Institute for Fiscal Studies

Martin O'Connell
Kate Smith

21/30 Working paper

Optimal sin taxation and market power

Economic and Social Research Council

Optimal sin taxation and market power

Martin O'Connell and Kate Smith*

September, 2021

Abstract

This paper studies the design of sin taxes when firms exercise market power. We outline an optimal tax framework that highlights how market power impacts the efficiency and redistributive properties of sin taxation, and quantify these effects in an application to sugar-sweetened beverage taxation. We estimate a detailed model of demand and supply for the UK drinks market, which we embed in our tax design framework to solve for optimal sugar-sweetened beverage tax policy. Positive price-cost margins on drinks create allocative distortions, which act to lower the optimal rate compared with a perfectly competitive setting. However, since profits accrue to the rich, this is partially mitigated under social preferences for equity. Overall, ignoring market power when setting the optimal sugar-sweetened beverage tax rate leads to welfare gains that are 40% below those at the optimum. We show that moving from a single tax rate on sugar-sweetened beverages to a multi-rate system can result in further substantial welfare gains, with much of these gains realized by instead taxing sugar content directly.

Keywords: externality, corrective tax, market power, profits, redistribution

JEL classification: D12, D43, D61, D62, H21, H23, L13

Acknowledgements: The authors would like to thank Richard Blundell, Pierre Dubois, Rachel Griffith, Rory McGee, Aviv Nevo, Cormac O'Dea, Ariel Pakes, Imran Rasul, and participants at seminars at Bristol, Brussels, Chicago Booth, INRAE, Michigan, Oxford, Warwick, UCL, UIUC, WUSTL and Wisconsin-Madison for many very helpful discussions. The authors gratefully acknowledge financial support from the Economic and Social Research Council (ESRC) under the Centre for the Microeconomic Analysis of Public Policy (CPP; grant number ES/T014334/1), under the Open Research Area (ORA; grant number ES/VO13513/1) and the British Academy under pf160093. Data supplied by Kantar FMCG Purchase Panel. The use of Kantar FMCG Purchase Panel data in this work does not imply the endorsement of Kantar FMCG Purchase Panel in relation to the interpretation or analysis of the data. All errors and omissions remain the responsibility of the authors. This article supersedes O'Connell and Smith (2020).

*O'Connell is at the University of Wisconsin-Madison and the Institute for Fiscal Studies and Smith is at University College London and the Institute for Fiscal Studies. **Correspondence**: moconnell9@wisc.educ and kate.smith@ifs.org.uk

1 Introduction

Taxes that aim to discourage socially costly consumption have long been applied to alcohol and tobacco, and are increasingly being used to tackle the harms associated with unhealthy eating and carbon emissions. There is a long literature that considers the design of "sin" taxes, dating back to Pigou (1920) and Diamond (1973), who argue that the optimal sin tax rate equals the average social harm associated with marginal consumption. Subsequent work shows how optimal policy is affected when other distortionary taxes are used and when the policymaker places more weight on the welfare of the poor than the rich.[1] This work has largely focused on policymaking under perfect competition.

Yet many papers demonstrate the pervasive nature of market power, with some recent evidence pointing to its growing importance.[2] An emerging literature illustrates empirically the role that market power can play in optimal policy design, highlighting the importance of firms' strategic pricing responses and product markups.[3] Market power generally gives rise to positive profits, some of which flow to the government via corporate tax receipts, with the remainder accruing disproportionately to the rich.[4] The distribution of profit holdings across individuals gives rise to an important additional redistributive effect of taxation, not accounted for in the existing literature on sin tax design.

In this paper, we show conceptually how market power affects optimal sin taxation, and empirically quantify this in an application to sugar-sweetened beverage taxation. We analyze an optimal tax problem, which highlights how market power impacts both the efficiency and redistributive effects of policy, and motivates our empirical approach. Harnessing detailed consumer level data, we estimate an equilibrium model of the UK market for drinks, and embed this into the tax problem to solve for optimal sugar-sweetened beverage tax policy. We show that allocative distortions associated with positive margins on drinks reduce the optimal sin tax rate, compared with a perfectly competitive setting, but this reduction is partially offset by the fact that profits accrue disproportionately to those with lower social

[1] For instance, Sandmo (1975) and Kopczuk (2003) show that the logic of Pigou and Diamond continues to apply when other distortionary taxes are in place; Kaplow (2012) and Allcott et al. (2019a) consider the interaction of sin taxation and redistribution.

[2] See Einav and Levin (2010) for a survey of the industrial organization literature, and Syverson (2019) for recent macroeconomic research on market power, with De Loecker et al. (2020) a prominent example.

[3] See, for instance, Fowlie et al. (2016), Miravete et al. (2018, 2020), Conlon and Rao (2015), Tebaldi (2017) and Polyakova and Ryan (2019).

[4] See, for example, evidence in Cooper et al. (2016).

welfare weights (foreign residents and the rich). Overall, ignoring market power leads to welfare gains that are 40% below those achieved by optimal policy.

Our tax design problem, which we present in Section 2, builds on the canonical framework (see, for instance, Kleven (2020)) by incorporating strategic firms. We study a government's choice over a set of linear commodity taxes levied in a specific market of interest. A set of heterogeneous individuals choose their labor supply and a consumption bundle. The product set includes: "sin" products, which generate externalities; alternatives, which do not create externalities; and a composite good, which represents consumption outside the market of interest. The products are supplied by firms that exercise market power, strategically reoptimize prices in response to tax policy, and earn positive profits, which are distributed to individuals in a potentially unequal manner. The government sets policy anticipating the strategic responses of firms (as, for instance, in Miravete et al. (2018)[5]), accounts for spillovers to other tax bases, and may place more weight on the welfare of the poor than the rich. We characterize the optimal single-rate sin tax, which facilitates comparison to existing sin tax results under perfect competition and clarifies the key forces that determine optimal tax policy.

The optimal rate can be written as an implicit function of four components, three of which reflect efficiency considerations, with the fourth reflecting equity. The first is an externality correcting term, echoing Diamond (1973) – the more the tax achieves large falls in the most socially costly consumption, the more effective it is at combating externality distortions, and all else equal, the higher will be the optimal rate. The second component reflects distortions from the exercise of market power. If equilibrium price-cost margins for the sin products are relatively high, this will act to lower the optimal tax rate, echoing the argument made by Buchanan (1969).[6] However, we show that it is the *relative* margins of the sin compared with the untaxed alternative products, and the strength of switching between them, that are relevant for the optimal rate. The third component captures the interaction of the sin tax with the pre-existing non-linear labor tax. When the latter is held fixed, any fiscal externality due to erosion of the labor tax base from higher commodity taxation, all else equal, acts to lower the optimal tax rate.

The final component of the optimal sin tax rate reflects distributional concerns, which affect policy through two channels. The first reflects consumption patterns; the more the sin products are consumed by low income individuals, the more re-

[5]Miravete et al. (2018) show how firms' pricing responses changes the peak and shape of the Laffer curve in the liquor market

[6]Buchanan (1969) argues that since a monopolist will choose a quantity below the competitive level, levying a Pigouvian tax (which ignores this distortion) is suboptimal.

gressive is the tax and the lower is the optimal rate. This channel is highlighted by Allcott et al. (2019a). The second channel reflects the distribution of net-of-tax profit holdings. The more that any reduction in profits due to tax policy is incident on rich individuals, all else equal, the more progressive is the tax and the higher is the optimal rate.

