Let them Eat Cake? The Net Consumer Welfare Impact of Sin Taxes

Di Cosmo, Valeria and Tiezzi, Silvia

University of Turin, Department of Economics and Statistics,
University of Siena, Department of Economics and Statistics

10 January 2023

Let Them Eat Cake?
The Net Consumer Welfare Impact of Sin Taxes*

Valeria di Cosmo† Silvia Tiezzi‡

University of Turin University of Siena

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 averted internalities. Using data on Italian consumers we 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 health benefits after the tax. Our results suggest costs from taxation larger than benefits at all income levels. As a fraction of income, the net impact would be regressively distributed.

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

JEL classification: O12, D12, I15.

*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

Corrective taxes, also known as sin taxes, on overconsumed goods, such as cigarettes, alcohol, soft drink and junk food, are applied in various countries (Cawley et al., 2019; Wright et al., 2017) and share the goal of increasing social welfare by curbing consumption and its associated health care costs.

From an economic standpoint, price incentives can be justified as a way to internalize externalities and internalities due to obesity and related chronic diseases, such as cardio-vascular diseases and diabetes (O'Donoghue & Rabin, 2006). In addition, food and beverage taxation based on the content of bad nutrients of concern can generate incentives for reformulating the ingredients. Finally, fiscal revenues from these taxes can be used to fund public health and to compensate for undesired distributional effects from such policies. The possibility of extracting such dividends, i.e. correcting static and dynamic inefficiencies and compensating adverse distributional effects, depends crucially on fiscal policy design, consumer behavioral reactions, and the distribution of consumption of bad nutrients in the population. However, unhealthy food taxes remain controversial policy measures with low political palatability.

One challenge is their appropriate design. In the spirit of Pigovian taxation, the rationale for policy intervention should be based on the presence of market failures or inefficient market outcomes that lead to excess consumption, and the tax rate should be set to reduce consumption to a socially optimal level. Another dilemma is whether to tax the product or the bad nutrient that is causing health problems. Let us consider sugar taxes as an example. Since damage to health from drinks sweetened with sugar is proportional to their "bad" content, (sugar in this case), a sugar tax should be levied directly on sugar content (Grummon et al., 2019) [1]. Linking the tax closely to the source of the external and internal costs should create incentives for firms to reformulate drink ingredients, and for consumers to switch from high-sugar to low-sugar drinks (Griffith et al., 2020). In fact,

[1] A similar design warning applies to other types of sin taxes such as fat taxes (Leicester & Windmeijer, 2004) for example.

sugar-sweetened drink taxes may be designed as volumetric taxes (i.e. fixed tax per liter of soft drink) or tiered volume-based taxes (i.e. approximate sugar-based taxes proportional to soft drink sugar content) (WorldBank, 2020). This design does not provide the maximum possible health benefits, because consumers have no incentive to replace high-sugar with low-sugar drinks and producers have no incentive to change drink ingredients. Another objection is that sin taxes are regressive because the poor consume sin goods disproportionately to the rich, so that the deadweight welfare loss falls more heavily on lower income than on higher income individuals. This is the prediction of many empirical studies focusing on foregone sin-good consumption and on the distribution of welfare costs after a price increase. But it is only part of the story. Allcott et al. (2019b) argue that the full welfare effects of a sin tax can only be evaluated if both costs and benefits of such taxes are considered, where the benefits are averted external and internal costs plus public revenues redistributed by the tax. If lower income individuals consume the most and also reduce consumption the most after a price increase, sin taxes may imply larger benefits for disadvantaged social groups. If larger welfare benefits overcome larger welfare costs, the traditional regressivity argument against taxation of unhealthy foods is undermined. The trade-off between the welfare costs and benefits of sin taxes is therefore pivotal for assessing their net welfare and distributional impact.

Here we estimate the net welfare impacts of a fat tax, accounting for both the monetary value of individuals' health benefits measured as savings from averted internalities and for its welfare costs. There is little evidence of welfare impacts of sin taxes net of benefits (Allcott et al., 2019b); the bulk of the literature focuses either on welfare costs or benefits, and the benefits are seldom measured in monetary terms. In addition, we are the first to provide evidence from Italy, where the government proposed introduction of a sugar tax in 2019.

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. While individual product's household scanner data are increasingly used to estimate rich models of consumer demand (Dubois et al., 2022), no institution offers free access to such large amount of households scanner data for academic research purposes. So we first collect nationally representative pooled cross-sections of Italian households consumption expenditures and associated prices indices. 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 a sin tax.

To compute savings from averted internalities, 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 sin tax, on individual monthly health expenditure. This is our money metric of internality benefits from the sin tax.

We make the following advances with respect to the existing literature. First, unlike studies based on household-level purchase data [2], we use a sample of single households[3] to ensure a unique correspondence between an individual, the recorded expenditure on the

[2] 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.

[3] 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

different food categories, and the costs and the health benefits of taxation. Our data suggests that the usual picture of poor individuals spending more of their budget on unhealthy food and drink than the rich is overturned in this case. It is high income persons who overconsume unhealthy nutrients the most. This distribution of bad nutrient consumption does not change when households with more than one adult, or with one or two children are considered. Second, we do not decide *a priori* which nutrient should be taxed, but adopt a Pigouvian approach to corrective taxation and tax the bad nutrient, consumption of which most exceeds WHO guidelines. Our simulated sin tax is therefore country-specific and targets the level of overconsumption of the various nutrients in the country in question. This makes it possible to account for prevailing social and cultural norms in designing sin tax policies. It turns out that if the objective is to tackle socially costly consumption, a sugar tax (as proposed by the Italian government in 2019) is not the best option for Italy. In the case of sugar, we find no evidence that most individuals exceed official recommendations on how much is acceptable. Saturated fat may be a better target for taxation. We choose the level of taxation necessary to reduce saturated fat consumption to the WHO threshold.

Third, we assess the monetary value of internality benefits associated with the tax by estimating individual health expenditure averted 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. The individual health expenditures measured in the EHIS and used in this study are, therefore, in excess of the SSN coverage and borne by the individual. We consider averted health expenditure in excess of the SSN coverage as our proxy for internality 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). The first part estimates the probability of positive medical expenditures, while the second part estimates the amount of medical expenditures, if any.

Fourth, we estimate sin tax costs and benefits both in absolute terms and relative to income. Our findings, previewed here, suggest a reversal of the hypothesis that sin taxes bring net benefits for lower income individuals (Allcott et al., 2019b; Dubois et al., 2020). Assuming full pass-through of the tax policy, we find that a fat tax aimed at reducing saturated fat consumption by 30% results in a small net welfare cost for the average Italian consumer. In terms of distributional impacts, considering net welfare impacts relative to total expenditure, low-income individuals experience a larger net loss from the fat tax, relative to income, than high-income individuals, suggesting regressive relative net impacts.

Interestingly, we also show that a small increase in the existing value added tax (VAT) on selected groups of foods results in net welfare and distributional impacts very close to those of the nutrient tax based on the saturated fat contents of foods.

Our work contributes to two strands of literature that seek to understand the effects of sin taxes. The first is the empirical literature focused on the impacts of sin taxes using a demand system approach. Estimation of a complete demand system is the ideal setting for computation of theoretically grounded price elasticities of quantities demanded, and for fully accounting the behavioral responses of consumers following a price increase, i.e. reallocation of consumer spending on the entire consumption basket after a price change. The associated monetary metric of the variation in welfare after a price change accounts for said behavioral 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 within food groups when assessing the impact of food and beverage tax policies by estimating a utility-theoretic demand system.

In these studies consumption is measured at household level including that of adults and children. Welfare changes caused by the tax are therefore measured at household level or as per capita averages[4]. Since health benefits are individual-specific, it is not possible to ensure a unique correspondence between purchases, consumption, and the health benefits from taxation for each individual in the household. Indeed, studies in this strand of the literature focus on welfare costs and/or on health benefits at the household level, and the health benefits are not expressed in units of dollars.

We depart from these studies by using single-household data to ensure alignment of purchases, consumption, welfare costs and welfare benefits at individual level, and by estimating health benefits in monetary terms. Allcott et al. (2019b) and Dubois et al. (2020) suggest that, although sin taxes are in generally found to be regressive, the effect could be reduced or eliminated by considering future overconsumption costs averted and future savings in out-of-pocket healthcare costs.

