Appendix:
For online publication

Optimal sin taxation and market power

Martin O'Connell and Kate Smith
July, 2023

A Optimal tax formulae

A.1 Optimal policy

The social welfare function (equation (1)) is:

$$W = \int_i \omega^i V^i + \lambda(T_D^i + T_\Pi(\delta^i \Pi) - \Phi^i) di,$$

where ω^i are Pareto weights, λ is the marginal value of government revenue, and where individual indirect utilities, V^i, tax revenue from commodity and labor taxes, T_D^i, tax revenue from profits taxation, $T_\Pi(\delta^i \Pi)$ and the budgetary externality, Φ^i, are as defined in Section IA of the paper.

Suppose the government has available a set of linear taxes, $\tau_1, \ldots, \tau_K$, where $K \leq J$ and $\mathcal{J}_k \subseteq \mathcal{M}$ is the set of products subject to rate τ_k. Let the product tax rates be a function of a policy parameter, θ, where changes in θ can capture any arbitrary changes in the tax rates. The optimal excise tax system satisfies $dW/d\theta = 0$,[1] which implies:

$$\int_i \frac{dV^i/d\theta}{\alpha^i} + \frac{dT_D^i}{d\theta} + \tau_\Pi^i \delta^i \frac{d\Pi}{d\theta} - \frac{d\Phi^i}{d\theta} + \left(g^i - 1\right) \frac{dV^i/d\theta}{\alpha^i} di = 0,$$

where α^i is the marginal utility of income of individual i and g^i is the social marginal welfare weight for the individual. The utility, tax revenue, profit and externality derivatives

[1] We focus here on the case where the government does not face any constraints on its policy choice (aside from the excise taxes being linear). The optimal single rate tax entails an improvement in the government's budget. However, in the multi-rate systems that we consider the unconstrained optimal entails a deterioration in the government's budget. When simulating the optimal multi-rate system we therefore add a constraint to the problem – either the government's budget net of externality costs must not deteriorate, or, the stronger condition, that all tax rates must be non-negative.

are:

$$
\begin{aligned}
\frac{dV^i/d\theta}{\alpha^i} &= -\sum_{k=1}^{K}\sum_{j\in\mathcal{J}_k}\left(\frac{d\tilde{p}_j}{d\theta}+\frac{d\tau_k}{d\theta}\right)x_j^i+\delta^i(1-\tau_\Pi^i)\frac{d\Pi}{d\theta}\\
\frac{d\Pi}{d\theta} &= \sum_{j\in\mathcal{M}}\left(\mu_j\frac{dX_j}{d\theta}+\frac{d\tilde{p}_j}{d\theta}X_j\right)+\mu_O\frac{dX_0}{d\theta}\\
\frac{dT_D^i}{d\theta} &= \sum_{k=1}^{K}\sum_{j\in\mathcal{J}_k}\left(\frac{d\tau_k}{d\theta}x_j^i+\tau_k\frac{dx_j^i}{d\theta}\right)+\tau_z^i\frac{dz^i}{d\theta}\\
\frac{\partial T_D^i}{\partial\theta} &= \sum_{k=1}^{K}\sum_{j\in\mathcal{J}_k}\frac{d\tau_k}{d\theta}x_j^i\\
\frac{d\Phi^i}{d\theta} &= \sum_{j\in\mathcal{S}}\phi_j^i\frac{dx_j^i}{d\theta}.
\end{aligned}
$$

Substituting these into the first order condition gives:

$$
\underbrace{\int_i\left(\frac{dT_D^i}{d\theta}-\frac{\partial T_D^i}{\partial\theta}\right)di}_{\text{fiscal externality}}+\underbrace{\sum_{j\in\mathcal{M}\bigcup O}\mu_j\frac{dX_j}{d\theta}}_{\text{market power distortions}}-\underbrace{\int_i\sum_{j\in\mathcal{S}}\phi_j^i\frac{dx_j^i}{d\theta}di}_{\text{externality distortions}}+\underbrace{\int_i(g^i-1)\frac{dV^i/d\theta}{\alpha^i}di}_{\text{distributional concerns}}=0.
$$

A tax reform impacts welfare through four channels; it entails a fiscal externality as behavioral adjustments alter tax revenue raised from commodities and labor, it alters the allocative distortions from both market power and externalities due to changes in commodity demands, and it affects the distribution of utilities across individuals (who may have different social marginal welfare weights). When policy is set optimally these factors must exactly offset in response to a small change in the tax system.

Equivalently, we can write the first order condition in terms of the tax rates, $\tau_1,\ldots,\tau_K$: for all $k'=1,\ldots,K$

$$
\begin{aligned}
\sum_{k=1}^{K}\tau_k\sum_{j\in\mathcal{J}_k}\frac{dX_j}{d\tau_{k'}}+\int_i\sum_{j\in\mathcal{M}}\left(\mu_j-\phi_j^i\right)\frac{dx_j^i}{d\tau_{k'}}di+\mu_O\frac{dX_0}{d\tau_{k'}}+\int_i\tau_z^i\frac{dz^i}{d\tau_{k'}}di+&\\
\int_i(g^i-1)\left(-\sum_{j\in\mathcal{M}}x_j^i\frac{dp_j}{d\tau_{k'}}+\delta^i(1-\tau_\Pi)\frac{d\Pi}{d\tau_{k'}}\right)di=0.&\qquad\text{(A.1)}
\end{aligned}
$$

A.2 Single sin tax rate

When a single tax rate is applied to the set of sin products, $\mathcal{S}$, condition (A.1) reduces to:

$$
\begin{aligned}
\int_i\sum_{j\in\mathcal{S}}\left(\mu_j+\mathbb{1}\{j\in\mathcal{S}\}\tau_{\mathcal{S}}-\phi_j^i\right)\frac{dx_j^i}{d\tau_{\mathcal{S}}}di+\mu_O\frac{dX_0}{d\tau_{\mathcal{S}}}+\int_i\tau_z^i\frac{dz^i}{d\tau_{\mathcal{S}}}di+&\\
\int_i(g^i-1)\left(-\sum_{j\in\mathcal{M}}x_j^i\frac{dp_j}{d\tau_{\mathcal{S}}}+\delta^i(1-\tau_\Pi)\frac{d\Pi}{d\tau_{\mathcal{S}}}\right)di=0.&
\end{aligned}
$$

Letting $d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}} \equiv \sum_{j\in\mathcal{S}} dX_j/d\tau_{\mathcal{S}}$ and $\bar{\phi} \equiv \int_i \frac{1}{n(\mathcal{S})}\sum_{j\in\mathcal{S}} \phi^i_j di$, defining

$$\begin{aligned}
\bar{\mu}_{\mathcal{X}} \equiv& \sum_{j\in\mathcal{X}} \mu_j \frac{dX_j/d\tau_{\mathcal{S}}}{\sum_{j'\in\mathcal{X}} dX_{j'}/d\tau_{\mathcal{S}}} \qquad \text{for } \mathcal{X} = \{\mathcal{S}, \mathcal{N}, O\} \\
\Theta_{\mathcal{X}} \equiv& \frac{d\mathbb{X}_{\mathcal{X}}/d\tau_{\mathcal{S}}}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}} \qquad \text{for } \mathcal{X} = \{\mathcal{S}, \mathcal{N}\} \\
\rho_j \equiv& \frac{dp_j}{d\tau_{\mathcal{S}}}
\end{aligned}$$

and rearranging the first order condition, we obtain equation (2):

$$\begin{aligned}
\tau^*_{\mathcal{S}} =& \underbrace{\bar{\phi} + \frac{\text{cov}(\phi^i_j, dx^i_j/d\tau_{\mathcal{S}})}{(1/n(\mathcal{S})) \times d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}}_{\text{externality correction}} - \underbrace{(\bar{\mu}_{\mathcal{S}} - \bar{\mu}_{\mathcal{N}}\Theta_{\mathcal{N}} - \mu_O\Theta_O)}_{\text{marjet power correction}} \\
&+ \underbrace{\frac{1}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}\left[\text{cov}\left(g^i, \sum_{j\in\mathcal{M}} x^i_j\rho_j - \delta^i(1-\tau^i_{\Pi})\frac{d\Pi}{d\tau_{\mathcal{S}}}\right)\right]}_{\text{distributional concerns}} \\
&- \underbrace{\frac{d(\int_i \mathcal{T} di)/d\tau_{\mathcal{S}}}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}}_{\text{tax base erosion}}.
\end{aligned}$$

A.3 Characterization of the tax base erosion component

The base erosion term in equation (2) can be expressed in terms of income and price elasticities. To see this first note that the base erosion term can be written:

$$\frac{d(\int_i \mathcal{T}(z^i)di)}{d\tau_{\mathcal{S}}} = \int_i \tau^i_z \frac{dz^i}{d\tau_{\mathcal{S}}} di = \int_i \tau^i_z \sum_{j\in\mathcal{M}} \frac{\partial z^i}{\partial p_j}\frac{dp_j}{d\tau_{\mathcal{S}}} di$$

Assume income effects on labor supply are negligible (see Saez et al. (2012) for support of this). Using Slutsky symmetry and the Slutsky decomposition we can re-write $\frac{\partial z^i}{\partial p_j}$:

$$\frac{\partial z^i}{\partial p_j} = -\frac{\partial \tilde{x}^i_j}{\partial(1-\tau^i_z)} = -\frac{\partial x^i_j}{\partial(1-\tau^i_z)} + z^i\frac{\partial x^i_j}{\partial Y^i}, \tag{A.2}$$

where $\tilde{x}^i_j$ denotes compensated demand for good j, and, as in the paper, Y^i is the sum of the consumer's virtual labor income and their profit income.

We make use of the conditional cost function (see Browning (1983)). The consumer's conditional cost function is defined:

$$e^i(\mathbf{p}, u, \bar{z}^i) = \min_{\mathbf{x}^i}\{\mathbf{p}\mathbf{x}^i : \text{s.t. } U^i(\mathbf{x}^i, \bar{z}^i) = u\},$$

and gives the minimum expenditure necessary to achieve a given level of utility, *holding labor supply fixed* at $\bar{z}^i$. The associated conditional compensated demand for product j is given by $\tilde{x}^i_j = \frac{\partial e^i(\mathbf{p},u,\bar{z}^i)}{\partial p_j}$. Inverting the conditional expenditure function yields the conditional indirect utility function: $V^i(\mathbf{p}, Y^i_{\bar{z}}, \bar{z}^i)$, where $Y^i_{\bar{z}} \equiv Y^i + (1-\tau^i_z)\bar{z}^i = \mathbf{p}\mathbf{x}^i$.

Substituting this into the conditional compensated demand for product j, yields the conditional uncompensated demand $x^i_j = \tilde{f}^i_j(\mathbf{p}, \tilde{Y}^i_{\bar{z}}, \bar{z}^i)$. Let z^i denote the optimal labor supply choice and $\tilde{Y}^i(z^i) \equiv Y^i + (1-\tau^i_z)z^i$ denote total income at this level of labor supply. Then $x^i_j = \tilde{f}^i_j(\mathbf{p}, \tilde{Y}^i(z^i), z^i) = f^i_j(\mathbf{p}, 1-\tau^i_z, Y^i)$ where f^i_j is the unconditional compensated demand.

Consider the derivative of $x^i_j = \tilde{f}^i_j(\mathbf{p}, \tilde{Y}^i(z^i), z^i)$ with respect to $(1-\tau^i_z)$:

$$\begin{aligned}
\frac{\partial x^i_j}{\partial(1-\tau^i_z)} =& \frac{\partial \tilde{f}^i_j}{\partial \tilde{Y}^i}\frac{d\tilde{Y}^i}{d(1-\tau^i_z)} + \frac{\partial \tilde{f}^i_j}{\partial z^i}\frac{\partial z^i}{\partial(1-\tau^i_z)} \\
=& \frac{\partial \tilde{f}^i_j}{\partial \tilde{Y}^i}\left(z^i + (1-\tau^i_z)\frac{\partial z^i}{\partial(1-\tau^i_z)}\right) + \frac{\partial \tilde{f}^i_j}{\partial z^i}\frac{\partial z^i}{\partial(1-\tau^i_z)} \\
=& \frac{\partial \tilde{f}^i_j}{\partial \tilde{Y}^i}z^i + \left(\frac{\partial \tilde{f}^i_j}{\partial \tilde{Y}^i}(1-\tau^i_z) + \frac{\partial \tilde{f}^i_j}{\partial z^i}\right)\frac{\partial z^i}{\partial(1-\tau^i_z)} \\
=& \frac{\partial \tilde{f}^i_j}{\partial \tilde{Y}^i}z^i + \frac{\partial f^i_j}{\partial z^i}\frac{\partial z^i}{\partial(1-\tau^i_z)} \qquad \text{(A.3)}
\end{aligned}$$

Where: equality (2) follows from the definition of $\tilde{Y}^i(z^i)$ and our assumption that $\mathcal{T}$ is piecewise linear; equality (3) follows from rearranging; and equality (4) follows from the definition of $\tilde{Y}^i(z^i)$ and $\tilde{f}^i_j(\mathbf{p}, \tilde{Y}^i(z^i), z^i) = f^i_j(\mathbf{p}, 1-\tau^i_z, Y^i)$.