In addition to highlighting how market power impacts optimal sin tax policy, our theoretical analysis motivates our empirical strategy. Often, work in empirical public finance considers the welfare implications of tax reforms by expressing their effect in terms of externally valid and estimable elasticities (or "sufficient statistics"). This approach, which extends to optimal tax analysis under the assumption of isoelastic preferences, has the advantage of allowing relatively transparent identification arguments and the use of quasi-experimental estimation methods (Chetty (2009)). However, under market power, the optimal tax expression does not straightforwardly map into externally valid elasticities, as it depends on product level price-cost margins, which are typically not directly observable (see Bresnahan (1989)) and depend on the market structure and full set of product level own- and cross-price elasticities. Our empirical approach is therefore to estimate a detailed market level model of consumer choice and firm pricing competition and to embed this into our optimal tax framework.

Our application is to the taxation of sugar-sweetened beverages. Consumption of these products is strongly linked to diet-related disease, which creates externalities through increased societal costs of funding both public and insurance based health care (see Allcott et al. (2019b)). In recent years, motivated by public health concerns, a number of countries and localities have introduced taxes on these products.[7] As we illustrate in Section 3, the UK drinks market is highly concentrated, and characterized by large multi-product firms that offer highly recognizable branded products. Market power is therefore likely to be important.

We specify an empirical demand and supply model of the drinks market, outlined in Section 4, in the broad tradition of Berry et al. (1995). We model the discrete choice that consumers make between the many products in the market (or allocating all of their spending outside the market), and estimate the model using micro level longitudinal data. We incorporate rich preference heterogeneity, including with income and a measure of total dietary sugar, which are important for capturing both redistributive and externality correcting aspects of tax policy. We exploit price changes that are agreed in advance by drinks firms and retailers, but that create differential variation across consumers, as a source of identifying price variation.

[7]As of May 2021, 44 countries and 8 US cities had some form of sugar-sweetened beverage tax (GFRP (2021)).

We identify product level marginal costs by coupling our demand estimates with a game-theoretic pricing model.

Our demand and supply estimates, which we summarize in Section 5, highlight that drinks products are highly differentiated and, due to this and the multi-product nature of their portfolios, drinks firms exercise a substantial degree of market power. We find that the average Lerner index at observed prices is around 0.5 for both sugar-sweetened drinks and alternatives. Our estimates show that, in response to a change in the price of one sugar-sweetened drink, consumers are most willing to switch to a similar sugar-sweetened drink. However, in response to a rise in the price of all sugary drinks, substitution to alternative (non sugar-sweetened) drinks is substantial. Market power among alternative drinks is therefore relevant for sugar-sweetened beverage tax policy. Our model allows us to determine how firms adjust prices (and hence price-cost margins) in response to tax policy. We find that pass-through of a tax on sugar-sweetened drinks is slightly above 100%, and show that this is in line with (out-of-sample) observed price changes following the introduction of the UK's Soft Drinks Industry Levy, as well as with findings in other jurisdictions.

In Section 6 we present our empirical tax results. We first consider the rate set by a government without redistributive concerns. The efficiency maximizing tax rate on sugar-sweetened beverages is 4 pence per 10 oz, which leads to a 19% rise in their average price. If the government ignores all distortions from market power it would set a suboptimally high rate of 12 pence per 10 oz, which leads to efficiency *losses*.[8] The positive price-cost margins of sugar-sweetened drinks act to lower the efficiency maximizing tax rate, though the strength of this effect is offset by the positive margins of alternative drinks.

We then consider optimal sin tax policy, assuming that the government does have redistributive concerns. In this case both the share of profits collected by the government in corporate and dividend taxes, and the distribution of net-of-tax profits across individuals affect optimal policy. We measure the allocation of corporate profits using information from national accounts and the distribution of dividend income across households. We find that under our baseline social preferences, the optimal sin tax rate on sugar-sweetened beverages is 6 pence per 10 oz,[9] which is 50% higher than the efficiency maximizing rate. This increase reflects the net impact of two offsetting forces. On the one hand, sugar-sweetened beverage consumption is highest among relatively low income consumers, which acts to lower the

[8] Fowlie et al. (2016) find a similar result for carbon abatement policy in the cement industry (where entry costs lead sellers to exercise market power).

[9] In practice, US and UK sugar-sweetened beverage taxes range from 7 to 15 pence per 10 oz.

optimal rate. On the other hand, post-tax profits are disproportionately owned by high income consumers (or flow overseas), which increases the progressivity of the tax, thus raising the optimal rate. Overall, the second effect dominates. We show that these results apply under a wide range of social preferences. In addition, we show that even if the government places zero welfare weight on the post-tax profits flowing to individuals, ignoring market power when setting policy would lead it to overshoot the optimal rate by 70% (with the resulting welfare gains 26% below optimal) due to spillovers to the corporate and dividend tax bases.

Most jurisdictions that tax sugar-sweetened beverages do so using a single rate levied volumetrically (i.e. per oz). We study the extent to which alternative tax instruments might improve welfare. In some settings, retailing is run by a state monopoly and therefore the government can set product-specific prices, which can lead to substantial gains (see Conlon and Rao (2015) and Miravete et al. (2020)). Product specific taxes are not feasible in the unregulated non-alcoholic drinks market. We therefore consider a multi-rate tax system that entails 12 tax rates, each applying to different drinks types (including non sugar-sweetened beverages).[10] If the government is free to choose the 12 rates, subject to policy not leading to a deterioration in the government's budget, the associated gains are 80% larger than under the single rate system. However, this tax system leads to negative drinks tax revenue (made up for by reduced public externality costs), and entails negative tax rates for some drinks types. Requiring instead that all tax rates be non-negative – a feature of most commodity taxes – lowers the welfare gains, though they remain higher than under the single rate. Varying rates allows the government both to better target the most socially costly consumption,[11] and to account for variation in margins across sugar-sweetened beverage types.

An alternative to volumetric sugar-sweetened drinks taxes is to directly tax the sugar in these products. Grummon et al. (2019) argue that since this tax is more closely related to the source of social harm, it would result in substantial welfare gains.[12] A tax on a production input may also induce input substitution (see Ganapati et al. (2020)), which could enhance the externality correcting role of the tax. We show that if there is no input substitution, a sugar tax would result in welfare gains that are 25% above those under the single rate volumetric tax. When firms do respond to the tax by reducing product sugar content, the welfare consequences depend on whether larger falls in externalities exceed increased production costs. We

[10]The drinks types are colas, diet colas, lemonades, diet lemonades etc. Taxation at this level is not unprecedented: alcohol taxes typically set separate rates for beers, wine and liquor.

[11]A point made by Griffith, O'Connell, and Smith (2019) in the context of alcohol taxation.

[12]Relatedly, Jacobsen et al. (2020) document substantial deadweight losses when externality correcting taxes are not directly levied on the source of externality.

consider an extension to our empirical model of firm competition (following Barahona et al. (2021)), which allows firms to reformulate their products in response to the sugar tax. We show that for all plausible reformulation costs, firms' privately optimal decisions to reformulate products act to enhance social welfare.

In addition to contributing to the tax design literature, our work relates to a growing literature on quantifying the impacts of sugar-sweetened beverage taxation. This includes a set of papers that use the implementation of these taxes to estimate their effect on prices and quantities,[13] and a set that use estimates of consumer demand based on periods and locations with no tax in place to simulate the introduction of taxes similar to those used in practice.[14] Like us, Allcott et al. (2019a) ask: what is the optimal sugar-sweetened beverage tax? They develop a novel characterization of the optimal corrective commodity tax rate under general preference heterogeneity and an optimally set labor tax schedule, and incorporate consumer misoptimization into the corrective tax motive. Our work complements theirs by showing how market power impacts optimal sin tax policy.

2 Sin tax design

We consider a setting in which a government chooses what tax rate(s) to levy in a market with multiple products, including a set associated with externalities. We allow for the possibility that these products are sold by firms that exercise market power. The government sets the tax rates while accounting for interactions with other parts of the tax system, and balances distortions from externalities and market power with equity considerations.