Our study also connects to the set of papers focusing on money metric estimation of health benefits from sin taxes. Recent studies from high-income countries including Australia, Canada and the USA report equal or greater health benefits in monetary terms for lower income groups (Kao et al., 2020; Lal et al., 2017; Wilde et al., 2019). In these studies, the monetary values of health benefits is computed as healthcare expenditures saved, based on the predicted reduction in mortality and morbidity from a set of diseases associated with overconsumption of the unhealthy nutrient targeted for taxation (a proxy for averted externalities), plus the saving in out-of-pocket healthcare costs, i.e. healthcare costs paid for by individuals, including medicines, medical services, medical practitioners and hospital costs, used as a proxy for averted internalities. These studies assume zero substitution between the group of goods targeted for taxation and the other food or beverage groups. In addition, elasticities used to estimate changes in weight and BMI driving the predicted reduction in mortality or morbidity are imported from outside sources. As a result, consumer costs and benefits could be to some extent misaligned. We depart from

[4]An exception is Xiang et al. (2018) who estimate welfare costs of a SSB tax for different household types including single households. However, the demand system here is highly aggregated and does not allow for substitutions within food groups.

this strand of the literature in two ways. First, both own- and cross-price elasticities are estimated for our sample of individuals, and the consumer surplus change associated with the tax takes into account all complementarities and substitutions that occur as a consequence of the introduction of the tax. We measure pure substitutions (or complementarities) by computing compensated cross-price elasticities, net of any income effect. Second, for each individual in our sample, we proxy averted internalities by estimating expected savings in out-of-pocket healthcare costs resulting from predicted weight changes one year after introduction of the tax.

Very few papers have linked 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 for evaluating an optimal sugar-sweetened beverages (SSBs) tax for the USA in the presence of internalities and externalities. Their study accounts the three key elements for evaluating welfare benefits from sin taxes: correcting consumer bias, externalities and revenue recycling through income transfers. Measures of consumer bias are computed adopting a "counterfactual normative consumer" strategy to predict consumption in the case that people had nutrition information from dietitians and nutritionists and perfect self-control. The "Internality Correction" is the increase in (money-metric) welfare due to the change in consumption resulting from the tax. 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. Our ambition is to contribute to this research cluster by providing new empirical evidence about a fat tax and about Italy, a country characterized by the Mediterranean diet, a recognized balanced model of healthy eating rooted in cultural and gastronomic tradition, where adult consumption patterns may reflect these specific social norms.

The rest of the paper is organized as follows. Section 2 describes the different data

sources. Section 3 describes the demand model, the estimation procedure and the elasticities derived. In Section 4 we discuss the welfare costs and distributional implications of our counterfactual sin tax simulations. Section 5 delves into the monetary value of averted internalities and assesses the net consumer welfare and distributional impact of the simulated tax policy. Section 6 concludes.

2 Data

2.1 Expenditures

We use data on household characteristics and food consumption expenditure from the Household Budget Survey (HBS) released by the Italian National Statistical Office (ISTAT) each year[5]. Each annual cross-section includes monthly consumption expenditures of about 23,000 Italian households (the exact number varies from year to year) in approximately 480 Italian municipalities (the smallest Italian administrative territorial grouping). ISTAT uses a weekly diary to collect expenditure data on frequently purchased items and face-to-face interview to collect data on large and durable expenditures. Two weeks in each month are randomly selected. Households sampled each month are divided into two groups of equal number and assigned to one of the two randomly selected weeks. Current expenditures are classified into about 280 elementary goods and services, the exact number changing from year to year due to minor adjustments in the list of items. The survey also includes detailed information on household structure and sociodemographic characteristics (such as regional location, number, gender, age, education and employment condition of each household member). All annual samples are drawn independently according to a two-stage design[6]. Since the survey only provides expenditures at the household level, we select households with only one member to avoid any ambiguity in consumption. A unique correspondence between individuals and recorded expenditure on each food category is of crucial importance here, as our aim is to match costs and benefits of the simulated tax

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

[6]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.

policy for each individual.

The above survey design makes it impossible to track individuals over time and to use panel data. We therefore use the series of independent cross-sections of micro-data for the period January 2014 - December 2018.

Our final sample includes a total of 12,369 single households. The households are classified into 21 regions and three urban types (metropolitan areas, medium-size cities and small cities). The household food consumption module assembles data on expenditure for about 200 items based on 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. The 17 groups are 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 in addition to the residual food items (such as condiments and spices), 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). For most food categories, at least 5% of households in the data did not record any purchase, and those zero values give rise to econometric issues that we discuss in the next section.

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. We report descriptive statistics for the average budget share of each food group and log prices in Table B1, Appendix B.

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 nutrient consumption from these sources. We can, however, disaggregate the budget share of food_afh over total food consumption 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 consumption, 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 (Section 3.2) that although an increase in the price of fat and cheese, sweets and snacks, or sweetened beverages causes substitution towards food away from home, such substitutions are small in magnitude.

2.2 Prices

Since the HBS does not provide information on prices paid by consumers, we use monthly consumer price indices (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 to be matched with our selected HBS expenditure categories. One concern is the small variation in and high collinearity of prices, which occurs in estimations of highly disaggregated demand systems on pooled cross-sectional data. To address this concern and to increase variation in prices, we compute 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 the single goods within each group are weighted with their expenditure shares in the commodity group. Since these shares vary for every household in our sample, price variation increases with the use of Stone-Lewbel prices [7].

2.3 Nutrients

We convert quantities of each food item purchased into its nutritional components (calories, fats and sugar) 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) [8]. This unique database allows us to calculate nutrient values per kilogram of each food group [9]. Table 1 shows the sugar, saturated fat and calorie content per kg of final product. As expected, sugar content is exceptionally high for sweets and snacks and sweetened beverages. Saturated fat is high in fat and cheese, oil, and processed meat.

[7] Figure B.1 in Appendix B shows monthly log price indices over time.

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

[9] 9The 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.

Table 1: Nutrients (g) 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 and saturated fat consumed per day by the individuals in our sample and the reference value, excluding fruit (Griffith et al. (2016)).

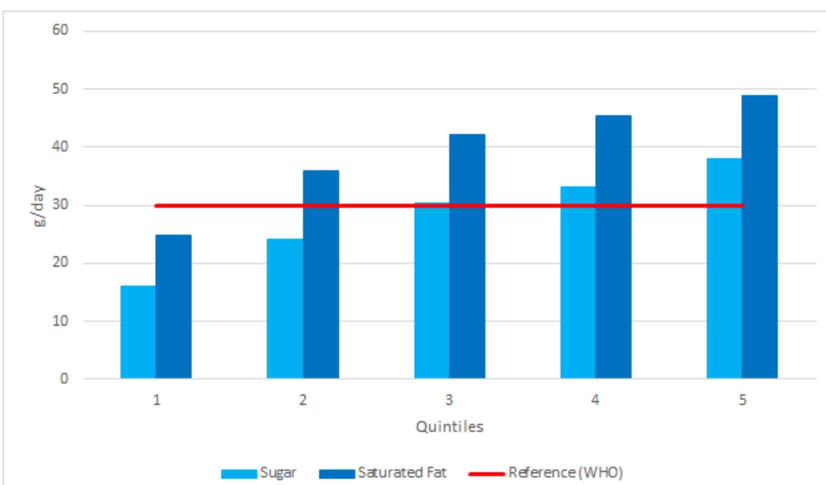


Figure 1: Sugar and saturated fat overconsumption by income quintile

Figure 1 shows that there is no overconsumption of sugar for individuals below the fourth quintile of the income distribution. The sugar consumption sample average (28 g/day) is also below the threshold recommended by the WHO (30 g/day). In 2019, led by the 5-star Movement, the Italian Government proposed a tax of about 3 cents for a 330 ml can of soft drink of average sugar content, a figure similar to that applied in France and Great Britain. The Italian sugar tax never came into force, but it was unlikely to promote reformulation of drink ingredients. Our data also fails to show any excess consumption of added sugar among single adults. The Italian measure, as conceived in 2019, would certainly have raised cash, but would have been of little use for reducing the amount of sugar per can or for protecting health.