As we have assumed that there are no income effects on labor supply:

$$\frac{\partial \tilde{f}^i_j}{\partial \tilde{Y}^i} = \frac{\partial f^i_j}{\partial Y^i}. \qquad \text{(A.4)}$$

Combining conditions (A.2)-(A.4) yields:

$$\begin{aligned}
\frac{\partial z^i}{\partial p_j} =& -\frac{\partial f^i_j}{\partial z^i}\frac{\partial z^i}{\partial(1-\tau^i_z)} \\
=& -\xi^i_j\zeta^i_z\frac{x^i_j}{(1-\tau^i_z)}
\end{aligned}$$

where $\xi^i_j \equiv \frac{\partial f^i_j}{\partial z^i}\frac{z^i}{x^i_j}$ is the elasticity of good j with respect to labor earnings and $\zeta^i_z \equiv \frac{\partial z^i}{\partial(1-\tau^i_z)}\frac{(1-\tau^i_z)}{z^i}$ is the elasticity of taxable earnings.

Hence the tax base erosion terms can be written:

$$\frac{d(\int_i \mathcal{T}(z^i)di)}{d\tau_{\mathcal{S}}} = \int_i \frac{\tau^i_z}{1-\tau^i_z}\zeta^i_z\sum_{j\in\mathcal{M}}\xi^i_j x^i_j \rho_j di.$$

A.4 Incorporating internalities

Suppose individual i's consumption of sugar-sweetened drinks gives rise to an internality, $I^i = I(\mathbf{x}^i_{\mathcal{S}})$, which impacts their utility but that they ignore when making decisions. In this case we write the consumer's indirect utility function: $V^i = V^i(\nu^i(\mathbf{p}, (1-\tau^i_z), Y^i), I^i)$.

Now there is an additional channel through which policy impacts social welfare: through its impact on internalities. The monetary impact of a marginal policy change on individual i is given by:

$$\frac{dV^i/d\theta}{\alpha^i} = -\sum_{j\in\mathcal{M}} x^i_j \frac{dp_j}{d\theta} + \delta^i(1-\tau^i_\Pi)\frac{d\Pi}{d\theta} + \sum_{j\in\mathcal{S}} \psi^i_j \frac{dx^i_j}{d\theta},$$

where $\psi^i_j \equiv \frac{\partial V^i/\partial I^i}{\alpha^i}\frac{dI^i}{dx^i_j}$ is the monetary cost the individual imposes on themself per additional unit of consumption of sin product j. The first order condition for optimal is then becomes:

$$\underbrace{\int_i \left(\frac{dT^i_D}{d\theta} - \frac{\partial T^i_D}{\partial\theta}\right) di}_{\text{fiscal externality}} + \underbrace{\sum_{j\in\mathcal{M}\bigcup O} \mu_j \frac{dX_j}{d\theta}}_{\text{market power distortions}} \underbrace{- \int_i \sum_{j\in\mathcal{S}} \left(\phi^i_j + \psi^i_j\right)\frac{dx^i_j}{d\theta} di}_{\substack{\text{externality and} \\ \text{internality distortions}}} + \underbrace{\int_i (g^i - 1)\frac{dV^i/d\theta}{\alpha^i} di}_{\text{distributional concerns}} = 0.$$

There are two differences with the condition under no internalities (equation (A.1)). First the externality distortion term is adjusted to also capture the internality distortion. This term reflects how policy reform affects economic efficiency through its impact on externality and internality distortions. The second difference is that the distributional concerns term is now influenced by the covariance of internality changes and social marginal welfare weights; all else equal, for a government with preferences for equity, the more internality reductions are concentrated among the relatively poor, the more valuable is a given fall in the average internality.

B Data

B.1 Sample

Our at-home sample (drawn from the Kantar Worldpanel) comprises 30,405 households and our on-the-go sample (drawn from the Kantar On-The-Go Survey) comprises 2,862 individuals. We omit a small number of consumers that record irregularly. Specifically, in the at-home segment we focus on households that record purchases of any groceries in at least 10 weeks per year and who make at least one drink purchase. In the on-the-go segment we focus on individuals who record at least 5 purchases each year. In each segment, this conditioning drops less than 3% of drinks transactions.

In Table B.1 we compare the demographic composition of the Kantar Worldpanel with the nationally representative Living Costs and Food Survey for a single year (2012). It shows that Kantar Worldpanel households are broadly representative of the UK population.

Table B.1: *Household demographics*

	Kantar	LCFS
Region		
North East	4.6	4.8
	[4.3, 4.9]	[4.3, 5.4]
North West	11.2	11.5
	[10.7, 11.6]	[10.6, 12.3]
Yorkshire and Humber	11.3	9.6
	[10.8, 11.7]	[8.8, 10.4]
East Midlands	8.4	7.8
	[8.0, 8.7]	[7.1, 8.6]
West Midlands	8.9	9.5
	[8.5, 9.3]	[8.7, 10.2]
East of England	10.5	10.4
	[10.1, 10.9]	[9.6, 11.2]
London	8.5	9.0
	[8.1, 8.9]	[8.3, 9.8]
South East	14.6	14.4
	[14.2, 15.1]	[13.5, 15.4]
South West	9.1	9.1
	[8.7, 9.5]	[8.3, 9.9]
Wales	4.6	4.9
	[4.4, 4.9]	[4.3, 5.5]
Scotland	8.2	8.9
	[7.9, 8.6]	[8.1, 9.7]
Socioeconomic status		
Highly skilled	20.9	17.4
	[20.3, 21.4]	[16.1, 18.7]
Semi skilled	55.8	53.0
	[55.1, 56.4]	[51.3, 54.7]
Unskilled	23.4	29.6
	[22.8, 23.9]	[28.1, 31.2]
Number of adults		
1	22.1	32.9
	[21.5, 22.6]	[31.7, 34.2]
2	60.8	55.8
	[60.1, 61.4]	[54.5, 57.2]
3+	17.2	11.3
	[16.7, 17.7]	[10.4, 12.1]
Number of children		
1	14.6	14.1
	[14.1, 15.1]	[13.2, 15.0]
2	15.1	11.0
	[14.6, 15.6]	[10.2, 11.8]
3+	6.1	5.1
	[5.8, 6.5]	[4.6, 5.7]

Notes: Table shows the share of households in the Kantar Worldpanel and Living Costs and Food Survey in 2012 by various demographic groups. Socioeconomic status is based on the occupation of the head of the household and is shown for the set of non-pensioner households. 95% confidence intervals are shown below each share.

B.2 Variation in SSB consumption by income

In Figure 1(b) of the paper we show that higher income households consume less sugar-sweetened beverages than poorer households. To what extent is this driven by heterogeneity in preferences or causal income effects? To answer this we estimate the following two regressions:

$$\text{volSSB}_{iyq} = \sum_{k=1}^{5} \beta_k^{NOFE} \text{income quintile}_{iy}^k + \epsilon_{iyq} \tag{B.1}$$

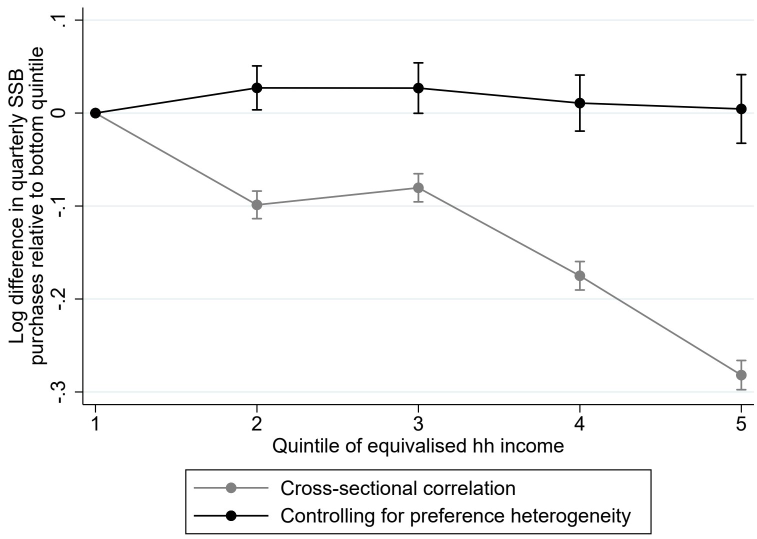
$$\text{volSSB}_{iyq} = \sum_{k=1}^{5} \beta_k^{FE} \text{income quintile}_{iy}^k + \mu_i + \epsilon_{iyq} \tag{B.2}$$

where volSSB_{iyq} denotes the volume of sugar-sweetened beverages purchased by household i in year-quarter (y, q) for at-home consumption. income quintile_{iy}^k is an indicator variable equal to 1 if household i is in income quintile k in year y, and μ_i is a household fixed effect. We estimate this over the period 2008 to 2012.

Figure B.1 plots the estimated $\hat{\beta}_k^{NOFE}$ and $\hat{\beta}_k^{FE}$. It shows that, although in the cross-section there is a negative relationship between household income and volume of sugar-sweetened beverage consumption, this relationship disappears when we control for household fixed effects. This indicates that preference heterogeneity accounts for the variation in sugar-sweetened beverage consumption across the income distribution, with little evidence of causal income effects.

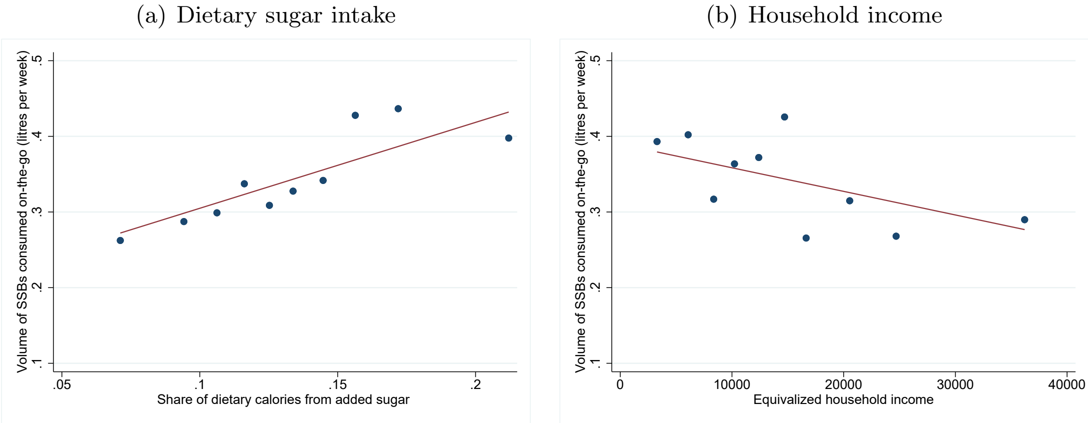
We repeat the analysis summarized Figure 1 in the paper, but using the on-the-go sample. This is shown in Figure B.2.

Figure B.1: *Income correlation with sugar-sweetened beverage consumption*



Notes: The light grey markets plot $\hat{\beta}_k^{NOFE}$ from equation (B.1) and the dark grey markets plot $\hat{\beta}_k^{FE}$ from equation (B.2).

Figure B.2: *Variation in volume of sugar-sweetened beverages consumed on-the-go*



Notes: The left hand panel shows mean volume of sugar-sweetened beverages purchased per person week and consumed on-the-go, by deciles of the share of dietary calories from added sugar (from food consumed at home). The right hand panel shows mean volume of sugar-sweetened beverages purchased per person per week and consumed on-the-go by deciles of equivalized (using the OECD-equivalence scale) household income.

B.3 Brand, products and retailers

In Table B.2 we list the main firms that operate in the drinks market and the brands that they own. The firms Coca Cola Enterprises and Britvic dominate the market, having a combined market share exceeding 65% in the at-home segment and close to 80% in the on-the-go segment. In Table B.3 we list the variants available for each brand. Most brands are available in a regular and diet variant (with some also having an additional zero sugar variant). The table also shows, for each brand-variant, the number of sizes available to consumers in the at-home and on-the-go segments. We include a size option corresponding to multiple units of a single UPC if that UPC-multiple unit combination accounts for at least 10,000 (around 0.2%) transactions. We refer to a brand-variant-size combination as a product.