2.1 Set-up

Individuals

There is a continuum of individuals (or consumers) indexed i. Individuals supply labor in a competitive labor market to generate pre-tax earnings, z^i, which are subject to a non-linear earnings tax, $\mathcal{T}(z^i)$. Each individual also potentially receives income arising from their holding of profits, which are generated by the sale of consumption goods. We denote total profits by Π and i's share of profits by $\delta^i \geq 0$,

[13] See, for instance, Bollinger and Sexton (2018) and Rojas and Wang (2017) who study the Berkeley tax, Seiler et al. (2021) and Roberto et al. (2019) who study the Philadelphian tax, and Grogger (2017) who study the Mexican tax. For a full survey of the recent literature see Griffith et al. (2019).

[14] These papers include Bonnet and Réquillart (2013), Wang (2015), Harding and Lovenheim (2017), Chernozhukov et al. (2019) and Dubois et al. (2020).

where $\int_i \delta^i di = 1$. Individual profit holdings are subject to a (potentially non-linear) tax, $T_\Pi(\delta_i \Pi)$.[15]

Net income is spent on consumption goods: $\mathbf{x}^i_S = \{x^i_j\}_{j\in\mathcal{S}}$ is a quantity vector of 'sin' products, belonging to the set $\mathcal{S}$, consumption of which potentially creates externalities; $\mathbf{x}^i_N = \{x^i_j\}_{j\in\mathcal{N}}$ is a quantity vector of products belonging to the set $\mathcal{N}$, which are in the same market as those in $\mathcal{S}$, but do not give rise to externalities; let $\mathcal{M} = \mathcal{S}\bigcup\mathcal{N}$ and $n(\mathcal{M}) = J$. x^i_O denotes the quantity of a composite consumption good that represents all goods outside of those in market $\mathcal{M}$. Consumers face the tax-inclusive price vector $\mathbf{p} = (\mathbf{p}_\mathcal{S}, \mathbf{p}_\mathcal{N}, 1)$, which embeds the normalization that the price of the composite good, x^i_O is 1. The individual's budget constraint is $\sum_{j\in\mathcal{M}} p_j x^i_j + x^i_0 = z^i - \mathcal{T}(z^i) + \delta^i\Pi - T_\Pi(\delta^i\Pi)$. We assume that the earning tax is piece-wise linear, denote $\frac{d\mathcal{T}}{dz^i} \equiv \tau^i_z$, and define virtual labor income by $G \equiv \tau^i_z z^i - \mathcal{T}(z^i)$. We denote the sum of virtual and unearned (profit) income by $Y^i \equiv G^i + \delta^i\Pi - T_\Pi(\delta^i\Pi)$.

Each individual chooses a bundle, $(\mathbf{x}^i_S, \mathbf{x}^i_N, x^i_O, z^i)$, to maximize utility, $U^i(\mathbf{x}^i_\mathcal{S}, \mathbf{x}^i_\mathcal{N}, x^i_O, z^i)$, subject to their budget constraint.[16] Consumer i's product demands are denoted by $x^i_j = x^i_j(\mathbf{p}, (1-\tau^i_z), Y^i)$, earnings supply by $z^i = z^i(\mathbf{p}, (1-\tau^i_z), Y^i)$, and indirect utility by $V^i = V^i(\mathbf{p}, (1-\tau^i_z), Y^i)$. We denote the marginal utility of income by $\alpha^i \equiv \partial V^i/\partial Y^i$.

Firms

Products are defined such that each product j in market $\mathcal{M}$ is produced by a single firm (firms can produce multiple products). We denote the market demand for product j by $X_j(\mathbf{p}_\mathcal{M}, \boldsymbol{\Upsilon}_\mathcal{M}) = \int_i x^i_j di$, where $\boldsymbol{\Upsilon}_\mathcal{M} = (\Upsilon_1, \ldots, \Upsilon_J)$ and Υ_j denotes all non-price attributes that influence demand for product j,[17] and we denote the product's marginal cost by c_j.[18] We consider a system of linear excise taxes that apply to the products in market $\mathcal{M}$; at its most general it specifies a vector of product level tax rates, $\boldsymbol{\tau}_\mathcal{M} = (\tau_1, \ldots, \tau_J)$. Note that constant marginal costs means that if market $\mathcal{M}$ were perfectly competitive, profits would be zero and prices would shift mechanically with commodity taxes.

[15] $T_\Pi(\delta_i\Pi)$ may capture both corporate taxes and individual (e.g., dividend) taxation. For simplicity we write the earning and profit taxes as additively separable. This is unimportant for our analysis, which would not be materially affected by non-separabilities in $\mathcal{T}(.)$ and $T_\Pi(.)$.

[16] We assume labor earnings (rather than hours worked) directly enter individual utility, implying that individuals are price takers in the labor market.

[17] This may include non-price attributes of the products, and consumer specific attributes that shift the individual level demand functions x^i_j.

[18] In principle a change in commodity taxation may impact marginal costs through consumer labor supply adjustments. We rule out such general equilibrium effects.

We allow for the possibility that firms exercise market power, meaning they can set price above marginal cost and face positive demand. In this section we can remain agnostic about the precise nature of the imperfect competition. For instance, the products may be differentiated and offered either by a monopolist, competing price-setting oligopolists, or oligopolists that collude to some extent, or the products may be homogeneous but offered by quantity setting oligopolists. In any case, in equilibrium (where all firms choose their strategies to maximize their profit function), the tax-exclusive price for any product j ($\tilde{p}_j \equiv p_j - \tau_j$), can be written $\tilde{p}_j(\mathbf{c}_{\mathcal{M}}, \boldsymbol{\Upsilon}_{\mathcal{M}}; \boldsymbol{\tau}_{\mathcal{M}}) = c_j + \mu_j(\mathbf{c}_{\mathcal{M}}, \boldsymbol{\Upsilon}_{\mathcal{M}}; \boldsymbol{\tau}_{\mathcal{M}})$, where μ_j denotes the equilibrium price-cost margin of product j.[19]

For the composite good, which has aggregate demand $X_O = \int_i x^i_O di$, we assume that its price remains fixed in response to the introduction of an excise tax system in market $\mathcal{M}$, but we allow for the possibility of a non-zero price-cost margin, μ_O. Total profits in the economy are given by $\Pi \equiv \sum_{j \in \mathcal{M} \bigcup 0} \mu_j X_j - FC$, where FC represent fixed costs.

Government

We consider a government that chooses a system of linear commodity taxes in market $\mathcal{M}$. The most general system entails a set of product specific taxes, $\{\tau_j\}_{j \in \mathcal{M}}$. In practice, tax rates tend not to vary across disaggregate products, due to prohibitive implementation issues. We therefore also consider more constrained systems closer to those used in practice. We assume that when introducing the excise tax system the government does not change other parts of the tax system (in particular, the earnings and profit taxes).

Such tax reforms are often motivated by the existence of consumption externalities. We allow for a pecuniary externality associated with consumer i's consumption of products in set $\mathcal{S} \subset \mathcal{M}$ that takes the form $\Phi^i \equiv \Phi^i(\boldsymbol{x}^i_{\mathcal{S}})$, where $\Phi^i(.)$ is a weakly increasing in each of its arguments. We denote the marginal externality of consumer i's consumption of good $k \in \mathcal{S}$ by $\frac{d\Phi^i}{dx^i_j} \equiv \phi^i_j$.