By contrast, the average intake of saturated fats is almost 40 g/day, exceeding the 30 g/day threshold recommended by the WHO by about 33%. Excess consumption increases with income, individuals in quintiles above the first displaying larger consumption excesses. This data suggests that in Italy the usual situation of lower income individuals showing higher consumption of unhealthy nutrients than higher income people is inverted. Although Figure 1 refers to one-member households, very similar patterns hold for larger households. Figure A1 in Appendix A shows the distribution of consumption of harmful nutrients in households with two adults, two adults and one child, two adults and two children. Although we cannot assess the intra-household allocation of consumption, Figure A1 suggests that in line with the distribution of consumption in single households studied here, families in higher quintiles of the income distribution consume more unhealthy nutrients than those in lower quintiles.

2.4 Health expenditures and body weight

Data on individual body weight in our sample comes from the 2015 Italian module of EHIS, a survey on the health of the population of EU member states conducted every four years. To match health expenditure from HBS data with data on individual body weight, 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 EASI (Lewbel & Pendakur, 2009) demand system with 16 food groups and a composite numéraire that incorporates all other consumption goods and services plus a residual food category. The estimated parameters of incomplete demand systems can be used to provide exact and correct measures of changes in welfare, unlike those of conditional demand models (LaFrance & Hanemann, 1989; Hanemann & Morey, 1992). Conditional demand systems also 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.

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.

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. Tractable multi-stage estimation procedures of censored demand models have been developed since Heien & Wesseils (1990) proposed their two-steps estimation. In the first step, a probit equation is estimated to model the binary decision to consume an item and, in a second step, the demand equations are augmented by the inverse Mills ratios extracted from the first-step regressions. Shonkwiler & Yen (1999) pointed out at an internal inconsistency of the Heien & Wesseils (1990) model and proposed the alternative two-step procedure that we adopt here.

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 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 where z_1 is a constant; 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 equations linear-in-the-parameters. This property allows us to circumvent the difficulties of simultaneously accommodating censoring of demands and estimation of non-linear 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. Lewbel & Pendakur (2009) show that the linearized version of the model estimated by OLS performs almost as well as fully-efficient nonlinear estimation.

To correct for endogeneity due to the introduction of budget shares into log real total expenditure, in line with Lewbel & Pendakur (2009), 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: $\hat{y} = lnx - \sum_{j=1}^{J} \bar{w}^j lnp^j$.

In addition to a constant, we specify the vector of demand shifters z_k to include the following binary and categorical 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 5 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.

3.1 Estimation and elasticities

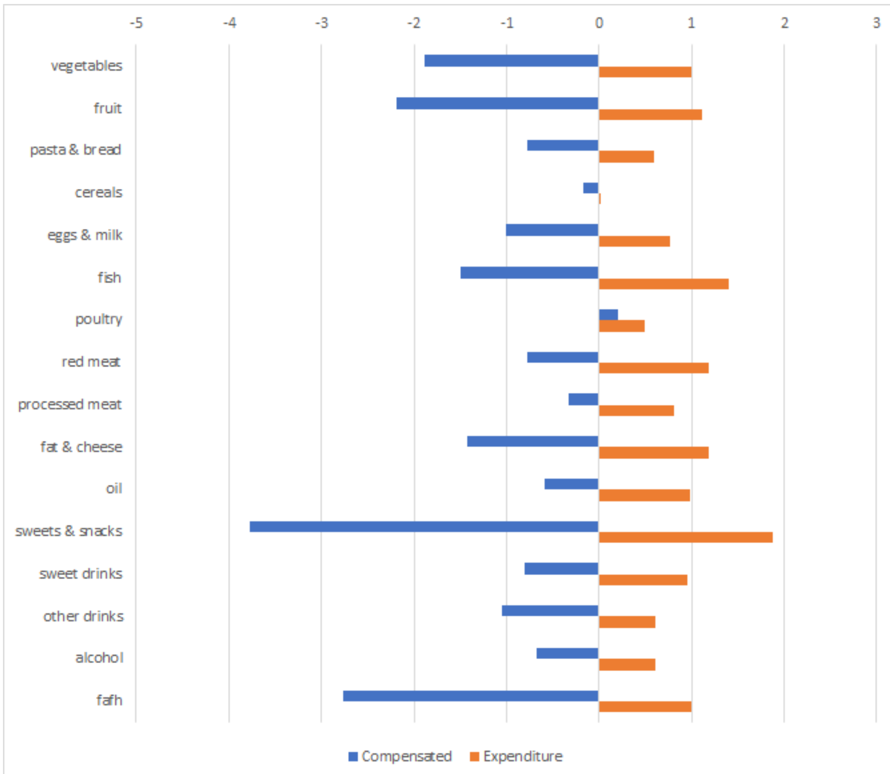
We estimate the model in (1) using seemingly unrelated regression methods. Figures C1 and C2 in Appendix C plot the Engel curves for our 16 commodities. Inspection of these Figures suggests that the Engel curve shapes cannot be adequately represented by a linear or quadratic function. To determine the degree of the income polynomials, we add a degree at a time starting from $L = 2$ and tested 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.

Behavioral reactions after a price change are measured by own and cross-price elasticities. Cross-price elasticities, in particular, highlight substitutions and complementarities among food products, i.e. changes in the quantities of other food products purchased after a price change. 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.

Our structural model estimates lead to a 16x16 matrix of 256 estimated price elasticities. Table D1 and Table D.1 in Appendix D show the full set of compensated (Hicksian) and uncompensated (Marshallian) elasticity point estimates, respectively. The full set of standard errors is available from the authors. Figure 2 focuses on own-price compensated elasticities and expenditure elasticities. Both are reported for each food group. The top bar in each aggregate indicates own-price compensated elasticity and the bottom bar, expenditure elasticity.

All estimated own-price elasticities have the expected negative sign and all except two are statistically significant at 1%. Among the food groups, sweets and snacks show the largest own-price elasticity (-3.770), implying that a 1% increase in their price would decrease the quantity purchased by about 3.8%. The quantity of fat and cheese purchased

Figure 2: Own-price compensated and expenditure elasticities.



is also elastic to its price (-1.432). The own-price elasticity for sweetened beverages is -0.805, which falls in the range -0.8 to -0.10 of the literature (Finkelstein et al., 2010). Studies such as Allcott et al. (2019a), focusing only on soft drinks, find higher own-price elasticities (-1.37). Expenditure elasticities are all positive and significant at 1%, except for cereals. Most food groups are necessities with an expenditure elasticity less than one. Fruit, fish, fat and cheese, red meat, sweets and snacks are luxuries with an expenditure elasticity greater than one. Table D1 shows the full set of compensated (Hicksian) and expenditure (last row) elasticities at the sample mean. Since compensated elasticities are utility-constant, cross-price elasticities provide information on net complementarities or substitutions, i.e. on the percentage change in the quantity demanded of good i after an increase in the price of good j, net of any income effect.

In other words, our elasticities in Table D1 measure pure substitution after a price change, which is the information we are looking for. 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. Cross-price elasticities can be read in the last row of the table. For example, the third entry in the first column (0.228) is the percentage change in the demand for bread and pasta 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 that arise when the price of fat and cheese, sweets and snacks and sweetened beverages increases, as these are the food groups most affected by our counterfactual fat tax. An increase in 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). An increase in the price of sweetened beverages causes substitution with fruit, fish, food_afh (0.096, 0.019, 0.181) and complementarities with oil, sweets and snacks, other drinks (-0.160, -0.112, -0.160). An increase in the price of fat and cheese causes substitution with bread and pasta (0.091) and complementarity with fruit (-0.064). We also find substitution between poultry and processed meat, between poultry and vegetables (0.133) and poultry and fat and cheese (0.098). 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 nutrient content of the latter. Inspection of Table D1 shows substitution of fat and cheese, sweets and snacks, and sweetened beverages with food afh, but these substitutions are small (0.032, 0.181, and 0.023, respectively). Finally, increasing the price of red meat reduces the quantity of eggs and milk, poultry, processed meat and alcohol purchased (-0.226, -0.609, -0.177, -0.213). Similar complementarities arise when the price of processed meat is increased with an additional complementarity with other drinks (-0.293).

Table 2: Compensated price elasticities and expenditure elasticities - 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 bootstrapped with 200 replications.

Table D2 in Appendix D shows compensated own-price elasticities and expenditure elasticities at low and high levels of total current consumption expenditure, our proxy for income. Low income individuals react more (larger own-price elasticities) than high income persons after a price change in vegetables, bread and pasta, cereals and rice, fish, red meat, sweets and snacks, sweetened beverages, food afh and fruit. Interestingly, however, we find that high income individuals, who are the highest consumers of fat and cheese, show greater own-price elasticity for fat and cheese (-1.633 against -1.088 for low income persons). Since this is the food aggregate highest in saturated fats, the narrative that poorer individuals consume more bad nutrients and react more than the rich to a price increase is inverted in Italy.