Table B.2: *Firms and brands*

			Market share (%)		Price (£/l)	
Firm	Brand	Type	At-home	On-the-go	At-home	On-the-go
Coca Cola Enterprises			*33.0*	*59.1*		
	Coke	Soft	20.4	36.4	0.86	2.09
	Capri Sun	Soft	3.1	–	1.08	–
	Innocent fruit juice	Fruit	2.1	1.6	2.04	7.09
	Schweppes Lemonade	Soft	1.7	–	0.44	–
	Fanta	Soft	1.7	5.3	0.79	2.10
	Dr Pepper	Soft	1.2	3.4	0.75	2.08
	Schweppes Tonic	Soft	1.1	–	1.22	–
	Sprite	Soft	1.0	2.8	0.77	2.08
	Cherry Coke	Soft	0.8	4.0	0.96	2.17
	Oasis	Soft	–	5.6	–	2.15
Pepsico/Britvic			*33.6*	*20.0*		
	Robinsons	Soft	10.7	–	1.09	–
	Pepsi	Soft	10.1	11.6	0.64	1.93
	Tropicana fruit juice	Fruit	6.1	3.8	1.62	3.63
	Robinsons Fruit Shoot	Soft	2.7	0.8	1.49	2.83
	Britvic fruit juice	Fruit	1.6	–	2.17	–
	7 Up	Soft	0.9	1.7	0.70	1.88
	Copella fruit juice	Fruit	0.8	–	1.68	–
	Tango	Soft	0.8	2.2	0.66	1.73
GSK			*7.6*	*12.7*		
	Ribena	Soft	3.3	3.4	1.69	2.20
	Lucozade	Soft	3.1	6.4	1.11	2.37
	Lucozade Sport	Soft	1.2	2.9	1.15	2.22
JN Nichols	Vimto	Soft	1.6	–	1.06	–
Barrs	Irn Bru	Soft	0.6	2.6	0.61	1.93
Merrydown	Shloer	Soft	2.0	–	1.79	–
Red Bull	Red Bull	Soft	0.2	3.5	3.66	5.27
Muller	Frijj flavoured milk	Milk	–	1.4	–	1.90
Friesland Campina	Yazoo flavoured milk	Milk	–	0.8	–	1.95
Store brand			*21.3*	*0.0*		
	Store brand soft drinks	Soft	13.1	–	0.62	–
	Store brand fruit juice	Fruit	8.1	–	1.05	–

Notes: Type refers to the type of drinks product: "soft" denotes soft drinks, "fruit" denotes fruit juice, and "milk" denotes flavored milk. The fourth and fifth columns display each firm and brand's share of total spending on all listed drinks brands in the at-home and on-the-go segments of the market; a dash ("–") denotes that the brand is not available in that segment. The final two columns display the mean price (£) per liter for each brand.

Table B.3: *Brands, sugar contents and sizes*

Firm	Brand	Variant	Sugar (g/100ml)	Number of sizes: At-home	Number of sizes: On-the-go
Coca Cola Enterprises	Coke	Diet	0.0	10	2
		Regular	10.6	9	2
		Zero	0.0	7	2
	Capri Sun	Regular	10.9	3	–
	Innocent fruit juice	Regular	10.7	4	1
	Schweppes Lemonade	Diet	0.0	2	–
		Regular	4.2	2	–
	Fanta	Diet	0.0	2	1
		Regular	7.9	2	2
	Dr Pepper	Diet	0.0	2	1
		Regular	10.3	2	2
	Schweppes Tonic	Diet	0.0	2	–
		Regular	5.1	2	–
	Sprite	Diet	0.0	2	–
		Regular	10.6	2	2
	Cherry Coke	Diet	0.0	2	1
		Regular	11.2	2	2
	Oasis	Diet	0.0	–	1
		Regular	4.2	–	1
Pepsico/Britvic	Robinsons	Diet	0.0	6	–
		Regular	3.2	6	–
	Pepsi	Diet	0.0	5	2
		Max	0.0	6	2
		Regular	11.0	5	2
	Tropicana fruit juice	Regular	9.6	4	1
	Robinsons Fruit Shoot	Diet	0.0	2	1
		Regular	10.3	2	–
	Britvic fruit juice	Regular	9.9	2	–
	7 Up	Diet	0.0	2	1
		Regular	10.8	2	2
	Copella fruit juice	Regular	10.1	3	–
	Tango	Regular	3.5	3	2
GSK	Ribena	Diet	0.0	2	1
		Regular	10.8	4	2
	Lucozade	Regular	11.3	3	2
	Lucozade Sport	Diet	0.0	1	1
		Regular	3.6	1	1
JN Nichols	Vimto	Diet	0.0	3	–
		Regular	5.9	4	–
Barrs	Irn Bru	Diet	0.0	1	2
		Regular	8.7	1	2
Merrydown	Shloer	Regular	9.1	3	–
Red Bull	Red Bull	Diet	0.0	–	1
		Regular	10.8	1	1
Muller	Frijj flavoured milk	Regular	10.8	–	1
Friesland Campina	Yazoo flavoured milk	Regular	9.5	–	1
Store brand	Store brand soft drinks	Diet	0.0	4	–
		Regular	10.3	2	–
	Store brand fruit juice	Regular	10.4	2	–

Notes: The final two columns displays the number of sizes of each brand-variant in the at-home and on-the-go segments of the market; a dash ("–") denotes that the brand-variant is not available in that segment.

Table B.4 lists retailers and the share of drinks spending that they account for in each segment. In the at-home segment, four large national supermarket chains account for almost 90% of spending, with the remaining spending mostly made in smaller national retailers. Each of these retailers offers all brands, with some variation in the specific sizes available in each retailer. The large four supermarkets are less prominent in the on-the-go segment, collectively accounting for less than 20% of on-the-go spending on drinks. The majority of transactions in the on-the-go segment are in local convenience stores.

Table B.4: *Retailers*

	Expenditure share (%)	
	at-home	on-the-go
Large national chains	87.0	19.9
of which:		
Tesco	34.7	–
Sainsbury's	16.8	–
Asda	19.8	–
Morrisons	15.7	–
Small national chains	10.7	16.4
Vending machines	0.0	9.2
Convenience stores	2.3	54.6
in region:		
South	–	13.6
Central	–	15.5
North	–	25.5

Notes: Numbers show the share of total drinks expenditure, in the at-home and on-the-go segment, made in each retailer.

C Equilibrium model details

C.1 Choice occasions and consumer groups

We observe households for an average of 36 at-home choice occasions and individuals for an average of 44 on-the-go choice occasions each year, with a total of 3.3 million at-home and 286,576 on-the-go choice occasions.

We estimate our demand model allowing all preference parameters to vary by the consumer groups shown in Table C.1. In the at-home segment we split households based on whether there are any children in the household. In the on-the-go segment we separate individuals aged 30 and under from those aged over 30. We also differentiate between those with low, high or very high total dietary sugar. This measure is based on the household's (or, for individuals in the on-the-go sample, the household to which they belong) share of total calories purchased in the form of added sugar across all grocery shops in the preceding year. We classify those that meet the World Health Organization (2015) recommendation of less than 10% of calories from added sugar as "low dietary sugar", those that purchase between 10% and 15% as "high dietary sugar", and those that purchase more than 15% of their calories from added sugar as "very high dietary sugar".

Table C.1: *Consumer groups*

	No. of consumers	% of sample
At-home segment (households)		
No children, low dietary sugar	7500	17
No children, high dietary sugar	11931	27
No children, very high dietary sugar	7292	17
With children, low dietary sugar	3561	8
With children, high dietary sugar	8382	19
With children, very high dietary sugar	5185	12
On-the-go segment (individuals)		
Under 30, low dietary sugar	240	6
Under 30, high dietary sugar	576	15
Under 30, very high dietary sugar	381	10
Over 30, low dietary sugar	601	16
Over 30, high dietary sugar	1319	34
Over 30, very high dietary sugar	757	20

Notes: Columns 2 and 3 show the number and share of consumers (households in the at-home segment, individuals in the on-the-go segment) in each group, respectively. If consumers move group over the sample period (2008-12) they are counted twice, hence the sum of the numbers of consumers in each group is greater than the total number of consumers. Dietary sugar is calculated based on the share of total calories from added sugar purchased in the preceding year; "low" is less than 10%, "high" is 10-15% and "very high" is more than 15%. Households with children are those with at least one household member aged under 18.

C.2 Dependence across at-home and on-the-go segments

Our demand model assumes independence between demand for drinks in the at-home and on-the-go segments of the market. A potential concern is that when people live in a household that has recently purchased drinks for at-home consumption, they will be less likely to purchase drinks on-the-go, thus introducing dependency between the two segments of the market.

We assess evidence for this by looking at the relationship between a measure of a household's recent at-home drinks purchases and the quantity of drinks an individual from that household purchases on-the-go. We construct a dataset at the individual-day level (we drop days before and after the first and last dates that the individual is observed in the on-the-go sample). The dataset includes the quantity of drinks purchased on-the-go (including zeros), and the total quantity of drinks purchased at home over a variety of preceding time periods.

We estimate:

$$\text{quantity on-the-go}_{it} = \sum_{s=1}^{4} \beta_s \text{week } s \text{ at-home volume}_{it} + \mu_i + \rho_r + \tau_t + \epsilon_{it}$$

$$\text{quantity on-the-go}_{it} = \sum_{d=1}^{7} \beta_d \text{daily } d \text{ at-home volume}_{it} + \mu_i + \rho_r + \tau_t + \epsilon_{it}$$

where week s at-home volume$_{it}$ is the total at-home purchases of drinks made by individual i's household in the s week before day t, and daily d at-home volume$_{it}$ is the total at-home purchases of drinks made by individual i's household on the d day before day t. We estimate both of these regression with and without individual fixed effects to show the importance of individual preference heterogeneity.

Table C.2 shows the estimates. The first two columns show the relationship between the volume of drinks purchased on-the-go and the volume of at-home purchases in the four weeks prior. When we do not include fixed effects, the results are positive and statistically significant. However, in the second column, once we include fixed effects, the results go to almost zero. We see a similar pattern in the final two columns, which show the relationship between volume purchased on-the-go and the daily volume of at-home purchases in the previous 7 days.

These descriptive results provide support for modeling the at-home and on-the-go segments as separate parts of the market. They are also consistent with the formal test of non-separability between the segments conducted in Dubois et al. (2020).

Table C.2: *Dependence across at-home and on-the-go*

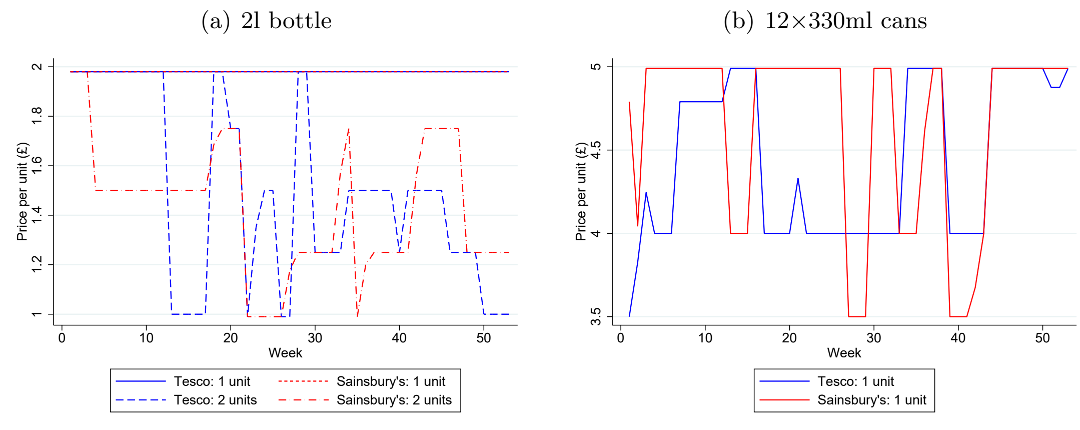
	(1) Volume	(2) Volume	(3) Volume	(4) Volume
At-home purchases 1 week before	0.0008***	0.0001**		
	(0.0000)	(0.0000)		
At-home purchases 2 weeks before	0.0008***	0.0001***		
	(0.0000)	(0.0000)		
At-home purchases 3 weeks before	0.0007***	0.0001*		
	(0.0000)	(0.0000)		
At-home purchases 4 weeks before	0.0007***	0.0001*		
	(0.0000)	(0.0000)		
At-home purchases 1 day before			0.0011***	-0.0002
			(0.0001)	(0.0001)
At-home purchases 2 days before			0.0014***	0.0000
			(0.0001)	(0.0002)
At-home purchases 3 days before			0.0012***	-0.0002
			(0.0001)	(0.0001)
At-home purchases 4 days before			0.0015***	0.0002
			(0.0001)	(0.0001)
At-home purchases 5 days before			0.0016***	0.0002
			(0.0001)	(0.0001)
At-home purchases 6 days before			0.0017***	0.0004**
			(0.0001)	(0.0001)
At-home purchases 7 days before			0.0018***	0.0005***
			(0.0001)	(0.0001)
N	2668585	2668585	2776989	2776989
Mean of dependent variable	0.0452	0.0452	0.0452	0.0452
Time effects?	Yes	Yes	Yes	Yes
Decision maker fixed effects?	No	Yes	No	Yes

Notes: Dependent variable in all regressions is the volume of drinks purchased on-the-go (in liters). An observation is an individual-day; data include zero purchases of drinks. Robust standard errors shown in parentheses.