The optimal choice of excise taxes requires the government to balance reducing inefficiencies associated with consumption externalities with the inefficiencies arising from the exercise of market power. It must also account for any spillovers to existing tax bases, and, depending on its preferences for equity, it may factor in distributional

[19] In Section 4 we assume the firms in the UK drinks market compete in a Nash-Bertrand pricing game. Let Ω be a $J \times J$ matrix, where the (j, j') element equals 1 if product j and j' are owned by the same firm and zero otherwise, and let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_J)$. In this case $\boldsymbol{\mu} = -\left[\Omega \circ \left(\frac{\partial \mathbf{X}_{\mathcal{M}}}{\partial \tilde{\mathbf{p}}_{\mathcal{M}}}\right)^T\right]^{-1} \mathbf{X}_{\mathcal{M}}$.

consequences of tax reform. We consider a government with the social welfare function:

$$W = \int_i \omega^i V^i + \lambda(T^i_D + T_\Pi(\delta^i \Pi) - \Phi^i) di, \tag{2.1}$$

where ω^i is the Pareto weight on consumer i, λ is the marginal value of government revenue, and where tax revenue raised from individual i is given by revenue from distortionary taxes:

$$T^i_D = \sum_{j \in \mathcal{M}} \tau_j x^i_j + \mathcal{T}(z^i)$$

and from the tax on their profit holdings, $T_\Pi(\delta^i \Pi)$.

Market power has important implications for tax design. First, the existence of positive margins distorts resource allocations. Second, the existence of positive profits, depending on how they are distributed across individuals, may impact the distributional consequences of taxation. Third, as firms re-optimize their strategies in response to a tax change, the tax-exclusive prices of all products in the market may change in response to a change in the tax rate levied on any one product (with one implication of this being tax changes are not necessary shifted one-for-one to the products subject to the tax change.)

2.2 Optimal policy

The general case

Consider the 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.[20]

The optimal excise tax system satisfies $dW/d\theta = 0$, which implies:

$$\underbrace{\int_i \left(\frac{dT^i_D}{d\theta} - \frac{\partial T^i_D}{\partial \theta} \right) di}_{\text{fiscal externality}} + \underbrace{\sum_{j \in \mathcal{M} \bigcup O} \mu_j \frac{dX_j}{d\theta}}_{\text{market power distortions}} - \underbrace{\int_i \sum_{j \in \mathcal{S}} \phi^i_j \frac{dx^i_j}{d\theta} di}_{\text{externality distortions}} + \underbrace{\int_i (g^i - 1) \frac{dV^i/d\theta}{\alpha^i} di}_{\text{distributional concerns}} = 0. \tag{2.2}$$

Optimal policy balances four forces, three that reflect efficiency considerations and one that reflects equity considerations. The fiscal externality captures the impact of the reform on tax revenue from commodity and labor taxation that is due to agents adjusting their behavior. The market power and externality distortion terms reflect

[20] This is purely an expositional device, which allows us to write the government's optimal condition as a single equation; $\frac{dW}{d\theta} = 0$. In practice, when solving for $K > 1$ optimal tax rates in Section 6, we use $\frac{dW}{d\tau_k} = 0$ for $k = 1, \ldots, K$.

the impact that tax reform has on these non-government distortions. The term reflecting distributional concerns equals the covariance of social marginal welfare weights (g^i) and the money metric impact of the reform on consumer utility. The social marginal welfare weights are defined $g^i \equiv \frac{\omega^i \alpha^i}{\lambda}$ and average to 1 ($\int_i g^i di = 1$). Variation in g^i across individuals reflects the government's preferences for equity: a government with a preference for reducing inequality will assign low weights to the rich and high weights to the poor. The reform impacts individual money metric utility through two channels. First, the tax alters prices, which impact utility from the individuals' chosen consumption bundle. Second, depending on their profit holdings, the tax may lead to a reduction in unearned income.

A single rate sin tax

In the empirical implementation in Section 6, we consider a number of tax systems that vary in the number of tax rates. Here we focus on the case where the government sets a single tax rate, applied to the set of sin goods, $\tau_{\mathcal{S}}$. This serves to clarify the key forces that determine optimal tax policy, facilitates comparison with existing sin tax results derived under perfect competition and is an interesting case as, in practice, governments often implement market specific excise tax systems that set a single rate.

We rearrange condition (2.2) for a single sin tax rate, to express the optimal rate, $\tau^*_{\mathcal{S}}$, as the following *implicit* formula:

$$\tau^*_{\mathcal{S}} = \underbrace{\bar{\phi} + \frac{\text{cov}(\phi^i_j, dx^i_j/d\tau_{\mathcal{S}})}{(1/n(\mathcal{S})) \times d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}}_{\text{externality correction}} - \underbrace{(\bar{\mu}_{\mathcal{S}} - \bar{\mu}_{\mathcal{N}}\Theta_{\mathcal{N}} - \mu_O\Theta_O)}_{\text{market power correction}} + \underbrace{\frac{1}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}\left[\text{cov}\left(g^i, \sum_{j\in\mathcal{M}} x^i_j\rho_j - \delta^i(1-\tau^i_\Pi)\frac{d\Pi}{d\tau_{\mathcal{S}}}\right)\right]}_{\text{distributional concerns}} - \underbrace{\frac{d(\int_i \mathcal{T}(z^i)di)/d\tau_{\mathcal{S}}}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}}_{\text{tax base erosion}}. \quad (2.3)$$

$\bar{\phi} \equiv \int_i \frac{1}{n(\mathcal{S})} \sum_{j\in\mathcal{S}} \phi^i_j di$ denotes the average marginal consumption externality in the population and $d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}} = \sum_{j\in\mathcal{S}} dX_j/d\tau_{\mathcal{S}}$ is the impact of a marginal tax change on total consumption of the set of sin products. $\bar{\mu}_{\mathcal{X}} \equiv \sum_{j\in\mathcal{X}} \mu_k \frac{dX_j/d\tau_{\mathcal{S}}}{\sum_{j'\in\mathcal{X}} dX_{j'}/d\tau_{\mathcal{S}}}$ is the weighted average margin for products in set $\mathcal{X} = \{\mathcal{S}, \mathcal{N}, O\}$, where the weights are each product's contribution to the marginal impact of the sin tax on equilibrium consumption of all products in that set. $\Theta_{\mathcal{X}} \equiv \frac{d\mathbb{X}_{\mathcal{X}}/d\tau_{\mathcal{S}}}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}$ is the fraction of reduced consumption of products in set $\mathcal{S}$ diverted to those in $\mathcal{X} = \{\mathcal{N}, O\}$ due to a marginal tax rise, and $\rho_j \equiv \frac{dp_j}{d\tau_{\mathcal{S}}}$ is the impact of a marginal tax change on the equilibrium consumer price of product j.

Equation (2.3) expresses the optimal tax rate as a function of wedges from non-government (externality and market power) distortions, tax derivatives (for quantities, prices and profits) and government distributional preferences. It has four components, which we describe in turn.

Externality correcting component. This equals the average marginal externality across consumers and sin products, plus an adjustment that reflects the covariance between the consumer-product specific marginal externality and the sensitivity of the individual's consumption of the product to a change in the sin tax rate. This covariance captures how effective the tax is at reducing the most socially costly consumption. The more a tax rise reduces consumption by consumers and/or of products associated with high marginal externalities, the better targeted it will be at the most externality generating consumption and the higher will be the optimal tax rate.[21] This mirrors the logic in the optimal externality correcting tax rate with heterogeneous externalities derived in Diamond (1973). However, an important difference is that the sensitivity of consumption to changes in the sin tax rate also depends on the equilibrium pricing response of firms; i.e., $\frac{dx_j^i}{d\tau_{\mathcal{S}}} = \sum_{j' \in \mathcal{M}} \frac{\partial x_j^i}{\partial p_{j'}} \frac{dp_{j'}}{d\tau_{\mathcal{S}}}$.