Expenditure elasticities identify luxuries and necessities. We do not find much variation in relation to income distribution. The only exception is cereals and rice which are luxury items for high-income (expenditure elasticity 1.203) but not for low-income consumers. Red meat and soft drinks are luxuries for low-income but not for high income individuals. Tables D3 and D4 in Appendix D show compensated own-price elasticities and expenditure elasticities, respectively, both by macro-region.

4 Counterfactuals

In the main counterfactual experiment we use our demand estimates to simulate the introduction of a specific (s) tax (τ) that is 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% fall in saturated fat purchased assuming a 100% pass-through of taxes to prices[10]. For each commodity (i.e. food group) $j, j = 1, ..., J$, the specific tax on saturated fat is:

$$\tau\eta^j = \frac{-0.30}{\epsilon^j} p_0^j \eta^j \tag{4}$$

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

We also separately simulate an easy to implement and to 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[11]. 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.

[10]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%

[11]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 away from home	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 that the variation in saturated fat consumption (grams) per month after application of the tax varies 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. Figure 3 shows that the fat tax achieves relatively large reductions in fat consumption among individuals with average (orange column) and high levels of total expenditure (grey columns), but it is not successful at targeting individuals in the lowest quintile of the expenditure distribution (blue columns). 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. Across the age distribution, young consumers are equally likely to be affected by the policy as adults.

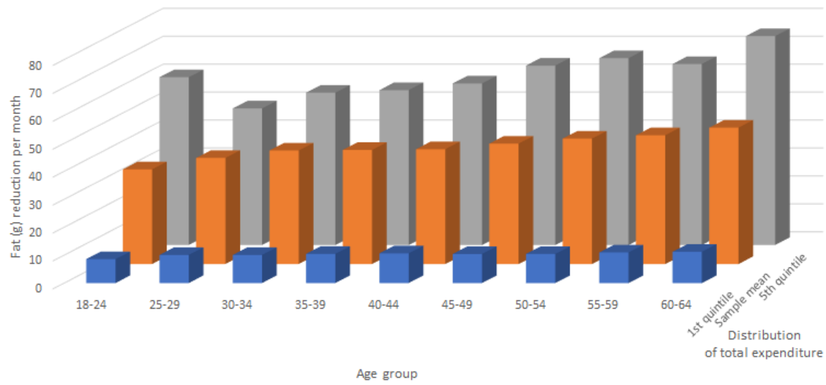


Figure 3: Reduction in saturated fat consumption (grams)

4.1 Consumer-Welfare Costs and Redistribution

We use our demand estimates to compute the compensating variation (CV), a monetary 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

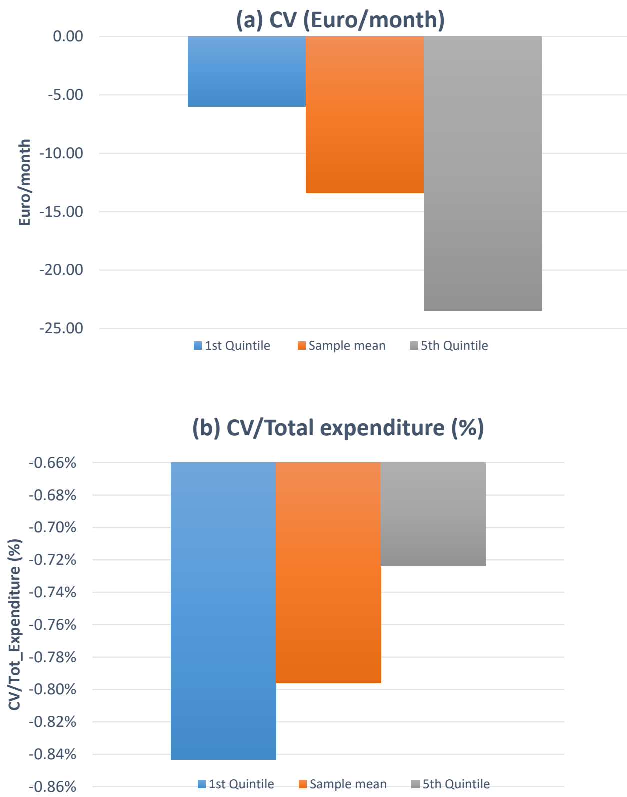
$J \times 1$ vector of new log prices after the tax is imposed, and Γ is a $J \times J$ matrix of parameters whose element Γ_{ij} equals a^{jk} in equation 1. Equation 5 captures two welfare effects of the fat tax on welfare. The first term on the right-hand-side is the Stone price effect that ignores any changes in budget shares of the taxed goods. By using observed rather than predicted budget shares, any unobserved heterogeneity absorbed into the error term during estimation is incorporated in the welfare analysis (Lewbel & Pendakur, 2009). 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. The welfare loss is progressively distributed since it increases with income. At the mean income, CV per month is 13.40 €. Relative to income, proxied by total monthly consumption expenditure, the welfare loss has a mildly regressive distribution. The CV provides a money value of the welfare cost to consumers resulting from the tax. However, if some consumers impose internalities on themselves, the compensating variation, based on revealed preferences, provides an incomplete picture of the welfare effects of the tax (Gruber & Koszegi, 2004).

5 Monetary value of averted internalities

One potential consequence of excess saturated fat consumption is weight gain. We proxy the monetary value of averted internalities with the value of health benefits associated with weight loss. The literature on the relationship between weight loss and health benefits calculates weight reduction after tax, starting from the variation in harmful nutrient intake after imposition of the tax (Hall et al., 2011; Lin et al., 2011; Harkanen et al., 2014; Xiang et al., 2018).

This is shown in Figure 3 as a function of age and across the distribution of total consumption expenditure (our proxy for disposable income). As explained in the previous Section, the tax achieves relatively large fat reductions among those individuals with an

Figure 4: Compensating Variation (CV)



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. to be larger at higher incomes. The impact of the tax on food consumption can be calculated by multiplying the matrix of uncompensated demand elasticities (Table D1 in Appendix D) by a vector containing the percentage changes in consumer prices. Table 4 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.

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)

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

with half the weight change achieved in about 1 year and 95% in about 3 years[12]. Table 4 shows the average reduction in body weight (in kg) one year after introduction of a fat tax aimed at reducing saturated fat consumption by 30% and the average change in energy intake (kJ/day). We obtain an average body-weight loss of 1.72 kg one year after 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 Cawley & Meyerhoefer (2012). The first part of the two-part model estimates the probability of positive health expenditure, while the second part estimates health expenditure, if any.

[12]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.

Monthly health expenditure at 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. The HBS data also includes expenditure on prescription and non-prescription drugs and sanitary articles such as medicines, plasters, syringes, first aid kits, bandages and the like, vitamins, minerals and homeopathic products. Health expenditure is not distributed evenly across respondents. In particular, health expenditure for the first quintile of the expenditure distribution 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, the first quintile of the income distribution spends less on healthcare than 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 Survey Interview (EHSI) by the matching procedure of Rubin (1986). Table E.2 in Appendix E shows the descriptive statistics for the variables resulting from the matching, used in our empirical analysis.