C.3 Price variation

In Figure C.1 we depict graphically price variation, by showing the path of price over one year for two example products in two different retailers. In the example shown in panel (a) a 2l bottle of Coke costs £2 in either retailer for the whole period. However, for most of the time two units of 2l Coke (which we treat as a separate product) is available on a multi-buy offer – where the price per liter is less when the consumer purchases two units. This kind of multi-buy offer is common, accounting for 30% of transactions. Both the depth and timing of the discount varies over time differentially by retailers. Panel (b) shows an example of a product, 12× 330ml cans of Coke, that does not have a multi-buy offer, but rather where the promotion takes the form of a ticket price reduction, or temporarily low price – this type of promotion accounts for 20% of transactions. Again the timing and depth of promotions vary across the retailers.

Figure C.1: *Examples of price variation for Coke options*



Notes: Panel (a) shows the weekly price series for a 2l bottle of Coke in Tesco and Sainsbury's when either one unit or two units are purchased. Prices are expressed per unit. Panel (b) shows the weekly price series for a pack of 12×330ml cans of Coke in Tesco and Sainsbury's when one unit is purchased.

C.4 Stockpiling

We present evidence regarding whether households in the at-home segment stockpile drinks by conducting a number of checks based on implications of stockpiling behavior highlighted by Hendel and Nevo (2006). Hendel and Nevo (2006) highlight the importance of controlling for preference heterogeneity across consumers; throughout our analysis, we focus on within-consumer predictions and patterns of stockpiling behavior.

We construct a dataset that, for each household, has an observation for every day that they visit a retailer. The data set contains information on: (i) whether the household purchased a drink on that day, (ii) how much they purchased, and (iii) the share of volume of drinks purchased on sale. To account for households who do not record purchasing any groceries for a sustained period of time (for instance, because they are on holiday), we construct "purchase strings" for each household. These are sequences that do not contain a period of non-reporting of any grocery purchases longer than 3 or more weeks.

Inventory. One implication of stockpiling behavior highlighted in Hendel and Nevo (2006) is that the probability a consumer purchases and, conditional on purchasing, the quantity purchased decline in the current inventory of the good. Inventory is unobserved; following Hendel and Nevo (2006) we construct a measure of each household's inventory as the cumulative difference in purchases from the household's mean purchases (within a purchase string). Inventory increases if today's purchases are higher than the household's average, and inventory declines if today's purchases are lower than the household's average.

Let i index household, $\tau = (1, \ldots, \tau_i)$ index days on which we observe the household shopping – we refer to this as a shopping trip – r index retailer and t index year-weeks. We estimate:

$$\text{buysoftdrink}_{i\tau} = \beta^{\text{inv,pp}}\text{inventory}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau}$$
$$q_{i\tau} = \beta^{\text{inv,q}}\text{inventory}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau} \quad \text{if buysoftdrink}_{i\tau} = 1$$

where $\text{buysoftdrink}_{i\tau}$ is a dummy variable equal to 1 if household i buys any drinks on shopping trip τ; $q_{i\tau}$ is the quantity of drink purchased, and $\text{inventory}_{i\tau}$ is household i's inventory on shopping trip τ, constructed as described above. μ_i are household-purchase string fixed effects, ρ_r are retailer effects and t_τ are year-week effects.

If stockpiling behavior is present we would expect that $\beta^{\text{inv,pp}} < 0$ and $\beta^{\text{inv,q}} < 0$; when a household's inventory is high it is less likely to purchase, and conditional on purchasing it will buy relatively little. The first two columns of Table C.3 summarize the estimates from these regressions. There is a small positive relationship between inventory and purchase probability and quantity purchased, conditional on buying. An increase in inventory of 1 liter is associated with an increase in the probability of buying of 0.001, relative to a mean of 0.23, and an increase in the quantity purchased, conditional on buying a positive amount, of 0.013, relative to a mean of 3.925. These effects are both very small and go in the opposite direction to that predicted by Hendel and Nevo (2006) if stockpiling behavior was present.

Time between purchases. The second and third implications of stockpiling behavior highlighted in Hendel and Nevo (2006) are that, on average, the time to the next purchase is longer after a household makes a purchase on sale, and that the time since the previous purchase is shorter.

We check for this by estimating:

$$\text{timeto}_{i\tau} = \beta^{\text{lead}}\text{sale}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau}$$
$$\text{timesince}_{i\tau} = \beta^{\text{lag}}\text{sale}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau}$$

where $\text{timeto}_{i\tau}$ is the number of days to the next drinks purchase, $\text{timesince}_{i\tau}$ is the number of days since the previous purchase, $\text{sale}_{i\tau}$ is the quantity share of drinks purchased on sale on shopping trip τ by household i, and μ_i, ρ_r, and t_τ are household-purchase string, retailer and time effects.

Stockpiling behavior should lead to $\beta^{\text{lead}} > 0$ and $\beta^{\text{lag}} < 0$. Columns (3) and (4) of Table C.3 summarize the estimates from these regressions. We estimate that purchasing on sale is associated with an increase of 0.14 days to the next purchase and 0.23 days less since the previous purchase. The sign of these effects are consistent with stockpiling, however their magnitudes are very small; the average gap between purchases of drinks is 12 days.

Probability of previous purchase being on sale. A fourth implication highlighted by Hendel and Nevo (2006) is that stockpiling behavior implies that if a household makes a non-sale purchase today, the probability of the previous purchase being non-sale is higher than if the current purchase was on sale.

We estimate:

$$\text{nonsale}_{i\tau-1} = \beta^{\text{ns}}\text{sale}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau}$$

where $\text{nonsale}_{i\tau} = \mathbb{1}[\text{sale}_{i\tau} < 0.1]$ indicates a non-sale purchase, and the other effects are as defined above.

The Hendel and Nevo (2006) prediction is that $\beta^{\text{ns}} < 0$. Column (5) shows the estimated β^{ns} from this regression. We find that there is a negative relationship between buying on sale today and the previous purchase not being on sale, however, the magnitude of this effect is relatively small.

Table C.3: *Stockpiling evidence*

	(1) Buys drink	(2) Vol. cond. on buying	(3) Days to next	(4) Days since previous	(5) Prev purch on sale
Inventory	0.0009***	0.0127***			
	(0.0001)	(0.0006)			
Purchase on sale?			0.1445***	-0.2272***	-0.0892***
			(0.0198)	(0.0198)	(0.0016)
Mean of dependent variable	0.2271	3.9250	12.1625	12.1625	0.4639
N	8027010	1823157	1692245	1692245	1712051
Time effects?	Yes	Yes	Yes	Yes	Yes
Retailer effects?	Yes	Yes	Yes	Yes	Yes
Decision maker fixed effects?	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable in column (1) is a dummy variable equal to 1 if the household purchases a non-alcoholic drink on shopping trip τ; in column (2) it is the quantity of drink purchased by household i on shopping trip τ, conditional on buying a positive quantity; in column (3) it is the number of days to the next drink purchase; in column (4) it is the number of days since the previous purchase; and in column (5) it is a dummy variable equal to 1 if the previous purchase was not on sale. Robust standard errors are shown in parentheses.

Sales and product switching. While the evidence suggests that people do not change the timing of their purchases when they buy on sale, this does not imply consumer choice does not respond to price variation resulting from sales. We quantify the propensity of people to switch brands, sizes and pack types (e.g. from bottles to cans) by estimating the following:

$$
\begin{aligned}
\text{brandswitch}_{i\tau} &= \beta^{brandswitch}\text{sale}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau} \\
\text{sizeswitch}_{i\tau} &= \beta^{sizeswitch}\text{sale}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau} \\
\text{packtypeswitch}_{i\tau} &= \beta^{packtypeswitch}\text{sale}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau} \\
\text{retailerswitch}_{i\tau} &= \beta^{retailerswitch}\text{sale}_{i\tau} + \mu_i + \rho_r + t_\tau + \epsilon_{i\tau}
\end{aligned}
$$

where $\text{brandswitch}_{i\tau}$ is a dummy variable equal to 1 if the household purchased a brand that is different from the brand they bought last, $\text{sizeswitch}_{i\tau}$ is a dummy variable equal to 1 if the household purchased a size that is different from the size they bought last, $\text{packtypeswitch}_{i\tau}$ is a dummy variable equal to 1 if the household purchased a pack type that is different from the pack type they bought last, and $\text{retailerswitch}_{i\tau}$ a dummy variable equal to 1 if the household shopped in a different retailer to their last shopping trip.

Table C.4 shows the estimated β coefficients. We find that buying on sale leads to an increase in the probability of switching brands, sizes and pack types. The percentage effect is largest for pack type switching: buying on sale is associated with an 12.5% (0.016/0.127) increase in the probability that the household switches to buying a new pack type (i.e., cans instead of bottles or vice versa). Buying on sale is associated with a 3.3% and 4.5% increase in probability of switching between brands and sizes, respectively. In contrast, although statistically significant, there is less than a 1% change in the probability of switching retailer. Switching across pack types, brand and sizes in response to sales contributes to the identification of the price preference parameters in our demand model.

Table C.4: *Sales and product switching*

	(1) Brand switch	(2) Size switch	(3) Pack type switch	(4) Retailer switch
Purchase on sale?	0.0181***	0.0234***	0.0160***	0.0035***
	(0.0012)	(0.0012)	(0.0007)	(0.0010)
Mean of dependent variable	0.5432	0.5183	0.1272	0.3566
N	1823157	1823157	1823157	1823157
Time effects?	Yes	Yes	Yes	Yes
Retailer effects?	Yes	Yes	Yes	Yes
Decision maker fixed effects?	Yes	Yes	Yes	Yes

Notes: The dependent variable in column (1) is a dummy variable equal to 1 if the household buys a brand on shopping trip τ that they did not buy on the last trip on which they made a soft drinks purchase; in column (2) it is a dummy variable equal to 1 if the household buys a size on shopping trip τ that they did not buy on the last trip on which they made a soft drinks purchase; in column (3) it is a dummy variable equal to 1 if the household buys a pack type on shopping trip τ that they did not buy on the last trip on which they made a soft drinks purchase; in column (4) it is a dummy variable equal to 1 if the household visits a different retailer on shopping trip τ to their previous trip. Robust standard errors are shown in parentheses.

To summarize, we find very limited evidence of stockpiling behavior in our data; although we cannot conclusively rule it out, any effects are likely to be extremely small.

D Additional demand and supply estimates

We allow all parameters to vary by consumer group and estimate the choice model separately by groups. For estimation: in the at-home segment, for each group, we use a random sample of 1,500 households and 10 choice occasions per household; in the on-the-go sample we use data on all individuals in each group and randomly sample 50 choice occasions per individual, weighting the likelihood function to account for differences in the frequency of choice occasion across consumers.

To calculate the confidence intervals, we obtain the variance-covariance matrix for the parameter vector estimates using standard asymptotic results. We then take 50 draws of the parameter vector from the joint normal asymptotic distribution of the parameters and, for each draw, compute the statistic of interest, using the resulting distribution across draws to compute Monte Carlo confidence intervals (which need not be symmetric).

Demand parameter estimates

Table D.1 summarizes our demand estimates. The top half of the table shows estimates for the at-home segment of the market and the bottom half shows estimates for the on-the-go segment. These include a set of random coefficients over price, a dummy variable for drinks products, a dummy for variable for whether the product contains sugar, a dummy variable for whether the product is 'large' (more than 2l in size for the at-home segment, and 500ml in size in the on-the-go segment), and dummy variables for whether the product is a cola, lemonade, fruit juice, store brand soft drink (at-home only), or a flavored milk (on-the-go only).

Conditional on consumer group, the price random coefficient is log-normally distributed and the other random coefficients are normally distributed; the unconditional distribution of consumer preferences is a mixture of normals. We normalize the means of the random coefficients for the drinks, large, cola, lemonade, store soft drinks and fruit juice effects to zero as they are collinear with the brand-size effects. We allow for correlation within consumer group between preferences for sugar and drinks. For the coefficients on price, branded soft drinks, store brand soft drinks, fruit juice and sugar we allow the mean preferences (within consumer group) to vary by household equivalized income. Note that across 10 of the 12 consumer groups (across both segments) the interaction with the price coefficient is negative and statistically significant – this indicates higher income households are less price sensitive than lower income households. Higher income households also tend to have weaker preferences for store brand soft drinks and products with high sugar content, but stronger preferences for pure fruit juice.

D.1 Elasticities, costs and margins

Table D.2 reports mean market elasticities for a set of popular products in the at-home and on-the-go segments of the market. For each segment, we show elasticities for the most popular size belonging to each of the 10 most popular brand-variants (where variants refer to regular/diet/zero versions). Table D.3 reports the average price, marginal cost and price-cost margin (all per liter) for each brand, as well as the average price-cost mark-up. Numbers in brackets are 95% confidence intervals.