Market power correcting component. All else equal, higher equilibrium margins on the products in set $\mathcal{S}$ act to lower the optimal tax rate. This reflects the classic argument made by Buchanan (1969), who pointed out the appropriate tax rate on an externality generating monopolist lies below the full Pigouvian rate. However, if taxing products in set $\mathcal{S}$ induces an increase in consumption of other products also supplied non-competitively, then distortions arising from the exercise of market power on these alternatives, all else equal, act to raise the optimal tax rate. The strength of this effect depends on the weighted average margin on these non-taxed alternatives, and the extent to which a marginal tax rise diverts equilibrium consumption towards them. For products in the set $\mathcal{N}$, these two forces are captured by the terms $\bar{\mu}_{\mathcal{N}}$ and $\Theta_{\mathcal{N}}$ (analogous expressions capture the influence of any market power distortions from the numeraire good).

Distributional concerns. These enter equation (2.3) through the covariance of consumers' social marginal welfare weights with the reduction in their utility

[21] Note, $d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}$ will generally be negative; an increase in the sin tax rate will lower equilibrium consumption of those products. If $\text{cov}(\phi_j^i, dx_j^i/d\tau_{\mathcal{S}})$ is negative, so a tax rise tends to achieve relatively large consumption reductions among products/consumers with large marginal externalities, the externality correction will exceed $\bar{\phi}$. Hence, all else equal, the more the total reduction in sin good consumption is concentrated among the most socially costly consumption, the higher will the externality correction component of the optimal tax rate.

resulting from a marginal increase in the tax rate, scaled by the marginal effect of the tax on total consumption of the sin products. The more a marginal tax rise results in relatively large utility falls among those with high social marginal welfare weights (low income consumers when the planner has a preference for equity), the lower will be the optimal tax rate. The strength of the distributional channel depends on how responsive equilibrium consumption of the set of sin products is to a marginal sin tax rise: the more sensitive is equilibrium consumption, the smaller the impact of equity considerations on the optimal tax rate.

To highlight how market power distortions interact with distributional concerns, it is informative to consider this term under perfect competition: $\frac{1}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}$ $\text{Cov}\left(g^i, \sum_{j\in\mathcal{S}} x_j^i\right)$. The reduction in a consumer's utility due to a marginal tax rise equals their total consumption on the taxed sin goods. All else equal, the more that those with high social marginal welfare weights consume a lot of the set of sin goods, the lower is the optimal tax rate on them.[22] Imperfect competition has two consequences for the utility impact of a marginal tax rate increase, and hence on the distributional component of equation (2.3). First, consumption of the sin products is replaced by consumption of all products in market $\mathcal{M}$, weighted by the marginal impact of tax on each product's equilibrium price. If firms hold their tax-exclusive prices fixed in response to tax changes, then this term collapses back to total consumption of the sin goods. Second, the impact of the tax on the size of individuals' net-of-tax profit holdings also matters. If the tax leads to a reduction in profits, and profit holdings are disproportionately held by those with low social marginal welfare weights (the wealthy), this will act to make the tax more progressive and will increase the optimal rate.

Tax base erosion. This terms arises because we assume that the government holds fixed the earning tax schedule. All else equal, the more that a marginal increase in the sin tax leads to a reduction in labor tax revenue, the lower will be the optimal rate. Whether this term (and hence the loss from not re-optimizing the earning tax alongside introducing the sin tax) is large or small is context dependent. To highlight what drives this term we assume that income effects on labor supply are negligible (see Saez et al. (2012) for empirical support of this), which allows us

[22] Allcott et al. (2019a) show that if the government also optimizes the earnings tax schedule, what matters is the cross sectional correlation in social marginal welfare weights and consumption of the taxed goods *net of income effects*.

to re-write $d(\int_i \mathcal{T}(z^i)di)/d\tau_{\mathcal{S}}$ as

$$\frac{d(\int_i \mathcal{T}(z^i)di)}{d\tau_{\mathcal{S}}} = \int_i \frac{\tau_z^i}{1-\tau_z^i}\zeta_z^i \sum_{j\in\mathcal{M}} \xi_j^i x_j^i \rho_j di,$$

where ζ_z^i is the individual (compensated) elasticity of taxable earnings and ξ_j^i is the individual elasticity of demand for product j with respect to earnings (see Appendix A). This expression highlights that a key determinant of the tax base erosion component of equation (2.3) is the strength of income effects for products in market $\mathcal{M}$. In Appendix B we provide empirical evidence that these are very small in the context of the UK drinks market.

Note, although we have written the "market power correction" as additively separable from the other components of the tax formula (equation (2.3)), the additive separability is illusory. Market power influences each component of the formula, because the strategic responses of firms are key in driving the quantity, price and profit tax derivatives.

2.3 Discussion of empirical implementation

A common approach to empirical tax analysis is to write the expression of interest in terms of a set of externally valid elasticities, or sufficient statistics. An advantage of this approach is that the elasticities can be estimated using quasi-experimental variation, with transparent identification arguments. However, an important restriction of the application of sufficient statistics to optimal tax formulae (as opposed to marginal tax reforms) is that they require implicitly assuming a structural (iso-elastic preference) model – see Kleven (2020).

There are two challenges with implementing a sufficient statistics approach in our context (with market power). First, the tax derivatives do not straightforwardly map into price elasticities. For a number of the derivatives this could potentially be overcome by making simplifying assumptions about tax pass-through.[23] Alternatively, one could use data covering the introduction of a new (or change in an existing) tax to directly estimate the tax derivatives (policy elasticities in the language of Hendren (2016)). The second challenge is that the optimal tax formula depends on equilibrium product-level price-cost margins. Marginal costs, and hence

[23]For instance, in general, the impact of a marginal tax rise on consumption of the set of sin products takes the form $\frac{d\mathbb{X}_{\mathcal{S}}}{d\tau_{\mathcal{S}}} = \sum_{j\in\mathcal{S}}\sum_{j'\in\mathcal{M}} \frac{\partial X_j}{\partial p_{j'}}\frac{dp_{j'}}{d\tau_{\mathcal{S}}}$. However, under the assumption of fixed tax pass-through (denoted ρ) across products in $\mathcal{S}$ and fixed tax-exclusive prices for other goods, this collapses to $\rho \times \triangledown\mathbb{X}_{\mathcal{S}}$ where we use $\triangledown\mathbb{X}_{\mathcal{S}}$ to denote the marginal impact of consumption goods in $\mathcal{S}$ with respect to a marginal price rise for all these products.

margins, are typically not straightforwardly observable in economic data. However, they can be inferred based on a profit maximizing model of the firms operating in the market, coupled with estimates of the own- and cross-price elasticities of all the products in the market.[24]

Our approach is therefore to specify and estimate an equilibrium model of the market of interest (the UK drinks market). This enables us to simulate the impact of an arbitrary tax policy on equilibrium margins, consumption and profits. To validate the model, we compare, where possible, its predictions to existing estimates of relevant elasticities, as well as to evidence from the introduction of the UK's sugary drinks tax. We embed the equilibrium model into the tax problem and consider optimal policy under different tax systems of varying degrees of flexibility (e.g., single rate vs. multi-rate system) and under different government preferences (e.g., efficiency maximizing, inequality averse).

3 The drinks market

Sugar sweetened beverage taxation is a natural setting in which to study how taxes on externality generating products interact with market power. In many jurisdictions, taxes on drinks have explicitly been motivated as a tool for improving public health, in part due to substantial external costs associated with their consumption. We discuss the nature and measurement of these external costs in Section 6. It is also the case that the market is concentrated (the two largest firms together have over 50% market share) and comprises a set of highly recognizable branded products. It is likely therefore that firms exercise considerable market power.