Let he_i denote monthly health expenditures (in Euro) by household i. Let α be the constant term and let $\mathbf{X}_i$ denote a vector of explanatory variables including the age (9 classes), gender, education level (5 classes), employment position (4 classes: 1= manager, 2= unskilled worker, 3= entrepreneur, 4=self-employed); employment status (8 classes), marital status (5 classes), macro-region (5 classes), income quintile and weight (kg) of each individual. Let ε_i be the idiosyncratic error term. Our base regression specification for estimating the marginal impact of weight on health expenditures is:

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

Table 5 lists regression results for the sample resulting from the matching. The cells of the table indicate the marginal effects (reflecting both parts of the two-part model) and the standard errors of the marginal effects at the sample mean, for the first quintile, and for the fifth quintile of the expenditure distribution. The table indicates that weighing an additional kilogram raises health expenditure by 0.01 € per month (which is not statistically significant) for individuals in the first quintile of the expenditure distribution, by almost 3€ per month on average, and by 4.30 € 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

Variable	**Sample mean**		**1st quintile**		**5th quintile**	
	Probit	GLM	Probit	GLM	Probit	GLM
Weight	0.02*	2.923**	0.020***	0.010	0.022*	4.307***
	(0.009)	(0.821)	(0.005)	(0.375)	(0.011)	(1.054)
Sex (1=male)	0.766***	66.903***	0.814***	-0.820	0.826***	108.939***
	(0.155)	(14.024)	(0.145)	(10.319)	(0.229)	(20.956)
Age group	0.058***	3.797***	0.027	4.814**	0.068***	5.089***
	(0.01)	(1.324)	(0.030)	(1.865)	(0.020)	(3.054)
Location	0.053***	-5.657**	0.125***	-6.371**	-0.009	-11.800**
	(0.013)	(1.924)	(0.043)	(2.921)	(0.027)	(4.371)
Income quintile	0.339***	34.380***				
	(0.015)	(2.131)				
Education	-0.068**	-4.662	-0.134	-6.263	-0.030	-6.192
	(0.025)	(3.437)	(0.093)	(3.517)	(0.047)	(4.305)
Marital status	-0.024	-5.589**	-0.020	-0.853	-0.019	-12.912**
	(0.016)	(2.020)	(0.053)	(3.912)	(0.027)	(4.630)
Empl. position	-0.028*	-7.545***	-0.087	-1.667	-0.041	20.492**
	(0.018)	(2.043)	(0.058)	(6.881)	(0.027)	(17.155)
Empl. status	-0.027	-8.658**	-0.221*	0.000	0.155*	0.000
	(0.066)	(6.542)	(0.112)	(0.000)	(0.112)	(0.000)
Constant	-3.554***	-306.319***	-2.723***	16.812	-2.374**	-265.869**
	(0.769)	(73.764)	(0.640)	(42.900)	(1.033)	(104.761)
#obs	7,781	7,781	640	640	2,499	2,499

$* = p < 0.10$; $** = p < 0.05$; $*** = p < 0.01$. 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 that of those in the highest quintile. Interestingly, in the first-step probit regression, the weight coefficient is significant at 5% and positive for the first and fifth quintiles, with an implied elasticity close to 0.07 for all groups. So the probability of positive health expenditure increases for individuals in the lowest quintile if they gain an extra kilogram, even if they do not benefit from a one-kg reduction in the case that 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: no benefits emerge for low income individuals. Benefits are progressively distributed, with high income individuals benefiting more than individuals at the sample mean of the expenditure distribution.

5.1 Net Welfare Impacts

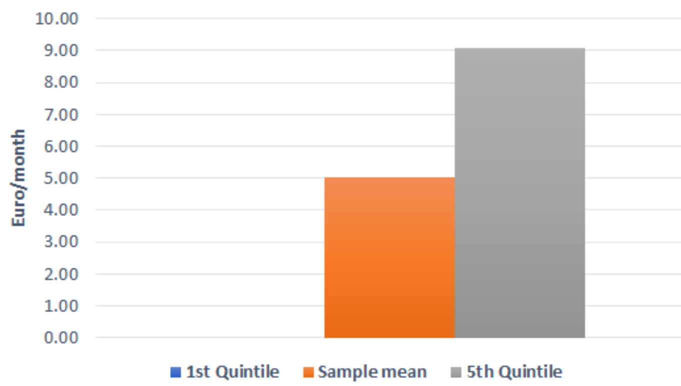
We combine the results of Section 4 with the empirical estimates of the monetary value of internalities averted to compute the net welfare impacts from the simulated tax.

The welfare effects can be decomposed into three distinct components: they are plotted in Figure 6a across the distribution of total expenditure. "Redistributed Revenues" are public revenues from the fat tax equally redistributed as lump-sum transfers. "Internality Correction" is the (money-metric) welfare benefit due to the weight loss resulting from 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 introduction of the tax indifferent to consumers. "Net Welfare Impact" is the difference between welfare costs and benefits.

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 but regressive net benefits. The lump sum returns only marginally offset the welfare costs of the tax.

Figure 5: Health benefits

(a) Health benefits (€/month)



(b) Health benefits/total expenditure

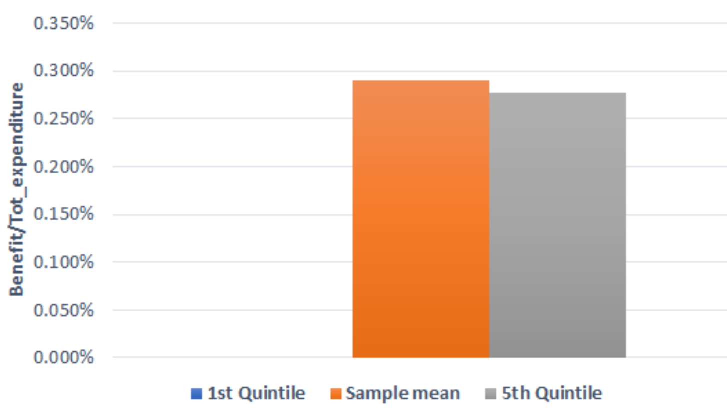
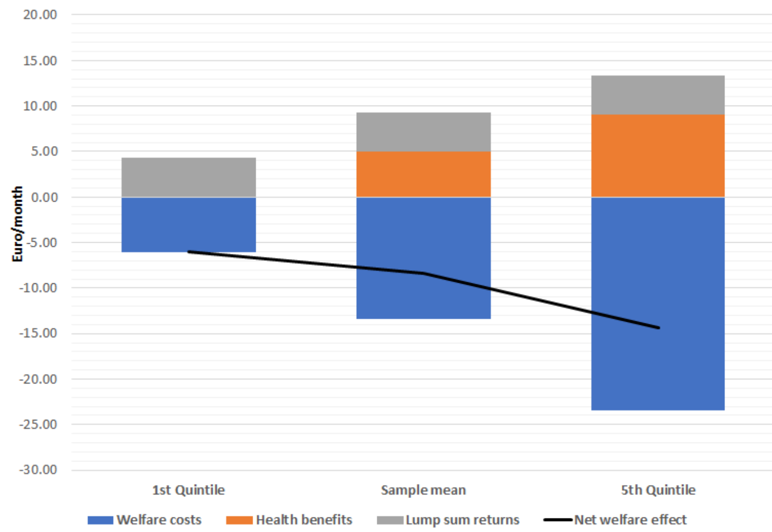


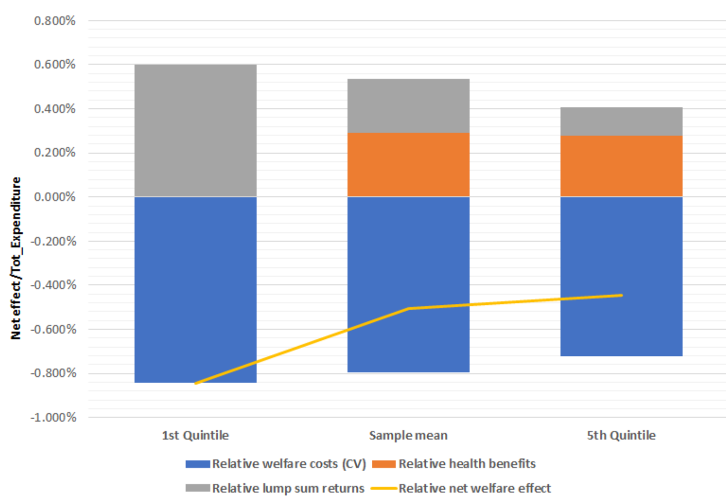
Figure 6b shows costs and benefits as a fraction of total expenditure, our proxy for income. Relative to total expenditure, the fat tax generates small and regressively distributed welfare losses, in line with Allcott et al. (2019a). In order to check whether an easy-to-implement ad valorem tax would 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). The results are shown in Appendix

Figure 6: Net Welfare Impacts

(a) Benefits and Costs (€/month)



(b) Benefits and Costs/total expenditure



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.

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. The net welfare effects from the two tax policies are very similar.

6 Summary and Conclusion

Modern economies often rely on excise taxation to reduce socially costly consumption. While it is quite common to investigate the potential welfare costs of such new tax policies, it is much less common to assess the welfare impacts net of potential benefits. This is, however, an important step in welfare change evaluations, as benefits assessment may enhance the social and political acceptability of the new tax policies. Our paper adds to the very scant literature by endeavoring to assess the net welfare impacts of taxes on unhealthy foods. By focusing on a fat tax, we investigate its individual benefits, proxied by averted internalities, and its individual costs.