In Figure D.1 we show how prices, marginal costs, and price-cost margins vary with product size. There is strong non-linear pricing; in per liter terms, smaller products are, on average, more expensive. Average marginal costs are broadly constant across the size distribution, with the exception of small single portion sizes, which, on average, have higher costs. Price-cost margins are declining in size – the average margin (per liter) is more than twice as large for the smallest options compared with the largest. This pattern has important implications for tax policy. A tax levied on the sugar in sweetened beverages will result in a higher tax burden (per liter) on large products. To the extent that this causes consumers to switch more strongly away from large products, relative to smaller products, consumers' baskets of taxed products will become more dominated by small, high margin products, which will exacerbate distortions associated with the market power of sugar-sweetened beverages.

Table D.1: *Estimated preference parameters*

At-home		No children			Children		
		low dietary sugar	med. dietary sugar	high dietary sugar	low dietary sugar	med. dietary sugar	high dietary sugar
Mean	Price	0.257	0.356	0.316	0.378	0.411	0.399
		(0.052)	(0.045)	(0.045)	(0.039)	(0.034)	(0.031)
	Sugary:<10g/100ml	1.076	1.048	1.119	0.507	0.851	1.001
		(0.135)	(0.123)	(0.125)	(0.112)	(0.106)	(0.099)
	Sugary:≥10g/100ml	0.541	0.441	0.645	0.102	0.507	0.843
		(0.110)	(0.099)	(0.102)	(0.089)	(0.087)	(0.080)
	Advertising	0.252	0.289	0.230	0.268	0.246	0.311
		(0.055)	(0.053)	(0.051)	(0.045)	(0.040)	(0.039)
Interaction with income	× Price	-0.008	-0.009	-0.011	-0.010	-0.012	-0.010
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	× Branded soft drinks	0.005	-0.005	0.009	-0.015	-0.015	-0.025
		(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
	× Store brand soft drinks	-0.013	-0.006	0.018	-0.029	-0.023	-0.036
		(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)
	× Pure fruit juice	0.051	0.021	0.022	0.033	0.037	0.018
		(0.008)	(0.007)	(0.008)	(0.009)	(0.009)	(0.009)
	× Sugary:<10g/100ml	-0.027	-0.009	-0.014	-0.009	-0.018	-0.006
		(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
	× Sugary:≥10g/100ml	-0.029	-0.008	-0.006	-0.022	-0.034	-0.016
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Variance	Price	0.127	0.175	0.165	0.069	0.061	0.075
		(0.019)	(0.020)	(0.019)	(0.010)	(0.009)	(0.009)
	Sugary	2.205	2.308	1.851	1.309	1.644	1.766
		(0.210)	(0.188)	(0.176)	(0.115)	(0.132)	(0.133)
	Drinks	2.217	1.790	1.422	1.296	1.750	1.481
		(0.220)	(0.165)	(0.177)	(0.134)	(0.149)	(0.142)
	Large	0.388	0.458	0.454	0.770	0.360	0.407
		(0.237)	(0.142)	(0.153)	(0.183)	(0.117)	(0.106)
	Cola	2.376	2.026	2.454	1.929	1.960	2.131
		(0.303)	(0.233)	(0.317)	(0.208)	(0.184)	(0.199)
	Lemonade	2.071	2.951	1.574	1.838	1.346	2.280
		(0.466)	(0.495)	(0.288)	(0.457)	(0.267)	(0.304)
	Store brand soft drinks	2.562	2.638	2.191	2.357	2.040	1.805
		(0.241)	(0.250)	(0.224)	(0.202)	(0.167)	(0.147)
	Pure fruit juice	3.176	3.283	3.993	2.823	2.449	2.613
		(0.286)	(0.327)	(0.459)	(0.271)	(0.240)	(0.218)
Covariance	Sugary-Drinks	-1.751	-1.607	-0.704	-0.851	-1.131	-1.041
		(0.192)	(0.157)	(0.146)	(0.100)	(0.117)	(0.121)
On-the-go		Aged under 30			Aged over 30		
		low dietary sugar	med. dietary sugar	high dietary sugar	low dietary sugar	med. dietary sugar	high dietary sugar
Mean	Price	1.069	1.207	0.966	0.868	1.499	1.263
		(0.129)	(0.088)	(0.146)	(0.123)	(0.054)	(0.083)
	Sugary:<10g/100ml	2.641	3.134	2.701	2.806	2.271	0.994
		(0.299)	(0.167)	(0.224)	(0.159)	(0.118)	(0.144)
	Sugary:≥10g/100ml	0.629	1.230	1.215	1.566	0.821	0.064
		(0.205)	(0.104)	(0.130)	(0.119)	(0.095)	(0.090)
	Advertising	0.786	0.666	0.545	0.553	0.457	0.603
		(0.077)	(0.045)	(0.060)	(0.046)	(0.031)	(0.046)
Interaction with income	× Price	0.022	-0.013	0.013	-0.038	-0.076	-0.081
		(0.014)	(0.010)	(0.012)	(0.009)	(0.006)	(0.007)
	× Branded soft drinks	-0.016	0.042	0.047	0.025	-0.021	-0.108
		(0.016)	(0.010)	(0.014)	(0.010)	(0.007)	(0.008)
	× Pure fruit juice	0.148	0.154	0.065	0.028	0.009	-0.145
		(0.027)	(0.018)	(0.023)	(0.021)	(0.011)	(0.013)
	× Flavored milk	-0.090	0.089	-0.019	-0.003	-0.070	-0.106
		(0.023)	(0.015)	(0.020)	(0.018)	(0.011)	(0.014)
	× Sugary:<10g/100ml	-0.001	-0.050	-0.015	-0.081	-0.054	0.103
		(0.010)	(0.006)	(0.008)	(0.007)	(0.005)	(0.006)
	× Sugary:≥10g/100ml	0.037	-0.037	-0.011	-0.080	-0.036	0.082
		(0.009)	(0.005)	(0.007)	(0.006)	(0.004)	(0.005)
Variance	Price	0.531	0.083	0.030	0.273	0.120	0.115
		(0.117)	(0.013)	(0.009)	(0.049)	(0.011)	(0.018)
	Sugary	8.770	4.380	7.230	8.576	7.970	6.333
		(0.724)	(0.235)	(0.534)	(0.423)	(0.320)	(0.354)
	Drinks	3.411	5.532	3.495	5.551	2.968	3.300
		(0.396)	(0.282)	(0.320)	(0.301)	(0.168)	(0.203)
	Large	4.839	4.985	3.630	7.787	4.157	5.667
		(0.397)	(0.299)	(0.248)	(0.412)	(0.171)	(0.304)
	Cola	4.927	5.358	3.660	7.536	7.191	7.472
		(0.398)	(0.294)	(0.300)	(0.418)	(0.284)	(0.346)
	Lemonade	3.383	4.984	6.205	0.793	1.285	5.507
		(0.408)	(0.680)	(0.666)	(0.183)	(0.160)	(0.492)
	Pure fruit juice	17.307	3.295	4.448	8.997	3.006	2.728
		(2.501)	(0.513)	(0.613)	(0.776)	(0.304)	(0.381)
	Flavored milk	5.667	2.251	9.466	4.636	4.140	2.727
		(1.015)	(0.485)	(1.097)	(0.919)	(0.556)	(0.503)
Covariance	Sugary-Drinks	-4.422	-4.503	-3.877	-5.903	-3.879	-3.494
		(0.514)	(0.227)	(0.416)	(0.319)	(0.222)	(0.239)
Brand-size effects		Yes	Yes	Yes	Yes	Yes	Yes
Brand-retailer effects		Yes	Yes	Yes	Yes	Yes	Yes
Size-retailer effects		Yes	Yes	Yes	Yes	Yes	Yes
Brand-time effects		Yes	Yes	Yes	Yes	Yes	Yes
Size-time effects		Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors are reported below the coefficients. For the variance estimates: "Sugary" is an indicator variable equal to 1 if the drink contains any sugar; "Drinks" is an indicator variable equal to 1 for all the inside drinks options; "Large" is an indicator variable equal to 1 if the option is larger than 2l in the at-home segment and 500ml in the on-the-go-segment. We group together some minor brands in the brand-retailer and brand-time effects.

Table D.2: *Price elasticities for popular products*

At-home		Coca Cola Enterprises						Pepsico/Britvic			GSK
		Coke		Capri Sun	Schweppes	Robinsons		Pepsi		Tropicana	Lucozade
		Reg. 2l	Diet 2x2l	10x200ml	Reg. 2x2l	Squash 1l	Fruit diet 1l	Reg. 2l	Max 2x2l	1l	Reg. 6x380ml
Coke	Regular 2l	-2.204	0.020	0.016	0.007	0.019	0.004	0.028	0.027	0.024	0.021
	Diet 2x2l	0.009	-2.721	0.009	0.005	0.008	0.008	0.010	0.058	0.016	0.016
Capri Sun	10x200ml	0.007	0.010	-2.582	0.009	0.025	0.007	0.011	0.013	0.025	0.033
Schweppes Lemonade	Regular 2x2l	0.006	0.010	0.017	-2.261	0.019	0.005	0.007	0.013	0.028	0.037
Robinsons	Squash 1l	0.007	0.008	0.023	0.009	-1.297	0.007	0.011	0.012	0.027	0.028
	Fruit diet 1l	0.004	0.015	0.013	0.005	0.015	-1.319	0.006	0.021	0.019	0.016
Pepsi	Regular 2l	0.023	0.020	0.019	0.007	0.023	0.005	-1.391	0.032	0.022	0.021
	Max 2x2l	0.009	0.047	0.010	0.005	0.010	0.009	0.013	-2.379	0.016	0.016
Topicana	1l	0.004	0.006	0.010	0.006	0.012	0.004	0.005	0.008	-1.944	0.016
Lucozade	Regular 6x380ml	0.006	0.010	0.021	0.012	0.020	0.005	0.007	0.013	0.026	-2.555
Outside option		0.004	0.007	0.009	0.005	0.014	0.006	0.006	0.010	0.019	0.012

On-the-go		Coca Cola Enterprises						Pepsico/Britvic		GSK	
		Coke		Fanta	Dr Pepper	Cherry Coke	Oasis	Pepsi		Ribena	Lucozade
		Reg 500ml	Diet 500ml	Reg 500ml	Reg 500m	Reg 500m	Reg 500ml	Reg 500ml	Max 500ml	Reg 500ml	Reg 330ml
Coke	Regular 500ml	-1.785	0.135	0.051	0.041	0.035	0.079	0.211	0.067	0.022	0.019
	Diet 500ml	0.229	-2.186	0.023	0.019	0.016	0.038	0.067	0.218	0.010	0.008
Fanta	Regular 500ml	0.202	0.054	-2.419	0.113	0.095	0.208	0.066	0.031	0.062	0.042
Dr Pepper	Max 500ml	0.215	0.059	0.150	-2.580	0.104	0.245	0.069	0.034	0.063	0.041
Cherry Coke	Regular 500ml	0.204	0.054	0.138	0.115	-2.401	0.197	0.061	0.028	0.056	0.046
Oasis	Regular 500ml	0.196	0.055	0.129	0.116	0.084	-2.165	0.059	0.029	0.049	0.038
Pepsi	Regular 500ml	0.761	0.143	0.060	0.048	0.038	0.086	-2.256	0.090	0.027	0.024
	Regular 500ml	0.248	0.479	0.029	0.024	0.018	0.043	0.092	-2.355	0.013	0.010
Ribena	Regular 500ml	0.220	0.060	0.159	0.121	0.097	0.202	0.075	0.035	-2.606	0.046
Lucozade	Regular 330ml	0.099	0.026	0.056	0.042	0.043	0.081	0.035	0.014	0.024	-1.813
Outside option		0.065	0.045	0.029	0.023	0.019	0.042	0.023	0.025	0.011	0.042

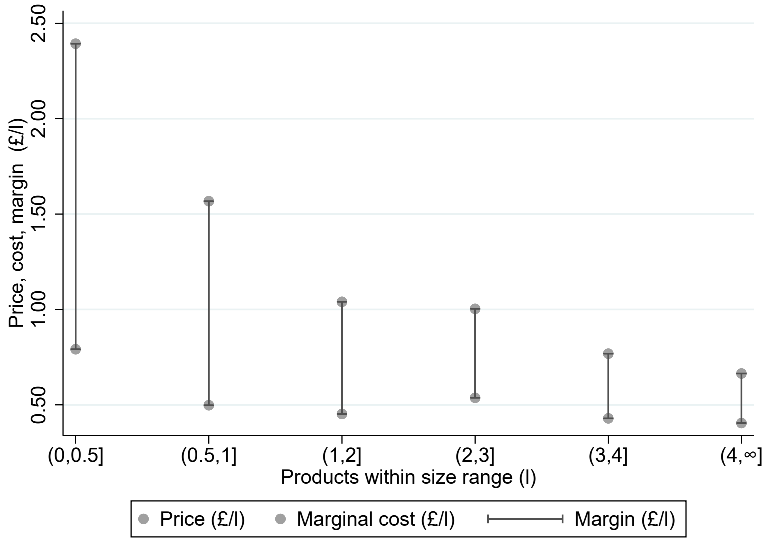
Notes: Numbers show the mean price elasticities of market demand in the most recent year covered by our data (2012). Number shows price elasticity of demand for option in column 1 with respect to the price of option in row 1.