3.1 Data

We model behavior in the UK market for non-alcoholic drinks. This market includes carbonated drinks (often referred to as sodas), fruit concentrates, and sports and energy drinks. We refer collectively to these as soft drinks. Typically they come in sugar-sweetened and artificially sweetened (i.e., diet) varieties. Sugar sweetened beverage taxes typically apply to sugar-sweetened varieties of soft drinks. The market also includes pure fruit juices and flavored milk. We use micro data on the drinks purchases of a sample of consumers living in Great Britain collected by the market research firm Kantar. The data contain information on household level

[24] Note we do make use of limited data on accounting margins. While these do not neatly conform to the economic concept of a price-cost margin, they are a useful cross-check on margins inferred from a model.

purchases for "at-home" consumption (Kantar Worldpanel), as well as purchases made by individuals for "on-the-go" consumption (Kantar On-The-Go Survey). Together "at-home" and "on-the-go" consumption account for over 90% of drinks consumption by volume.[25]

Households in the at-home sample record, by means of a barcode scanner, all grocery purchases made and brought into the home. The data are broadly representative of British households (in Appendix B we compare them with the nationally representative Living Cost and Food Survey) and cover 2008 to 2012. Individuals in the on-the-go data record all purchases they make from shops and vending machines for out-of-home consumption using a cell phone app. The data cover 2010 to 2012 and comprise individuals (aged 13 and upwards) randomly drawn from the Worldpanel households. In both datasets, we observe households/individuals over many months. The data contain detailed information – including brand, flavor, size and nutrient composition – on the UPCs (barcodes) purchased, the store in which the purchase took place, and transaction level prices.

3.2 Consumers and purchasing patterns

We use the term consumer to refer to households in the at-home segment, and individuals in the on-the-go segment. Our at-home sample of consumers comprises 30,405 households and our on-the-go sample comprises 2,862 individuals.[26]

Figure 3.1 highlights variation in soft drinks purchases across two key dimensions. Panel (a) shows that consumers that get a high fraction of their *total* dietary calories from added sugar purchase significantly more sugar-sweetened beverages than other consumers. Policymakers have typically focused on changing the behavior of consumers with dietary sugar above a particular threshold, due to elevated health risks (e.g., the World Health Organization (2015) advice focuses primarily on those with dietary sugar above 10%). The more that a sugar-sweetened beverage tax is able to achieve large consumption reductions among those consumers that create relatively high marginal externalities through their sugar-sweetened beverage intake, the more effective it will be at reducing externality distortions.

Panel (b) shows that there is a negative cross-sectional correlation between sugar-sweetened beverage consumption and equivalized household income – richer

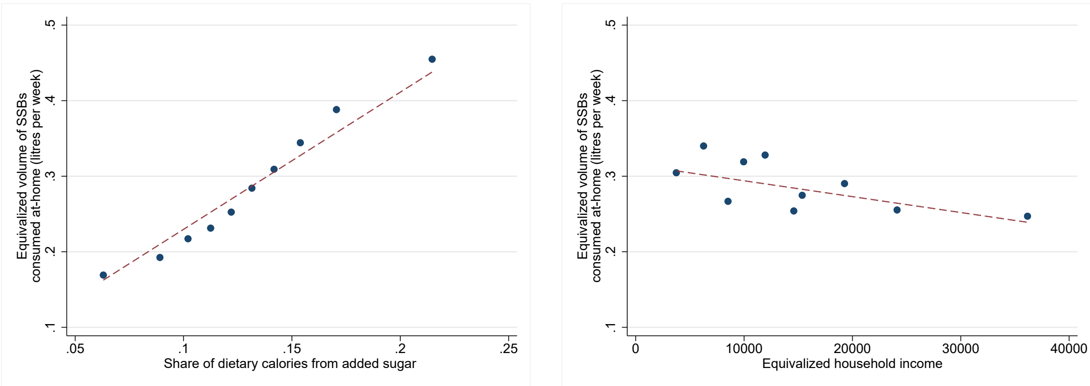
[25]The remainder occurs in restaurants and bars, which are not covered by our data. Numbers are computed using the Living Cost and Food Survey.

[26]We omit a small number of consumers that record irregularly. Specifically, in the at-home segment we focus on households that record purchases 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 5 purchases each year. In each segment, this conditioning drops less than 3% of transactions.

households consume less sugar-sweetened beverages, and therefore have consumption baskets less exposed to a sugar-sweetened beverage tax, than lower income consumers. The extent to which this is driven by preference heterogeneity (correlated with income) versus causal income effects will impact optimal policy. In Appendix B we show that after removing consumer fixed effects (and hence relying on within household income transitions to estimate any income effects), the consumption gradient in equivalized income flattens completely. In our demand model we control flexibly for equivalized income, and in our optimal tax analysis, we treat variation in drinks demand across the income distribution as preference heterogeneity.

Figure 3.1: *Variation in volume of sugar-sweetened beverages consumed at-home*

(a) Dietary sugar intake (b) Household income



Notes: The left hand panel shows mean volume of sugar-sweetened beverages purchased per person per week and consumed at home 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 at home by deciles of equivalized (using the OECD-equivalence scale) household income. Analogous figures for the sugar from soft drinks consumed on-the-go are shown in Appendix B.

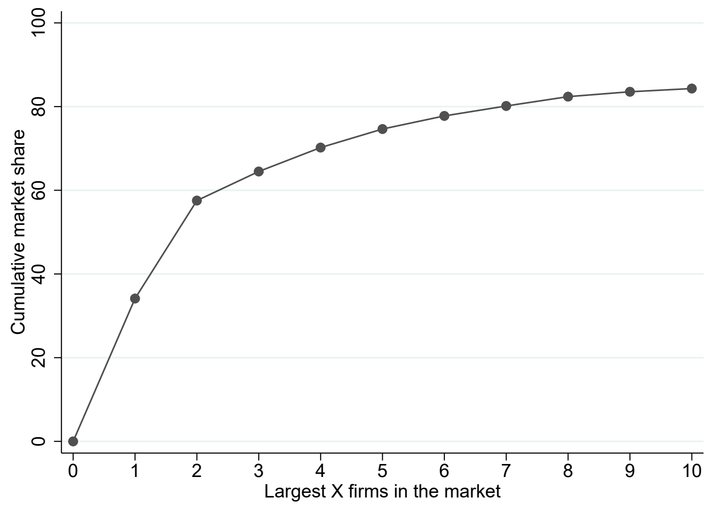
3.3 Firm and brands

The drinks market is highly concentrated. In Figure 3.2 we show the cumulative market share of the ten largest UK producers. The two largest firms (Coca Cola Enterprises and Britvic have a combined market share of almost 60%).[27]

[27] Drinks producers are known in the industry as bottlers. They buy concentrate from upstream firms (e.g., Coca Cola Enterprises obtains concentrate from The Coca Cola Company) and use this as an input to produce soft drinks products (see Luco and Marshall (2020)). For modeling firm behaviour in the UK drinks market, it is the bottlers – all of which are national – who are the relevant agents: as well as producing the products, they are responsible for negotiating product shelf prices and placement with retailers, and for promotional activity (see Competition Commission (2013)).

Firms each typically own several separate brands. For instance, Coca Cola Enterprises most popular brand is Coca Cola, but it also owns 9 other brands with market share of at least 1%. Each soft drinks brand is typically available in sugar-sweetened ("regular") and artificially sweetened ("diet" and/or "zero") variants. Each brand-variant is available in multiple pack-sizes. In our equilibrium model of the market we focus on the set of main brands in the market, which together comprise over 75% of total spending on non-alcoholic drinks.[28] Table B.3 in Appendix B lists brand-variants and the number of sizes they are available in.