We restrict the analysis to single adults to ensure close alignment of expenditure, consumption, costs and benefits of taxation. We focus on Italy, where a sugar tax was proposed in 2019. Our results predict that high-income individuals would lower their fat consumption more than persons of low income in response to a fat tax. As a consequence, individuals in the highest quintile of the expenditure distribution are the ones who lose out most in terms of direct consumer surplus loss due to the tax, but they are also the ones who benefit most from the tax. The prediction that sin taxes bring net benefits, especially for those with lower incomes, is inverted in this case. Instead we find that an Italian fat tax would result in progressive (i.e. larger for high-income individuals) net losses.

These findings should be taken with caution. One limit of our study is that we only consider obesity a potential consequence of excess saturated fat consumption. We ignore disability-adjusted life-years averted. The effects on other health outcomes, such as type 2 diabetes and cardiovascular disease, and the secondary effects of obesity, such as cancer and arthritis, are not modelled. Another limit is that we do not count savings from any averted externalities (such as lower public costs of healthcare). So our estimated benefits from sin taxation should be regarded as a lower bound of the true benefits.

Finally, we only consider single adults, in spite of the fact that child obesity is an increasing concern in Italy (Crudu et al., 2021).

Despite these weaknesses, we trust that the results of our study, together with those of other researchers, will help shift the discussion of sin taxes from mere welfare-cost calculation to a more comprehensive assessment.

7 Acknowledgements

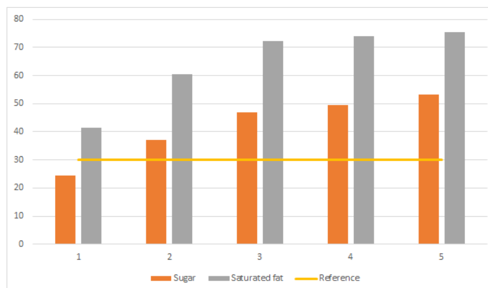
We thank Daniel Stoehlker and seminar participants at the University of Turin (I), and at the 2022 Italian Economic Association Meeting for helpful comments. Declaration of interest: None

Appendix

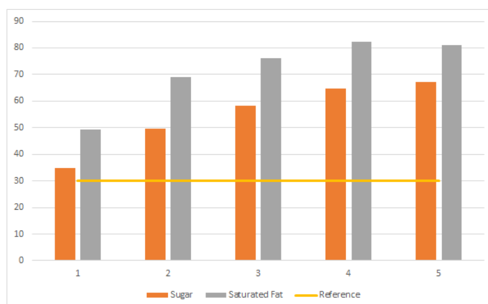
A Excess consumption of sugar and fat

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 more than one member (i.e. two adults, two adults and one child, two adults and two children) across income quintiles. For each household type, excess consumption of added sugar and saturated fats increases along the distribution of income, consistently with what we observe for singles.

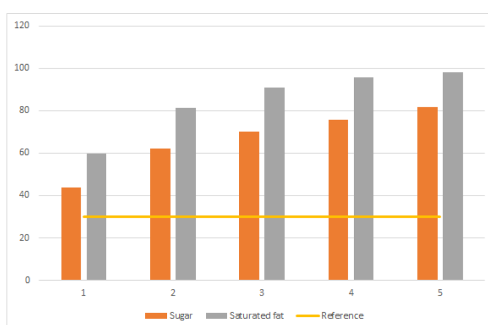
Figure A1: Consumption of sugar and saturated fats in households of different sizes



Two adults



Two adults, one child



Two adults, two children

B Descriptive Statistics

Table B1: Summary Statistics

Variable	Obs.	Mean	Std.Dev.	Min	Max
Exp Share					
Alcohol	12,369	0.010	0.018	0	0.223
Bread	12,369	0.009	0.010	0	0.191
Cereals & rice	12,369	0.003	0.005	0	0.092
Eggs & Milk	12,369	0.010	0.010	0	0.133
Fat & Cheese	12,369	0.012	0.012	0	0.176
Fish	12,369	0.013	0.018	0	0.278
Fafh	12,369	0.051	0.064	0	0.671
Fruit	12,369	0.015	0.014	0	0.183
Oil	12,369	0.005	0.009	0	0.243
Other	12,369	0.816	0.105	0.240	1
Otherdrinks	12,369	0.010	0.010	0	0.309
Pasta	12,369	0.004	0.006	0	0.076
Processed meat	12,369	0.012	0.014	0	0.159
Poultry	12,369	0.007	0.010	0	0.145
Red meat	12,369	0.014	0.018	0	0.203
Sweet drinks	12,369	0.004	0.006	0	0.102
Sweets & snacks	12,369	0.017	0.015	0	0.180
Vegetables	12,369	0.023	0.021	0	0.267
Log Prices					
Alcohol	12,369	-0.658	0.271	-1.609	0.889
Bread	12,369	0.004	0.004	-0.001	0.014
Cereals & rice	12,369	-3.437	0.272	-4.256	-2.773
Eggs & Milk	12,369	-1.670	0.188	-2.867	1.240
Fat & Cheese	12,369	-0.818	0.177	-1.686	-0.276
Fish	12,369	-1.190	0.227	-1.991	-0.609
Fafh	12,369	0.191	0.229	-0.797	0.826
Fruit	12,369	-0.400	0.243	-1.589	0.112
Oil	12,369	2.658	0.309	-3.636	-1.568
Other	12,369	3.075	0.269	1.737	4.114
Otherdrinks	12,369	-1.996	0.192	-2.639	-1.596
Pasta	12,369	0.002	0.014	-0.013	0.041
Processed meat	12,369	-1.472	0.169	-2.037	-0.967
Poultry	12,369	0.010	0.013	-0.004	0.038
Red meat	12,369	-1.162	0.186	-1.980	-0.688
Sweet drinks	12,369	-2.779	0.253	-3.494	-1.957
Sweets & snacks	12,369	0.710	0.315	-1.026	1.391
Vegetables	12,369	1.623	0.534	-0.757	3.048

Table B2: Summary Statistics cont'ed

Variable	Obs.	Mean	Std.Dev.	Min	Max
Control variables					
Total monthly expenditure	12,369	1869,209	1090.125	110	9697.53
Gender	12,369	1.465	0.499	1	2
Education	12,369	3.831	0.828	1	5
Marital status	12,369	2.315	1.786	1	6
Employment position	12,369	2.011	1.971	1	8
Age	12,369	6.008	2.209	1	9
Metropolitan area	12,369	0.154	0.361	0	1
Medium city	12,369	0.294	0.456	0	1
Small city	12,369	0.551	0.497	0	1

Table B3: Share of food expenditures by education level.

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
Fafh	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

Table B4: Share of food expenditures by geographic area.

Expenditure share	North	Centre	South	Islands
Alcohol	0.011	0.009	0.011	0.009
Bread & pasta	0.011	0.014	0.016	0.017
Cereals & rice	0.002	0.003	0.004	0.003
Eggs & Milk	0.009	0.010	0.012	0.011
Fat & Cheese	0.012	0.012	0.014	0.011
Fish	0.009	0.013	0.016	0.019
Fafh	0.061	0.046	0.038	0.044
Fruit	0.013	0.016	0.017	0.017
Oil	0.004	0.004	0.005	0.006
Other	0.840	0.820	0.779	0.793
Otherdrinks	0.008	0.009	0.012	0.014
Processed meat	0.011	0.013	0.015	0.012
Poultry	0.006	0.007	0.010	0.08
Red meat	0.011	0.016	0.018	0.018
Sweet drinks	0.004	0.009	0.004	0.005
Sweets & snacks	0.016	0.015	0.019	0.016
Vegetables	0.020	0.024	0.027	0.025
N	5954	2434	3148	833

Table B5: Share of food expenditures by gender

Expenditure share	Female	Male
Alcohol	0.006	0.014
Bread & pasta	0.013	0.013
Cereals & rice	0.003	0.003
Eggs & Milk	0.011	0.009
Fat & Cheese	0.013	0.012
Fish	0.013	0.012
Fafh	0.036	0.064
Fruit	0.016	0.014
Oil	0.005	0.004
Other	0.820	0.815
Otherdrinks	0.010	0.009
Processed meat	0.012	0.013
Poultry	0.007	0.007
Red meat	0.014	0.015
Sweet drinks	0.004	0.004
Sweets & snacks	0.018	0.016
Vegetables	0.025	0.021
N	5756	6613