Table D.3: *Average price-cost margins by brands*

Firm	Brand	Price (£/l)	Marginal cost (£/l)	Price-cost margin (£/l)	(Price-cost) /Price
Coca Cola Enterprises	Coke	1.13	0.34	0.79	0.60
			[0.31, 0.38]	[0.76, 0.83]	[0.58, 0.62]
	Capri Sun	1.17	0.57	0.60	0.50
			[0.55, 0.59]	[0.58, 0.62]	[0.49, 0.52]
	Innocent fruit juice	3.34	1.54	1.80	0.56
			[1.40, 1.65]	[1.69, 1.94]	[0.54, 0.59]
	Schweppes Lemonade	0.52	0.14	0.38	0.71
			[0.13, 0.16]	[0.36, 0.39]	[0.69, 0.74]
	Fanta	1.44	0.41	1.03	0.70
			[0.34, 0.47]	[0.96, 1.10]	[0.67, 0.74]
	Dr Pepper	1.33	0.42	0.91	0.66
			[0.38, 0.47]	[0.87, 0.95]	[0.64, 0.69]
	Schweppes Tonic	1.65	0.72	0.93	0.64
			[0.66, 0.75]	[0.89, 0.99]	[0.62, 0.67]
	Sprite	1.26	0.34	0.92	0.72
			[0.28, 0.38]	[0.88, 0.98]	[0.70, 0.76]
	Cherry Coke	1.53	0.50	1.04	0.62
			[0.43, 0.56]	[0.98, 1.11]	[0.60, 0.66]
	Oasis	2.31	0.53	1.79	0.77
			[0.33, 0.68]	[1.64, 1.98]	[0.71, 0.86]
Pepsico/Britvic	Robinsons	1.20	0.32	0.88	0.74
			[0.29, 0.34]	[0.86, 0.91]	[0.72, 0.76]
	Pepsi	1.02	0.40	0.62	0.61
			[0.37, 0.43]	[0.59, 0.65]	[0.59, 0.63]
	Tropicana fruit juice	2.20	0.94	1.25	0.56
			[0.87, 1.01]	[1.19, 1.33]	[0.54, 0.59]
	Robinsons Fruit Shoot	1.81	0.63	1.18	0.64
			[0.58, 0.68]	[1.13, 1.23]	[0.62, 0.66]
	Britvic fruit juice	2.06	0.90	1.16	0.56
			[0.86, 0.93]	[1.12, 1.20]	[0.55, 0.58]
	7 Up	1.22	0.45	0.77	0.68
			[0.41, 0.48]	[0.74, 0.82]	[0.65, 0.71]
	Copella fruit juice	1.40	0.23	1.17	0.83
			[0.19, 0.27]	[1.13, 1.21]	[0.80, 0.86]
	Tango	1.13	0.34	0.79	0.74
			[0.31, 0.38]	[0.75, 0.82]	[0.71, 0.76]
GSK	Ribena	1.77	0.89	0.88	0.50
			[0.85, 0.92]	[0.86, 0.92]	[0.48, 0.52]
	Lucozade	1.62	0.77	0.85	0.53
			[0.73, 0.82]	[0.80, 0.89]	[0.51, 0.55]
	Lucozade Sport	1.49	0.83	0.65	0.44
			[0.80, 0.86]	[0.63, 0.69]	[0.42, 0.46]
	Vimto	1.09	0.50	0.59	0.57
			[0.48, 0.51]	[0.58, 0.61]	[0.56, 0.59]
	Irn Bru	1.56	0.63	0.93	0.61
			[0.55, 0.72]	[0.84, 1.00]	[0.57, 0.65]
Merrydown	Shloer	1.59	0.71	0.88	0.55
			[0.68, 0.73]	[0.86, 0.91]	[0.54, 0.57]
Red Bull	Red Bull	4.74	2.60	2.15	0.44
			[2.32, 2.81]	[1.93, 2.42]	[0.40, 0.49]
Total		1.44	0.55	0.89	0.62
			[0.51, 0.58]	[0.85, 0.93]	[0.60, 0.64]

Notes: We recover marginal costs for each product in each market. We report averages by brand for the most recent year covered by our data (2012). Margins are defined as price minus cost and expressed in £ per liter. 95% confidence intervals are given in square brackets.

Figure D.1: *Price-cost margins, by product size*



Notes: We group products by size. The figure shows the mean price, cost, and margin (all expressed in £/l) across products within each size range. Numbers are for the more recent year covered by our data (2012).

D.2 Model validation

We use data on the price changes of drinks following the introduction of the UK's Soft Drinks Industry Levy (SDIL) in 2018 to validate our empirical model's tax pass-through predictions. We use a weekly database of UPC level prices and sugar contents for drinks products, collected from the websites of 6 major UK supermarkets (Tesco, Asda, Sainsbury's, Morrisons, Waitrose and Ocado), that cover the period 12 weeks before and 18 weeks after the introduction of the tax (on April 1, 2018).[2] We use data on all the brands included in our demand model, excluding data on minor brands (some of which benefit from a small producers' exemption from the levy).

The SDIL tax is levied per liter of product, with a lower rate of 18p/liter for products with sugar contents of 5-8g/100ml and a higher rate of 24p/liter for products with sugar content > 8g/100m. The tax applies to sugar-sweetened beverages; milk-based drinks and pure fruit juices are exempt from the tax.

We define three sets of products. First, the "higher rate treatment group" are those products with at least 8g of sugar per 100ml, at the time the tax was introduced and therefore are subject to the higher tax rate. Second, the "lower rate treatment group" are those products that have 5-8g of sugar per 100ml, and therefore are subject to the lower tax rate. The remaining set of products are exempt, either because their sugar content is less than 5g per 100ml, or because they are milk-based or fruit juice. There

[2]We are grateful to the University of Oxford for providing us with access to these data, which were collected as part of the foodDB project (University of Oxford (2018)).

was some reformulation in anticipation of the introduction of the SDIL. We categorize products based on the post reformulation sugar contents.[3]

We estimate price changes for the two treatment and the exempt groups. Let j index product, r retailer, and t week. We define the dummy variables $\text{TreatHi}_j = 1$ if product j is in the high treatment group, $\text{TreatLo}_j = 1$ if product j is in the low treatment group, and $\text{TreatExempt}_j = 1$ if product j is exempt from the tax. Let Post_t denote a dummy variable equal to 1 if $t >= 13$ i.e. weeks following the introduction of the tax. We estimate the following regression, pooling across products in each of the three groups:

$$p_{jrt} = \beta^{hi}\text{TreatHi}_j \times \text{Post}_t + \beta^{lo}\text{TreatLo}_j \times \text{Post}_t + \sum_{t \neq 12} \tau_t + \xi_j + \rho_r + \epsilon_{jrt} \tag{D.1}$$

where p_{jrt} denotes the price per liter of product j in retailer r in week t,[4] τ_t are week effects, ξ_j are product fixed effects, and ρ_r are retailer fixed effects.

Figure D.2(a) plots the estimated price changes, relative to the week preceding the introduction of the tax, for the higher rate treatment group ($= \hat{\beta}^{hi} \times \text{Post}_t + \sum_{t \neq 12} \hat{\tau}_t$). Figure D.2(b) plots the analogous estimates for the lower rate treatment group ($= \hat{\beta}^{lo} \times \text{Post}_t + \sum_{t \neq 12} \hat{\tau}_t$). Figure D.2(c) plots the estimates for the group of products exempt from the tax ($\sum_{t \neq 12} \hat{\tau}_t$). The solid blue line plots the tax per liter. The data suggest that there was slight overshifting of the tax, with an average price increase among the high treatment group of 26p per liter (a pass-through rate of 108%), and the average price increase among the low treatment group of 19p per liter (a pass-through rate of 105%). The prices of products not subject to the tax do not change following its introduction.

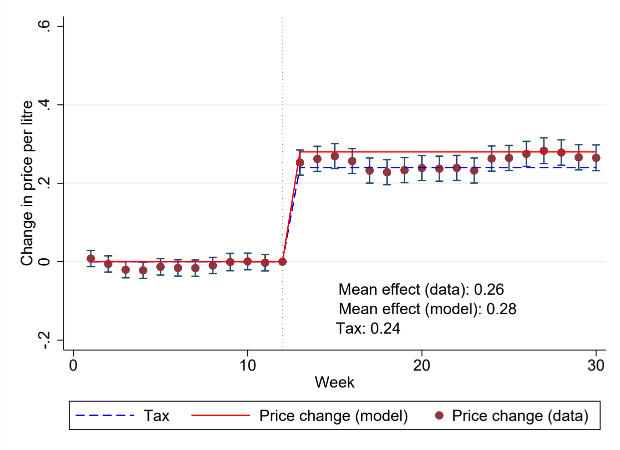
We simulate the introduction of the SDIL using our estimated model of demand and supply in the non-alcoholic drinks market (based on product sugar contents when the SDIL was implemented). The red lines plot the average price increase for each of the three group predicted by our model. These match very closely the actual price increases following the policy's introduction.

[3] We exclude a small number of products belonging to the Irn Bru and Shloer brands that were reformulated approximately 10 weeks after the introduction of the tax.

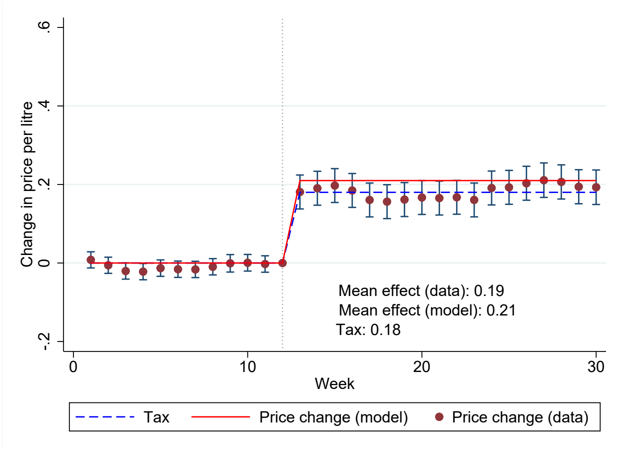
[4] This is the VAT-exclusive price per liter.

Figure D.2: *Out-of-sample model validation: UK Soft Drinks Industry Levy*

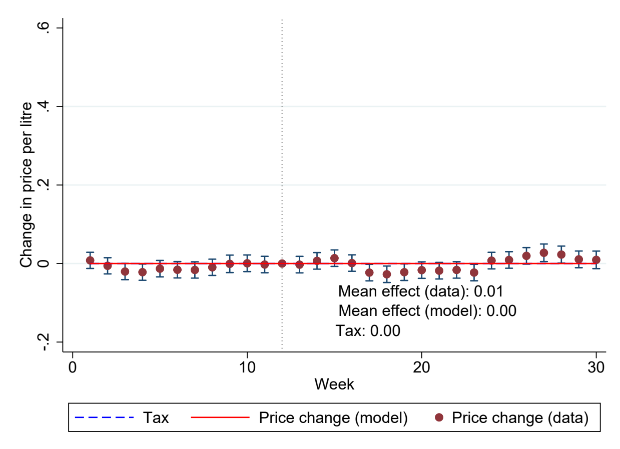
(a) High treatment group



(b) Low treatment group



(c) Exempt group



Notes: Grey markers show the estimated price changes (relative to the week preceding the introduction of the tax). For the higher rate treatment group (top panel), the estimated prices changes are $= \hat{\beta}^{hi} Post_t + \sum_{t \neq 12} \hat{\tau}_t$*, for the lower rate treatment group (middle panel), the estimated price changes are* $= \hat{\beta}^{lo} Post_t + \sum_{t \neq 12} \hat{\tau}_t$*, and for the exempt group (bottom panel) they are* $= \hat{\tau}_t$ *All coefficients are estimated jointly (equation (D.1)). 95% confidence intervals shown. The blue line shows the value of the tax, and the red line shows the predicted price changes from our estimated demand and supply model.*

E Implementation of optimal tax problem

E.1 Externality calibration

We use two different approaches to estimate the average externality costs associated with reductions in sugar-sweetened beverage consumption. Both yield similar estimates.

Approach using evidence from Wang et al. (2012)

Wang et al. (2012) consider the impact of a fall of approximately 15% in sugar-sweetened beverage consumption among adults aged 25-64 on health care costs in the US. They conclude it would result in savings of $17.1 billion realized over 10 years, discounted at a rate of 3% per year.