Figure 3.2: *Market share of the largest firms in the drinks market*



Notes: The line shows the cumulative market share for the X largest firms in the market, where X is shown on the horizontal axes. Market shares are shown for 2012.

The combination of large multi-product firms and differentiated, strongly branded, products means that it is likely that the firms in this market exercise considerable market power. Our empirical estimates shed light on the extent to which this is true.

3.4 Drinks firm-retailer relations

Drinks firms do not sell directly to consumers, rather retailers act as intermediaries between drinks firms and consumers. The majority of expenditure is undertaken in national grocery chains (see Table B.4 in Appendix B). In the UK the main gro-

[28]This include all soft drinks brands with more than 1% market share in either the at-home or on-the-go segment, as well as the main fruit juice and flavored milk brands. For some brands, there are only a very small number of transactions in one of the two segments of the market; we therefore omit these brands from the choice sets in that segment.

cery chains set prices nationally (see Competition Commission (2000)).[29] Retailers typically offer all brand-varieties, though the pack sizes on offer can vary.

We do not directly observe the contracting relationship between the drinks firms and retailers. However, a 2013 report into the soft drinks market by the UK competition authority provides evidence on the nature of these relations (see Competition Commission (2013)). They cite evidence that annual bilateral "Joint Business Plans" are agreed between a drinks firm and retailer setting out wholesale prices, payments related to product visibility, recommended retail prices, and agreements on the number, type and timings of promotions. This evidence of non-linear contracting suggests drinks firms and retailers avoid double marginalization. We therefore treat drinks firms as (effectively) setting final consumer prices, an outcome consistent with optimal non-linear contracting – see Villas-Boas (2007) and Bonnet and Dubois (2010). We also exploit the fact the promotions are agreed on in advance (and are not coordinated across retailers) as a useful identifying source of price variation (see Section 4.3).

4 Equilibrium model of the drinks market

We estimate a model of consumer demand in the drinks market using a discrete choice framework in which consumer preferences are defined over product characteristics (Gorman (1980), Lancaster (1971), Berry et al. (1995)). This enables us to model demand for the many differentiated products in the market, while incorporating rich preference heterogeneity, including by total dietary sugar and income. We identify product level marginal costs (and hence equilibrium price-cost margins) by combining the demand estimates with the equilibrium conditions from an oligopoly pricing game (Berry (1994), Nevo (2001)). The estimates of the primitives of demand and supply enable us to simulate the impact of tax policy on equilibrium quantities and prices, and hence consumer utilities and profits.

4.1 Consumer demand

We model which, if any, drink product a consumer (indexed i) chooses on a "choice occasion", where choice occasion refers to a week in which a household purchases

[29]This means that if we observe a transaction price for a particular product in a store belonging to one of the main retailers we know the price that consumers shopping in other stores belonging to that retailer at the same time faced for that product. Using the large number of transactions in our data we can construct the price vector consumers faced in each retailer in each week. For the smaller retailers we construct a mean transaction price for a product as a measure of the price faced by consumers.

groceries in the at-home segment, or a day on which an individual buys a cold beverage (including bottled water) in the on-the-go segment.[30] We treat the decisions that households make in the at-home segment and individuals make in the on-the-go segment separately, allowing for all preferences to vary freely with each type of choice situation. In Appendix C we provide evidence that recent purchases of drinks by a household in the at-home segment do not influence either the propensity to buy or quantity purchased by household members in the on-the-go segment.[31] Choice in the on-the-go segment is between single portion size of products (e.g., 330ml cans and 500ml bottles); choice in the at-home segment is between multi-portion sizes. For notational parsimony we suppress a market segment index.

We index the drinks products by $j = \{1, \dots, J\}$. Products vary by brand, indexed by $b = \{1, \dots, B\}$, whether or not they contain sugar (for instance, the brand Coke is available in Regular, Diet and Zero variants), and their size, indexed by $s = \{1, \dots, S\}$. Brand-variants can be purchased in different sizes for two reasons: (i) the availability of different pack sizes (or UPCs), and (ii) the purchase of multiple packs. For each brand-variant we define sizes as the set of available pack sizes and the most common multiple pack purchases of UPCs.[32] Our product definition embeds only minimal aggregation across very similar UPCs; for instance, across different flavors that have the same sugar contents and prices.[33] The consumer chooses between the available drinks products and choosing not to buy a drink, which we denote by $j = 0$. On around 42% of at-home and 60% of on-the-go choice occasions, a household purchases a drink (i.e., $j > 0$).[34] As there is some variation in available pack-sizes by retailer (indexed r), the set of available drinks products is retailer specific (and denoted Ω_r).

[30]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.

[31]This is consistent with the findings in Dubois et al. (2020), who, following Browning and Meghir (1991), formally test for non-separabilities between the two segments and find no evidence of demand dependence.

[32]Specifically, 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.

[33]For instance, "Diet Coke 12x330ml" and "Diet Coke Caffeine Free 12x330ml" – which are priced the same and have zero sugar – both belong to the product "Coke: Diet: 12x330ml".

[34]Consumers are sometimes observed purchasing multiple (typically) two brand-variants on a single choice occasion. On 40% (10%) of occasions in which a consumer chooses a drink in the at-home (on-the-go) segment, multiple are chosen. In this case, we randomly sample one, assuming that, conditional on consumer specific preferences, these purchases are independent, e.g., because they are bought for different household members.

Consumer i in period t, with total period budget y_{it}, solves the utility maximization problem:

$$V(y_{it}, \mathbf{p}_{rt}, \mathbf{x}_t, \epsilon_{it}; \boldsymbol{\theta}_i) = \max_{j \in \{\Omega_r \cup 0\}} \nu(y_{it} - p_{jrt}, \mathbf{x}_{jt}; \boldsymbol{\theta}_i) + \epsilon_{ijt}. \tag{4.1}$$

where $\mathbf{p}_{rt} = (\mathbf{p}_{1rt}, \ldots, \mathbf{p}_{Jrt})$ is the price vector faced by the consumer, $\mathbf{x}_{jt}$ are other characteristics of product j, and $\mathbf{x}_t = (\mathbf{x}_{1t}, \ldots, \mathbf{x}_{Jt})$ (note $p_0 = 0$ and $\mathbf{x}_{0t} = 0$); $\boldsymbol{\theta}_i$ is a vector of consumer level preference parameters; and $\epsilon_{it} = (\epsilon_{i0t}, \epsilon_{i1t}, \ldots, \epsilon_{iJt})$ is a vector of idiosyncratic shocks.

The function $\nu(.)$ captures the payoff the consumer gets from selecting option j. Its first argument, $y_{it} - p_{jrt}$, is spending on the numeraire good, i.e., spending outside the drinks market. We assume that preferences are quasi-linear, so $y_{it} - p_{jrt}$ enters $\nu(.)$ linearly. This means that y_{it} differences out when the consumer compares different options; we therefore suppress the dependency of $\nu(.)$ on y_{it}. An implication of quasi-linearity is that a change in the price of any drinks product does not induce an income effect. Given the small share of total consumer expenditure allocated to drinks products, this is a mild assumption.[35] However, we do allow equivalized household income to shift consumer preferences, θ_i. This enables our model to capture how demands vary across consumers with different incomes.

