimate` commands.

Table E2: Descriptive statistics, matched dataset

Variable	Obs	Mean	Std. dev.	Min	Max
Gender(1=male)	16,261	1.48	0.50	1	2
Age	16,261	9.55	2.48	4	13
Geographical location	16,261	2.63	1.31	1	5
Employment position	15,043	1.65	0.62	1	3
Education	16,261	3.75	0.89	1	5
Marital status	16,261	2.16	1.37	1	4
Income	16,261	3.29	1.38	1	5
Imputed health expenditures: sample mean	15,982	128.02	267.44	0	5711
Imputed health expenditures: 1st quintile	2,369	36.84	69.78	0	709
Imputed health expenditures: 5th quintile	3,986	261.72	447.69	0	5711
Imputed weight: sample mean	9,633	71.20	12.08	40	127
Imputed weight: 1st quintile	991	69.95	12.84	40	120
Imputed weight: 5th quintile	2,946	72.28	12.53	40	127

As shown in Table E2, the original EHIS dataset has 1498 observations on weight of individuals under 65. The Rubin procedure adds 8135 new observations for which a matching with the HBS is possible, which leads to 9633 observations on imputed weight in the final dataset. For health expenditures the original HBS dataset has 12419 observations, increased to 15982 by the matching algorithm. The final dataset with health expenditures, weight and a common set of control variables contains 8513 observations.

F Ad valorem tax

In addition to the main counterfactual experiment, we simulate an easy to implement increase in the existing Value Added Tax (VAT) on fat and cheese, processed meat and sweets and snacks (i.e. the food groups highest in saturated fat) that would cut fat consumption by 30%, resulting in a 4.3% increase in their initial prices. This amounts to the introduction of an ad valorem (av) fat tax (t), such that the after-tax price of a taxed food group j, $p^j_{1,av}$, is:

$$p^j_{1,av} = p^j_0(1 + t\eta^j) \tag{4}$$

Since fat and cheese, processed meat and sweets and snacks differ both in the per kg content of saturated fat and in the compensated price elasticity of quantity, we compute the ad valorem tax that brings about a 30% decrease in saturated fat consumption as:

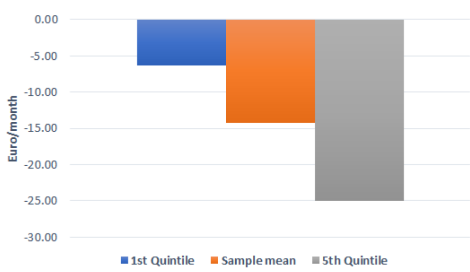
$$t\bar{\eta} = \frac{-0.30}{\bar{\epsilon}}\bar{\eta} \tag{5}$$

where $\bar{\epsilon}$ is the average of the own-price compensated elasticities of the three taxed food groups, and $\bar{\eta}$ is the average of the saturated fat content per kg of fat and cheese, processed meat, sweets and snacks.

Figure F1 shows the distribution of the compensating variation from the ad valorem tax in Euro (a) and as a share of total expenditure (b). Figure F2 shows the distribution of benefits and Figure F3 shows the net consumer welfare impact.

Figure F1: Compensating Variation, ad valorem tax

(a) CV (€/month)



(b) CV/total expenditure

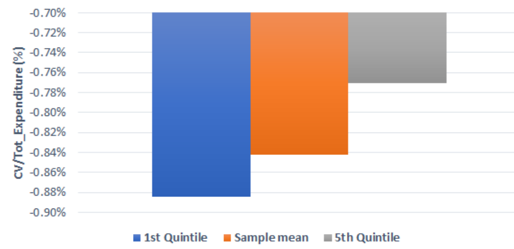
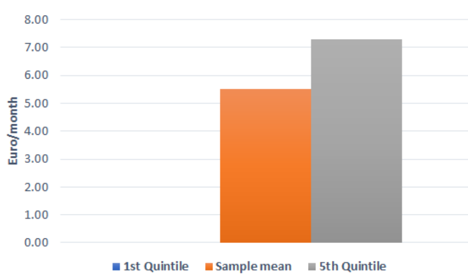


Figure F2: Health benefits, ad valorem tax

(a) Health benefits (€/month)



(b) Health benefits/total expenditure

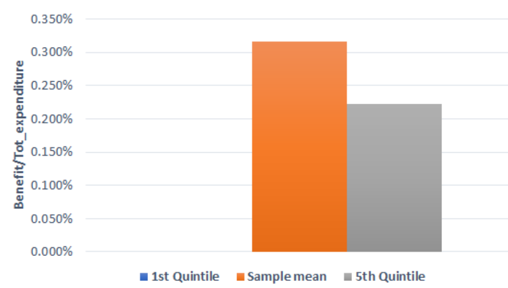
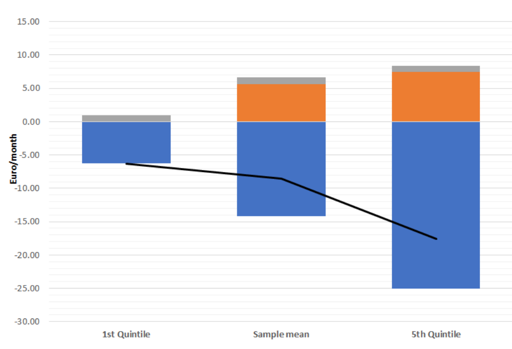
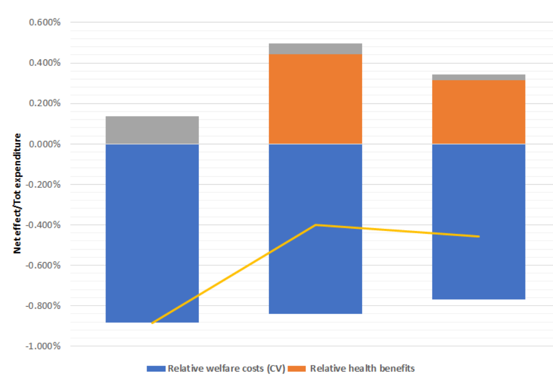


Figure F3: Net Welfare Impacts, ad valorem tax

(a) Benefits and Costs (€/month)



(b) Relative Benefits and Costs



References

Alpman, A. (2016). Implementing rubin's alternative multiple-imputation method for statistical matching in stata. *The Stata Journal*, *16*(3), 717-739. Retrieved from `https://doi.org/10.1177/1536867X1601600311` doi: 10.1177/1536867X1601600311

Hirukawa, M., Liu, D. & Prokhorov, A. (2021). msreg: A command for consistent estimation of linear regression models using matched data. *The Stata Journal*, *21*(1), 123–140.

Hirukawa, M. & Prokhorov, A. (2018). Consistent estimation of linear regression models using matched data. *Journal of Econometrics*, *203*(1), 344–358.

Irz, X. (2017). *Demand for food and nutrients and its climate impact: A micro-econometric analysis of economic and socio-demographic drivers* (Natural Resources and Bioeconomy Studies No. 28). Natural Resources Institute Finland, Helsinki.

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics*, *4*(1), 87–94. Retrieved 2022-05-26, from `http://www.jstor.org/stable/1391390`

Wooldridge, J. W. (2010). *Econometric analysis of cross section and panel data (second edition)*. The MIT Press.