As a baseline, they use an average daily serving of 0.56 and serving size of 170kcal. They simulate a reduction in sugar-sweetened beverage consumption to 0.47 daily servings, which translates into a fall in calories from these products of (0.56-0.47)*170=15kcal per adult per day.[5] This corresponds to a 3.75g fall in sugar per adult per day. Their estimate of health care cost savings of $17.1 billion over 10 years corresponds to an average daily fall of $4.7 million, or 2.7¢ per adult (based on 171 million Americans aged 15-64). Hence, the implied health cost saving is $2.7/0.375 \approx 7$¢ per 10g of sugar.

We convert the average health care saving to UK numbers by applying a \$-£ exchange rate of 0.75 and deflating by an estimate of the cost of providing health care in the UK relative to US (equal to 0.83 and based on OECD (2019)). This yields an average health care cost saving of approximately 4 pence per 10g of sugar. Health care in the UK is almost entirely provided by the taxpayer funded National Health Service, so we treat this as an externality.

Approach using evidence from Briggs et al. (2013)

Briggs et al. (2013) use a comparative risk assessment model, which maps dietary changes into health outcomes, to estimate the impact of a 1g reduction in sugar per adult per day into changes in the prevelance of obesity and overweight.[6] They find that the fall in consumption would lead to a 1.3% fall in the number of obese adults and a 0.9% fall in the number of overweight adults. This implies a 1.1% reduction in the number of adults who are overweight or obese (27% of UK adults are obese and 36% are overweight).

Public Health England (2017) estimate that the NHS spent £6.1 billion on overweight and obesity-related ill health in 2014/15. A 1.1% reduction in these costs implies a fall of 0.37 pence per adult per day (based on 49.3 million adults in the UK in 2014). This yields a healthcare cost saving of approximately 3.7 pence per 10g sugar.

[5] Note, they assume 40% of the calories are displaced so refer to an 9kcal reduction.

[6] The 1g reduction corresponds to their estimate of the impact of 20% tax on sugar-sweetened beverages.

Distribution of the external costs

Based on the World Health Organization's official recommendation that individual added sugar consumption should be below 10% of dietary calories we assume that only consumers with dietary sugar above this threshold create externalities. This group comprises around 80% of consumers, so this implies an externality per 10g of sugar of 5 pence per 10g of sugar for this group. Since, on average, sugar-sweetened beverages have 26g of sugar per 10 oz, this implies an *average* externality of 14 pence per 10 oz of sugar-sweetened beverage for people in this group; however individual marginal externalities will vary with the sugar content of sweetened beverages chosen.

E.2 Distribution of profits

In Section VB of the paper we investigate how distributional concerns change the optimal tax rate on sugar-sweetened beverages. This requires information on how profits are distributed between the government, domestic residents and the portion that flow overseas. Measuring the distribution of profits across individuals is challenging and a topic of much recent work. It forms a key part of the "Distributional National Accounts" (DINA) approach, pioneered by Piketty et al. (2018), whose goal is to allocate all national income to individuals. Nonetheless, the methods developed in the literature are subject to considerable debate. We use a method to allocate profits that is inspired by this research, and implement it to the best extent possible using publicly available data.

The government's share of profits comprises corporate tax revenue, as well as revenue from the personal taxation of individual profit holdings. We set the share of profits that the government collects through corporate taxation to 25%, based on the corporate tax rate in 2012. We assume that 30% of profits flow overseas. We calculate this using information from the UK National Accounts: we take the ratio of income distributed by corporations that flows overseas to net operating surplus excluding imputed rents from owner-occupied housing.

We assume that the remaining 45% of profits are distributed to UK households in proportion to the share of dividend income that they receive. Saez and Zucman (2016) use a combination of dividend income and realized capital gains to estimate stock ownership. Smith et al. (2020) use a weighted average of dividend income and capital gains, but with most of the weight assigned to dividend income, which they find a better predictor of stock ownership. There is no publicly available data that contains information on the joint distribution of capital gains and taxable income for UK individuals. Instead, we use the Survey of Personal Incomes (SPI; HM Revenue and Customs, CS and TD KAI Personal Taxes. (2020)), which records dividend income received by individuals to estimate the relationship between individuals' total income and the amount they receive from dividends. Table E.1 shows the mean dividend income for individuals with different levels of total income. Individuals earning more than £40,000 (roughly the top 10%), receive approximately 70% of dividend income recorded on tax records.

We map this into the share of dividend income received by *households* (as opposed to individuals). To do this, we use the Living Costs and Food Survey, which contains information on the total (but not dividend) income received by individual household members. We use the mean dividend income by banded personal income shown in Table E.1 to impute dividend income for individuals in participating households in the Living Costs and Food Survey 2012, which we then sum for all members in the household. Table E.2 shows the distribution of dividend income across households. Note that the distribution across households is less skewed than the distribution across individuals, reflecting the fact that many households consist of one high and one lower earner.

Table E.1: *Mean dividend income by banded personal total income*

Total income	Mean dividend income
0-2.5k	39
2.5-5k	51
5-7.5k	46
7.500-10k	56
10-12.5k	73
12.5-15k	98
15-20k	143
20-30k	302
30-40k	758
40k+	3436

Notes: We use data from the Survey of Personal Incomes in 2012. The table shows the mean dividend income (excluding dividends received from owner-managed companies) for individuals with total personal incomes in the bands shown in the first column.

Table E.2: *Distribution of dividends across household equivalized income distribution*

(1) Equivalized hh income	(2) % hh	(3) Mean div income	(4) ATR divs	(5) % dividends Pre-tax	(6) Post-tax
0-5k	12.8	60	0.00	0.8	0.9
5-10k	11.9	141	0.00	1.8	2.0
10-15k	17.6	173	0.00	3.3	3.7
15-25k	29.6	508	0.00	16.3	17.7
25-35k	13.5	1721	0.04	25.2	25.8
34-45k	9.3	2933	0.10	29.4	28.5
45k+	5.3	4021	0.17	23.2	21.3

Notes: We use the mean dividend income by banded personal income shown in Table E.1 to impute dividend income for individuals in participating households in the Living Costs and Food Survey 2012. We sum dividend income for all members in the household. We construct equivalized total household income (using the OECD-modified equivalence scale) and put households into bands, listed in column (1). Column (2) shows the share of households in each band, column (3) shows the mean amount of dividend income per household for each band, and column (4) shows the average personal tax paid on dividends for households in each band. Columns (5) and (6) show the share of total dividend income (pre and post- dividend tax, respectively) that households in each band receive.

Dividend income is subject to personal taxation. Table E.1 reports the average tax rate on dividends for each household income band. After taking account of this (and corporate tax), the government share in profits is 29%. Post-tax profits are distributed to households according to column (6) in the table.

E.3 Solution algorithm

Obtaining the optimal tax rate (or vector of rates) entails solving an algorithm that consists of an outer loop and several inner loops. The solution of the outer loop is the optimal tax vector, the solution to the inner loops are, given a candidate tax vector, the equilibrium price vector and the matrix of derivatives of the optimal price vector with respect to the tax vector.

Inner loops Given the tax vector $(\tau_1, \ldots, \tau_K)$, equilibrium prices, $\boldsymbol{p'} = (p'_1, \ldots, p'_J)$, are obtained as the solution to the system of equations: for $j = 1, \ldots, J$

$$q_j(\boldsymbol{p'}) + \sum_{j' \in \mathcal{J}_f} (p'_{j'} - \mathbb{1}\{j' \in \mathcal{J}_k\}\tau_k - c_{j'}) \frac{\partial q_{j'}(\boldsymbol{p'})}{\partial p_j} = 0.$$

The $J \times K$ matrix of derivatives $\frac{d\mathbf{p}'}{d\boldsymbol{\tau}}$ is obtained by solving $k = 1, \ldots, K$ systems of equations of the form: for $j = 1, \ldots, J$

$$\sum_{j' \in \mathcal{M}} \frac{\partial q_j}{\partial p'_{j'}} \frac{dp_{j'}}{d\tau_k} + \sum_{j' \in \mathcal{J}_f} \left(\frac{dp'_{j'}}{d\tau_k} - \mathbb{1}\{j' \in \mathcal{J}_k\} \right) \frac{\partial q_{j'}}{\partial p_j} +$$
$$\sum_{j' \in \mathcal{J}_f} (p'_{j'} - \mathbb{1}\{j' \in \mathcal{J}_k\}\tau_k - c_{j'}) \sum_{j'' \in \mathcal{M}} \frac{\partial^2 q_{j'}}{\partial p_j \partial p_{j''}} \frac{dp'_{j''}}{d\tau_k} = 0.$$

Outer loop We use three alternative methods for solving the outer loop:

1. The optimal tax vector can be expressed in the form: $\boldsymbol{\tau}^* = G(\boldsymbol{\tau}^*)$ (see equation (2.2), for the case of a single sugar-sweetened beverage tax rate). One solution method involves iterating on this equation: (1) guess a tax vector $\boldsymbol{\tau}^r$, (2) solve the inner loops, (3) compute $G(\boldsymbol{\tau}^r)$, (4) set $\boldsymbol{\tau}^{r+1} = G(\boldsymbol{\tau}^r)$ and repeat until convergence. This method is relatively quick but has the disadvantage that it is not suitable for imposing constraints on the government's objective function.

2. When solving for the multi-tax rate system subject to constraints (see Section VC of the paper) we instead numerically maximize the social welfare function subject to the constraint. For each iteration of the algorithm, we must solve the inner loops.

3. A third solution method is a grid search over the tax rate (feasible in the single rate, but not multi rate, case). We use this method to draw Figure E.1.

E.4 Multi-rate tax system

We consider two variants of the multi-rate system. The first variant allows the government to vary taxes rates across 12 different drink types,[7] subject to there being no deterioration in the government's budget, inclusive of the budgetary externality. The second variant requires all tax rates be non-negative, ruling out subsidies on alternatives to sugar-sweetened beverages.

Table E.3 summarises the results. The optimal system under no budget deterioration involves subsidizing non-sugar sweetened drinks, and entails, on average, lower tax rates on sugar-sweetened beverages than under the optimal single rate system. This leads to increases in consumer surplus and profits, with the fall in tax revenue offset by a reduction in budgetary external costs. Overall, welfare is 80% higher than under the optimal single rate system. The third column shows the optimal policy when subsidies are prohibited – this leads to a smaller rise in welfare (of 17%) relative to the optimal single rate system.

Table E.3: *Multi-rate taxation*

	(1) Single rate	(2) Multi-rate: No budget deterioration	(3) Multi-rate: No subsidy
Tax rate (p/10oz) for:			
Sin products	5.97	4.28*	5.53*
Alternatives	0.00	-4.01*	0.00*
Price change for:			
Sin products	27.3%	22.5%	27.1%
Alternatives	-0.7%	-15.8%	-0.6%
Consumption change for:			
Sin products	-37.7%	-37.7%	-35.8%
Alternatives	9.4%	35.1%	8.5%
Welfare components (£m):			
Private welfare	-790	299	-670
Consumption	*-747*	*276*	*-634*
Profit holdings	*-43*	*23*	*-37*
Gov. budget	957	0	866
Excise tax rev.	*522*	*-572*	*424*
Profits tax rev.	*-74*	*40*	*-63*
Ext. cost savings	*509*	*532*	*505*
Total welfare (£m)	167	299	196

* *average tax rate*

Notes: Each column corresponds to a tax system as described in the text. Numbers summarize the effect of policy when the social marginal welfare weight on foreign individuals is 0 and on domestic individuals is $1/\tilde{y}_i$. Welfare numbers are per annum and report the change relative to no drinks taxation. Total welfare = Private welfare+Government budget.

[7] These are: 5 sugar-sweetened beverage drinks types (cola, lemonade, other sodas, juices and energy/sports drinks) and 7 drinks types comprising alternatives to sugar-sweetened drinks (pure fruit juices, milk drinks, plus diet counterparts of cola, lemonade etc.).

E.5 Sugar tax

Table E.4 compares the optimal sugar tax, assuming no changes in products' sugar contents, to the optimal single rate volumetric tax. The sugar tax, which entails a rate of 2.28 pence per 10g of sugar, results in a larger increase in the average price of sugar-sweetened beverages (30.3% relative to 27.3% under the volumetric tax). This leads to slightly larger falls in consumer welfare and lower tax revenue, but a considerably larger fall in externality costs (£572m compared with £509m under the volumetric tax). Overall the sugar tax raises welfare by £208m, which is higher than the £167m rise under the optimal single rate volumetric tax. Similarly to volumetric taxation, the costs, in terms of forgone welfare gains, of ignoring market power when setting the sugar tax are substantial (at 31%).