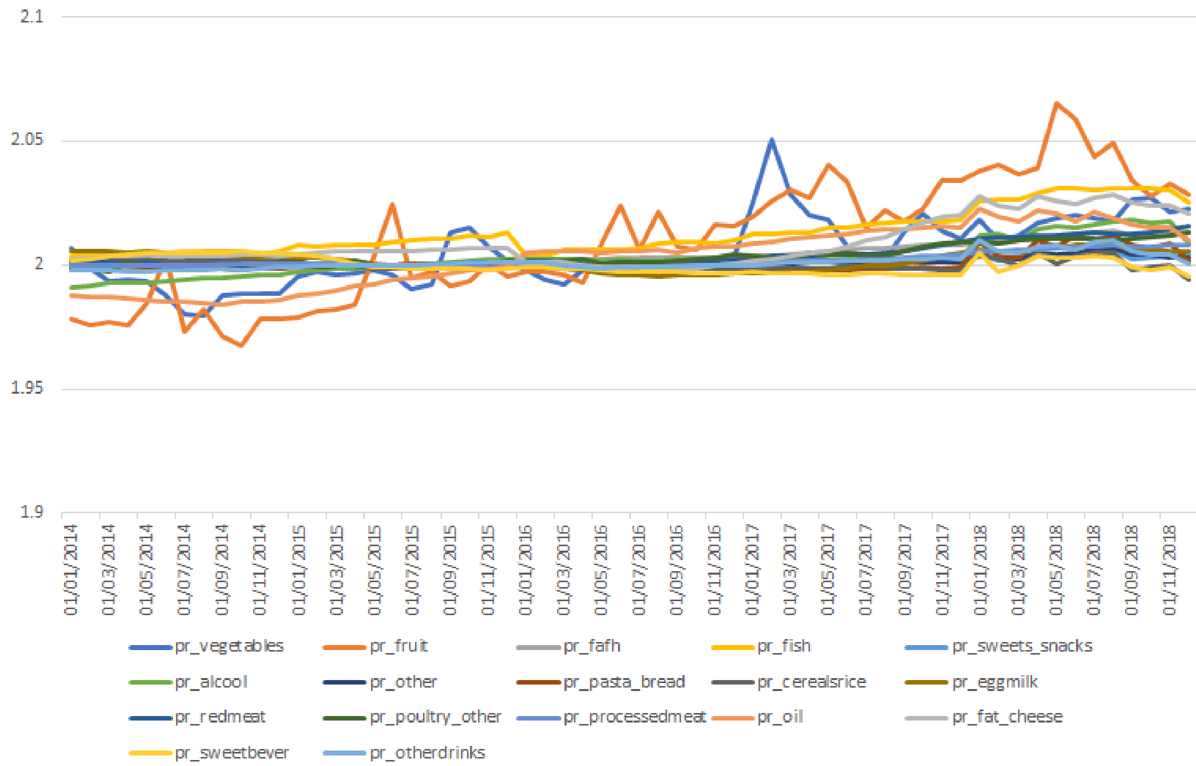
Table B6: Share of food expenditures across the distribution of total expenditure

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
Fafh	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
N	2474	7421	2474

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.

As shown in Figure B1, for some food groups the series of monthly price indices display little variation over time. This differential price variation over time coupled with no cross-sectional variation provides our motivation for using Lewbel prices.

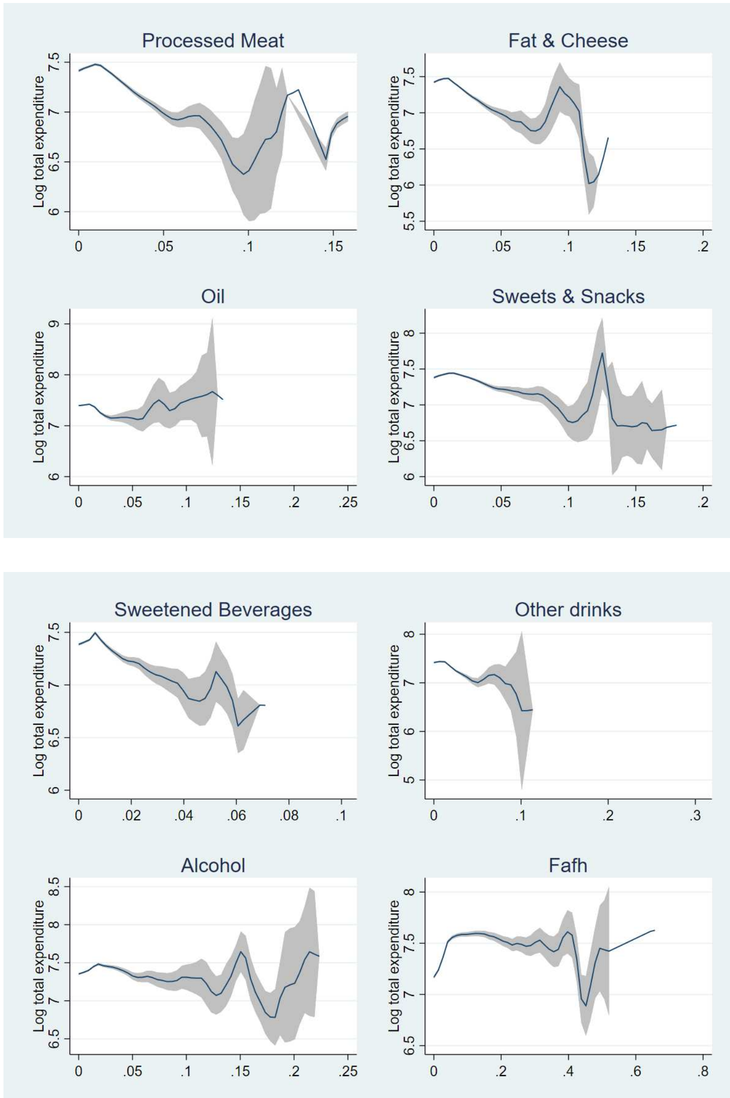
Figure B1: Monthly price indices, logs (2014-2018)



C 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):

$$\frac{\partial lnq^i}{\partial lnp^j} = \frac{a^{ij}}{w^i} + \bar{w}^j - \delta_{ij} - w^j \left[\sum_{r=1}^{R} b_r^i r \left(\hat{y} \right)^{r-1} + \frac{1}{w^i} + 1 \right] \tag{7}$$

$$\frac{\partial lnq^i}{\partial lnx} = \left[\sum_{r=1}^{R} b_r^i r \left(\hat{y} \right)^{r-1} \right] \frac{1}{w^i} + 1 \tag{8}$$

$$\left. \frac{\partial lnq^i}{\partial lnp^j} \right|_{\bar{u}} = \frac{a^{ij}}{\bar{w}^i} - \delta_{ij} + \bar{w}^j \tag{9}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise[13]. Standard errors of elasticities are bootstrapped with 200 replications.

[13]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 - 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***
Fafh	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

	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
Fafh	-2.569***	-0.179	-2.361***	-0.348
N	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

	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***
Fafh	-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
Sweet drinks	-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

	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***
Fafh	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
Sweet drinks	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 weight of each individual is included in the European Health Survey Interview (EHSI) 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.

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 EHSI 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 EHSI 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 in the estimated standard errors. [14] Consistently with our main empirical analysis, we consider single households aged less than 65 years. Summary statistics of both the initial

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

and matched datasets are shown below.