Table E.4: *Sugar taxation*

		Change (relative to zero tax) in:							
		Welfare components (£m)						
	% Δ in SSB	Private, welfare, from:		Tax revenue:		Ext. cost	Total	% loss under
	price	Cons.	Profits	Sin tax	Profit tax	savings	welfare	naivety
Volumetric tax	27.3%	-747	-43	522	-74	509	167	40%
Sugar tax	30.3%	-762	-42	511	-72	572	208	31%

Notes: Row (1) shows the effects of the optimal volumetric tax for reference. It repeats some of the information in row (1) of Table 5. Row (2) shows the effects of the optimal tax on the sugar in sweetened beverages under the assumption of no product reformulation. Numbers in the final column show the % of welfare gains from optimal policy forgone when the government sets policy ignoring market power. Numbers summarize the effect of policy when the social marginal welfare weight on foreign individuals is 0 and on domestic individuals is $1/\tilde{y}_i$. *Welfare numbers are per annum and report the change relative to no drinks taxation. Total welfare = Private welfare+Tax revenue+External cost savings. Numbers in the final column show the % of welfare gains from optimal policy forgone when the government sets policy ignoring market power.*

We also consider the effect of a sugar tax when firms reoptimize both price and the sugar content of their products. We model firms' decision over product sugar content following Barahona et al. (2020). In their model a firm can substitute sugar for an alternative input, whilst keeping the taste (and hence consumer valuation) of the product unchanged, but at the cost of increasing production costs. With no tax in place the firm will choose the cost minimizing sugar level. With a tax in place the firm may change product sugar content, trading-off increased production costs with a reduced tax liability.

Consider firm $f = 1, \ldots, F$, which owns products $\mathcal{J}_f$ – it chooses the vector of tax-inclusive prices for these products $\{p_j\}_{j\in\mathcal{J}_f}$. Denote the subset of products in $\mathcal{J}_f$ that are sugar-sweetened beverages by $\mathcal{J}_f^{\mathcal{S}}$ and the remaining products by $\mathcal{J}_f^{\mathcal{N}}$. The firm chooses the sugar content of each product in set $\mathcal{J}_f^{\mathcal{S}}$. We denote by z_j^* the production cost-minimizing sugar content of product $j \in \mathcal{J}_f^{\mathcal{S}}$ (conditional on the taste of the product).

In the absence of a sugar tax, the firm's problem is

$$\max_{\{p_j\}_{j\in\mathcal{J}_f},\{z_j\}_{j\in\mathcal{J}_f^S}} \sum_{j\in\mathcal{J}_f^S}(p_j - c_j(z_j))q_j(\mathbf{p}) + \sum_{j\in\mathcal{J}_f^N}(p_j - c_j)q_j(\mathbf{p})$$

The first order conditions are: for $f = 1,\dots,F$

$$q_j + \sum_{j'\in\mathcal{J}_f^S}(p_{j'} - c_{j'}(z_{j'}))\frac{\partial q_{j'}}{\partial p_j} + \sum_{j'\in\mathcal{J}_f^N}(p_{j'} - c_{j'})\frac{\partial q_{j'}}{\partial p_j} = 0 \quad \text{for all } j\in\mathcal{J}_f$$

$$c_j'(z_j) = 0 \quad \text{for all } j\in\mathcal{J}_f^S$$

By definition, the sugar contents that satisfy these conditions are $z_j = z_j^*$ for all $j\in\mathcal{J}_f^S$ and all f.

With a sugar tax in place, we can define the tax-inclusive marginal cost as $C_j(z_j) = \tau z_j + c_j(z_j)$ for all $j\in\mathcal{J}_f^S$ and f. The first order conditions that characterize the firms' optimal choices are then: for $f = 1,\dots,F$

$$q_j + \sum_{j'\in\mathcal{J}_f^S}(p_{j'} - C_{j'}(z_{j'}))\frac{\partial q_{j'}}{\partial p_j} + \sum_{j'\in\mathcal{J}_f^N}(p_{j'} - c_{j'})\frac{\partial q_{j'}}{\partial p_j} = 0 \quad \text{for all } j\in\mathcal{J}_f$$

$$C_j'(z_j) = 0 \quad \text{for all } j\in\mathcal{J}_f^S$$

Hence the optimal sugar choice of product k satisfies: $\tau + c_k'(z_k) = 0$.

We assume that the marginal costs function takes the following quadratic form:

$$c_j = \bar{c}_j + \frac{\nu}{z_j^*}(z_j^* - z_j)^2,$$

where $\bar{c}_j$ denotes the cost-minimizing marginal cost (which corresponds to production decisions in the absence of a sugar tax) and ν controls the marginal cost of reformulation. Along with the firm's first order condition for sugar choice, this implies that with a sugar tax in place:

$$\frac{(z_j^* - z_j)}{z_j^*} = \frac{\tau}{2\nu}.$$

Hence the percentage reduction in a product's sugar content is proportional to the sugar tax rate τ and inversely proportional to the reformulation cost ν. Under a sugar tax the increase in the tax-inclusive marginal cost of product j is:

$$\begin{aligned}\Delta C_j(\nu,\tau) =& \tau\left(z_j^* - \frac{z_j^*\tau}{2\nu}\right) + \frac{\nu}{z_j^*}\left(\frac{\tau z_j^*}{2\nu}\right)^2 \\ =& \tau z_j^* - \frac{z_j^*\tau^2}{4\nu}.\end{aligned}$$

Note that, the sugar tax changes the relative marginal cost of two sugary products according to:

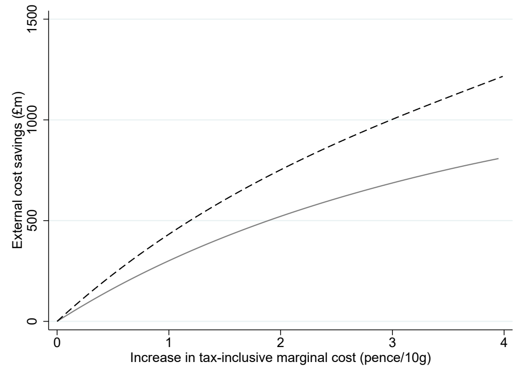
$$\frac{\Delta C_k(\nu,\tau)}{\Delta C_j(\nu,\tau)} = \frac{z_k^*}{z_j^*}.$$

Hence, for every reformulation cost, ν, there is a sugar tax rate that results in the same vector of tax-inclusive costs, $\{C_j(\nu,\tau)\}_{j\in\mathcal{M}}$, and hence equilibrium prices and quantities and consumer surplus and profits.

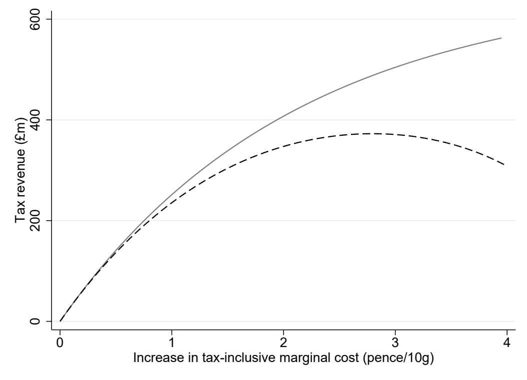
In Figure E.1 we compare the implications of a sugar tax when reformulation costs are prohibitive and when they are relatively low (ν=5 pence). Instead of plotting how welfare varies with the tax rate, τ, we plot how it varies with $\frac{\Delta C}{z^*}$ (which is a monotonically increasing function of τ). Conditional on $\frac{\Delta C}{z^*}$ the market equilibrium is the same (regardless of reformulation costs), and welfare differences as reformulation costs fall are driven purely by whether larger reductions in external costs offset higher production costs. In Figure E.2 we plot how the optimal sugar tax rate and associated welfare gain vary with the reformulation costs. As the cost of reformulation falls, firms choose to remove more of the sugar from their sugar-sweetened beverages. The figure shows that this results in larger welfare gains from optimal sugar taxation. This reason for this is that larger falls in external costs from sugar outweigh raised production costs. Even though firms make privately optimal decisions over product sugar content, the externalities form sugar are sufficiently large that these private decision improve social welfare.

Figure E.1: *Impact of sugar tax with input substitution*

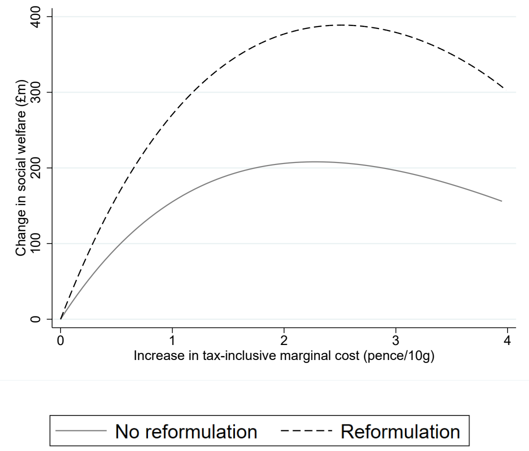
(a) External cost savings



(b) Tax revenue



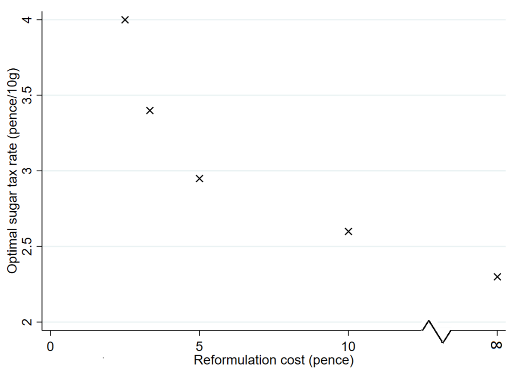
(c) Welfare



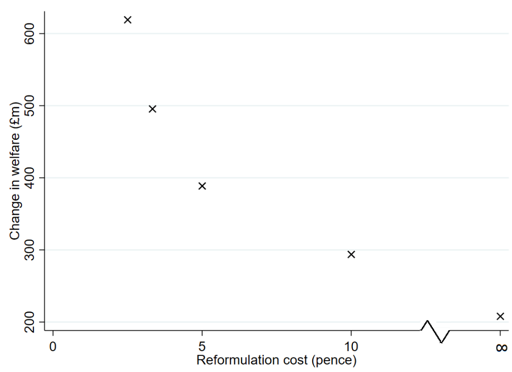
Notes: Graphs show the impact on external costs (panel (a)), tax revenue (panel (b)) and social welfare (panel (c)) from sugar taxation when firms do not reformulate products, and optimally reformulate. On the horizontal axis we plot the increase in the tax-inclusive marginal cost due to a sugar tax. This is a monotonic function of the sugar tax rate, and means that conditional on a given marginal cost increase, the implications of the sugar tax for private welfare are the same in the two scenarios.

Figure E.2: *Variation in optimal sugar tax and welfare gain with reformulation costs*

(a) Optimal tax rate



(b) Welfare gain



Notes: Graphs show how the optimal sugar tax rate (panel (a)) and associated welfare gain (panel (b)) vary with the reformulation cost parameter ν.

References

Barahona, N., C. Otero, S. Otero, and J. Kim (2020). Equilibrium Effects of Food Labeling Policies. *Available at SSRN: 3698473*.

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

Browning, M. J. (1983). Necessary and Sufficient Conditions for Conditional Cost Functions. *Econometrica 51*(3), 851–856.

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

Hendel, I. and A. Nevo (2006). Sales and Consumer Inventory. *RAND Journal of Economics 37*(3), 543–561.

HM Revenue and Customs, CS and TD KAI Personal Taxes. (2020). Survey of Personal Incomes, 2011-2012: Public Use Tape. *3rd Edition. UK Data Service. SN: 7472, DOI: http://doi.org/10.5255/UKDA-SN-7472-3*.

OECD (2019). Prices in the health sector. OECD Indicators.

Piketty, T., E. Saez, and G. Zucman (2018). Distributional National Accounts: Methods and Estimates for the United States. *Quarterly Journal of Economics 133*(2), 553–609.

Public Health England (2017). Health matters: Obesity and the food environment. https://www.gov.uk/government/publications/health-matters-obesity-and-the-food-environment/health-matters-obesity-and-the-food-environment–2.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature 50*(1), 3–50.

Saez, E. and G. Zucman (2016). Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. *Quarterly Journal of Economics 131*(2), 519–578.

Smith, M., O. Zidar, and E. Zwick (2020). Top Wealth in America: New Estimates and Implications for Taxing the Rich. pp. 117.

University of Oxford (2018). Food DB. *https://www.ndph.ox.ac.uk/food-ncd/research/fooddb*, (accessed May 29, 2019).

Wang, Y. C., P. Coxson, Y.-M. Shen, L. Goldman, and K. Bibbins-Domingo (2012). A Penny-Per-Ounce Tax On Sugar-Sweetened Beverages Would Cut Health And Cost Burdens Of Diabetes. *Health Affairs 31*(1), 199–207.

World Health Organization (2015). *Sugar Intake for Adults and Children.* Geneva.