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

Variable	Obs	Mean	Std. dev.	Min	Max
EHSI					
Weight (Kg)	1,964	71.79	14.02	40	130
Gender (1=male)	1,977	1.44	0.50	1	2
Age	1,977	5.82	2.25	1	9
Geographic area	1,977	2.56	1.29	1	5
Income quintile	1,977	3.49	1.34	1	5
Employment position	1,474	2.38	1.28	1	4
Education	1,977	3.87	0.84	1	5
Marital status	1,977	1.90	1.23	1	4
Employment status (1=employed)	1,977	3.49	1.34	1	5
HBS					
Health expenditures (Euro)	12,419	61.14	132.31	0	2855.47
Gender (1=male)	12,419	1.47	0.50	1	2
Age	12,419	6.00	2.21	1	9
Geographic area	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	1.90	1.26	1	4
Employment status (1=employed)	12,419	3.26	1.39	1	5

Table E2: Descriptive statistics, matched dataset

Variable	Obs	Mean	Std. dev.	Min	Max
Imputed health expenditures (Euro)	14,159	61.67	126.23	0	2855.47
Imputed weight (Kg)	8,774	71.63	12.62	40	130
Gender (1=male)	14,396	1.46	0.50	1	2
Age	14,396	5.98	2.22	1	9
Geographic area	14,396	2.61	1.31	1	5
Income quintile	14,396	3.29	1.39	1	5
Education	14,396	3.84	0.83	1	5
Marital status	14,396	1.90	1.25	1	4
Employment position	13,081	1.96	1.06	1	4
Employment status (1=employed)	14,396	0.74	0.44	0	1

As shown in Table E2, the original EHSI dataset has 1964 observations on weight of individuals under 65. The Rubin procedure adds 6810 new observations for which a matching with the HBS is possible, which leads to 8774 observations on imputed weight in the final dataset. For health expenditures the original HBS dataset has 12419 observations, increased to 14159 by the matching algorithm (which found 1740 matches). The final dataset with health expenditures, weight and a common set of control variables contains 8774 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:

$$p^j_{1,av} = p^j_0(1 + t\eta^j) \tag{10}$$

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:

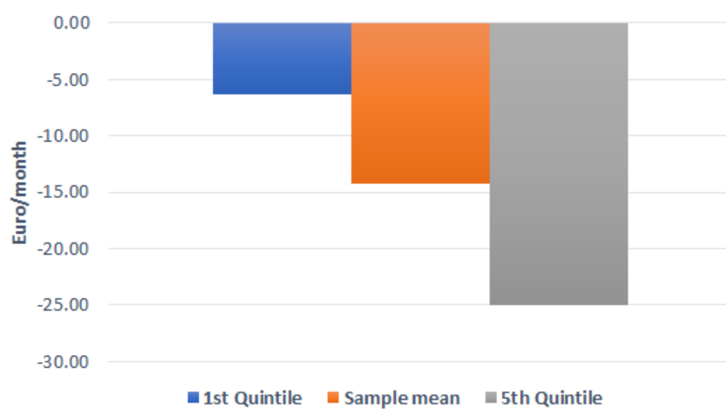
$$t\bar{\eta} = \frac{-0.30}{\bar{\epsilon}}\bar{\eta} \tag{11}$$

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

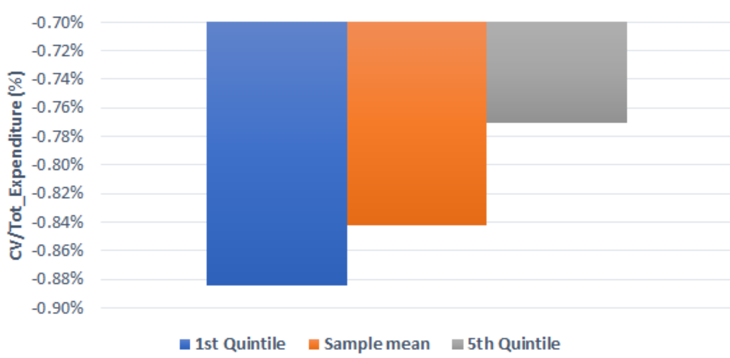
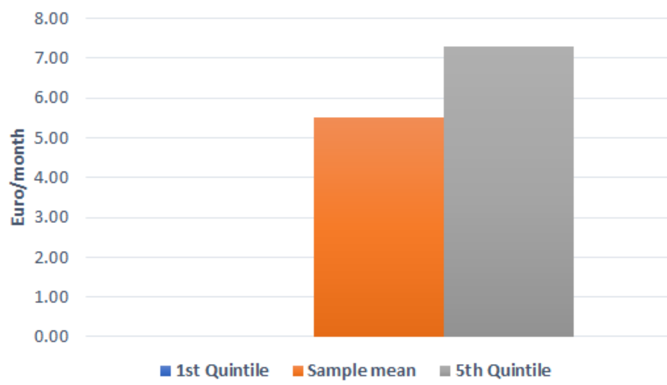


Figure F2: Health benefits, ad valorem tax

(a) Health benefits (€/month)



(b) Health benefits/total expenditure

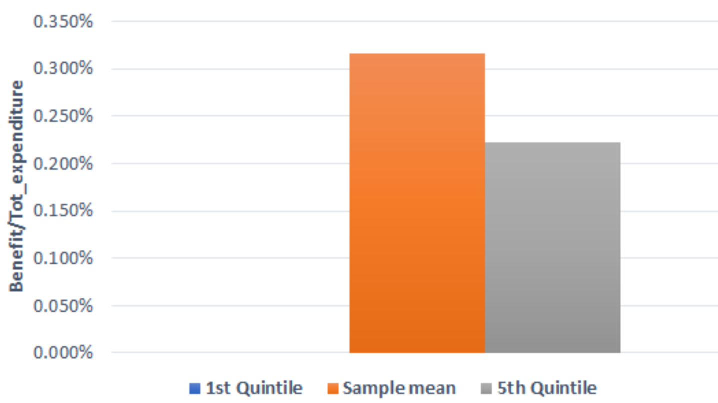
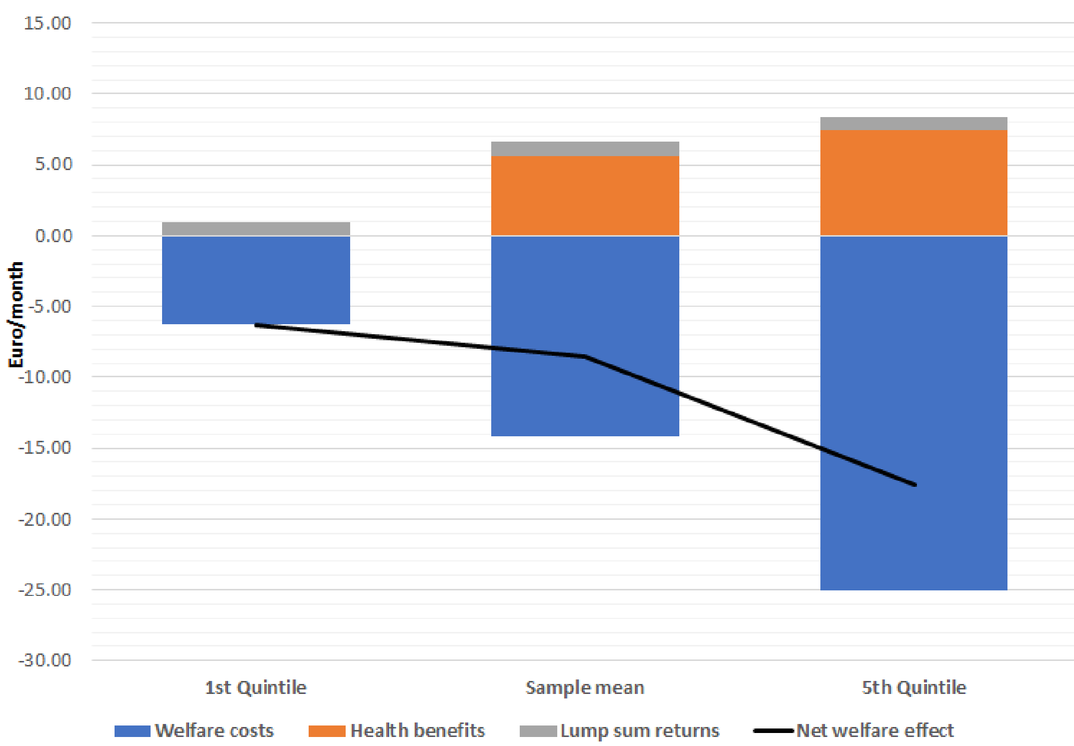


Figure F3: Net welfare effects, ad valorem tax



References

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.

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.

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.

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.

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.

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, M. e. a., G. an Moodie. (2017). Modelled health benefits of a sugar-sweetened beverage tax across different socioeconomic groups in australia: a cost-effectiveness and equity analysis. *PLoS Med*, *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`

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.

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.).

Wilde, P., Huang, Y., Sy, S., Abrahams-Gessel, S., Jardim, T., Paarlberg, R. & et al. (2019). Cost-effectiveness of a us national sugar-sweetened beverage tax with a multistakeholder approach: who pays and who benefits. *Am. J. 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.