We assume that ϵ_{ijt} is distributed i.i.d. type I extreme value. Under this assumption the probability that consumer i selects product j in period t, conditional on prices, product characteristics and preferences, is given by:

$$\sigma_j(\mathbf{p}_{rt}, \mathbf{x}_t; \boldsymbol{\theta}_i) = \frac{\exp(\nu(p_{jrt}, \mathbf{x}_{jt}; \boldsymbol{\theta}_i))}{1 + \sum_{j' \in \Omega_r} \exp(\nu(p_{j'rt}, \mathbf{x}_{j't}; \boldsymbol{\theta}_i))}, \tag{4.2}$$

and the consumer's expected utility is given by:

$$v(\mathbf{p}_{rt}, \mathbf{x}_t; \boldsymbol{\theta}_i) = \ln \sum_{j \in \Omega_r} \exp\{\nu(p_{jrt}, \mathbf{x}_{jt}; \boldsymbol{\theta}_i)\} + C, \tag{4.3}$$

where C is a constant of integration.

Specification details

We allow for rich preference heterogeneity, with both observed and unobserved consumer characteristics. This is important for two reasons. First, it is well established that the inclusion of rich, and, in particular, unobserved preference heterogeneity

[35] In general, the own price effect on demand for good j follows the Slutsky equation $\epsilon_{ij} = \epsilon^h_{ij} + \frac{p_j q_{ij}}{y} e_{ij}$, where ϵ_{ij} and ϵ^h_{ij} are the Marshallian and Hicksian own-price elasticities of demand, and e_{ij} is the income elasticity. For a small budget share good $\frac{p_j q_{ij}}{y} \approx 0$, meaning $\epsilon_{ij} \approx \epsilon^h_{ij}$ and preferences are approximately quasi-linear.

is crucial in enabling models of this type (mixed logit choice models) to recover realistic patterns of consumer substitution (i.e., own- and cross-price elasticities) across products (see Berry et al. (1995), Nevo (2001)).[36] Second, in our application it is important that we capture variation in preferences across different consumers that is relevant for tax policy. This includes variation across consumers whose consumption is likely to create different externalities at the margin, and across consumers to which a government may assign different social marginal welfare weights. We therefore partition consumers into demographic groups (indexed d) based on whether their total share of dietary calories (measured in the preceding year) is below 10%, between 10 and 15% and in excess of 15% (as well as on the basis of whether the household contains children), and allow all preference parameters to vary by these groups; see Table B.5 in Appendix B for details. We also allow preferences over key product attributes to vary with equivalized household income (which we denote $\tilde{y}_i$).

We specify that the payoff function $\nu(.)$ for consumer i belonging to consumer group $d(i)$ and for product j belonging to brand $b(j)$ and of size $s(j)$ takes the form:

$$\nu(.) = -\alpha_{i0} p_{jrt} + \sum_{k>0}^{K} \alpha_{ik} x_{jk} + \zeta_{d(i)b(j)s(j)rt},$$

where, for product attribute, $k = 0, \ldots, K$:

$$\alpha_{ik} = \bar{\alpha}_{d(i)k} + \alpha^{\tilde{y}}_{d(i)k} \tilde{y}_i + \sigma^{\alpha}_{d(i)k} \eta_{ik},$$

$\zeta_{d(i)b(j)s(j)rt}$ denotes an unobserved brand-size attribute (which may vary by demographic group, retailer and time) and η_{ik} is a standard normal random variable.[37]

$\bar{\alpha}_{d(i)k}$ denotes the baseline preference for product attribute k among demographic group d, $\alpha^{\tilde{y}}_{d(i)k}$ captures how preferences for the attribute vary across consumers with different equivalized household incomes, and $\sigma^{\alpha}_{d(i)k}$ captures the dispersion in unobserved preferences for the attribute. Product attributes include (in addition to price) sugar content, drink types (cola, lemonade, pure fruit juice etc.) size and advertising.[38] We allow preferences over price, sugar, branded soft drinks,

[36]Models of this sort achieve dimension reduction by mapping products into product attribute space. With no preference heterogeneity the model implies unrealistic substitution patterns that satisfy the independence of irrelevant alternatives (IIA) property. However, the inclusion of preference heterogeneity breaks the IIA property and allows for much richer substitution patterns. Indeed, in principle, the mixed logit model (if specified flexibly enough) can approximate to an arbitrary degree any underlying random utility model (McFadden and Train (2000)).

[37]For price we assume the coefficient is log-normally distributed.

[38]We measure monthly advertising expenditure across TV, radio and online in the AC Nielsen Advertising Digest. We compute product specific stocks based on a monthly depreciation rate of

non-branded drinks and pure fruit juice to vary with equivalized income, and for unobserved heterogeneity in preferences for drinks (relative to the outside option), price, size, cola, lemonade, non-branded drinks and fruit juice. As we allow both the baseline and dispersion parameters to vary across demographic groups d, the overall random coefficient distribution is a flexible mixture of normal distributions.

We decompose the unobserved product attribute, $\zeta_{d(i)b(j)s(j)rt}$, into a set of detailed fixed effects:

$$\zeta_{d(i)b(j)s(j)rt} = \xi^{(1)}_{d(i)b(j)s(j)} + \xi^{(2)}_{d(i)b(j)t} + \xi^{(3)}_{d(i)s(j)t} + \xi^{(4)}_{d(i)s(j)r} + \xi^{(5)}_{d(i)b(j)r}$$

All the fixed effects are demographic group specific. They include brand-size, brand-time, size-time, brand-retailer and size-retailer effects. These control for shocks to demand that may be correlated with price setting. They leave residual temporal brand-size price variation that is differential across retailer. In Section 4.3 we discuss why this is useful for identifying the price responsiveness of demand.

4.2 Supply model

We model drinks firms as setting prices in a simultaneous move Nash-Bertrand game.[39] We do not explicitly model drinks firms-retailer relationships, but, based on the evidence of non-linear contracting in these relations, we assume manufacturers set consumer prices, which is consistent with efficient contracting. Let $\boldsymbol{p}_m = (p_{1m}, \ldots, p_{Jm})$ denote the prices that drinks firms set in market (year) m.[40]

Market demand for product j is given by:

$$q_{jm}(\mathbf{p}_m) = \int_i \sigma_j(\mathbf{p}_m, \mathbf{x}_m; \boldsymbol{\theta}_i) dF(\boldsymbol{\theta}) M_m,$$

0.8. This is similar to the rate used in Dubois et al. (2018) on similar data in the potato chips market.

[39]Around one-fifth of the market consists of "store brands". These are no-frills, low priced alternatives to the branded products. We treat these as a competitive fringe, with (tax-exclusive) prices that remain fixed.

[40]In practice, for a given product-year a drinks firm and retailer agree on a base price $\bar{p}$ and a sale price p_S, with the former applying ρ proportion of weeks (see Section 3.4). Instead of modeling choice over $(\bar{p}, p_S, \rho)$, we model choice over $p = (1-\rho)\bar{p} + \rho p_S$. This average price exhibits little variation across retailers. Cross-retailer variation in the price of a given product at a point in time is driven by non-synchronization of sales (see next section). Hence, we specify the relationship between prices in the supply game, p_{jm}, and those faced by consumers in retailer r week $t \in m$ as $p_{jrt} = p_{jm} + \tilde{\mu}_{jrt}$, where $\mathbb{E}[\tilde{\mu}_{jrt}|(j,m)] = 0$.

where M_m denotes the potential size of the market.[41] We denote the marginal cost of product j in market m as c_{jm}.[42]

We index the drinks firms by $f = (1, \ldots, F)$ and denote the set of products owned by firm f by $\mathcal{J}_f$. Firm f's total variable profits in market m are

$$\Pi_{fm}(\boldsymbol{p}_m) = \sum_{j \in \mathcal{J}_f} (p_{jm} - c_{jm}) q_{jm}(\boldsymbol{p}_m). \tag{4.4}$$

Under Nash-Bertrand competition, the equilibrium prices satisfy the set of first order conditions: $\forall f$ and $\forall j \in \mathcal{J}_f$,

$$q_{jm}(\boldsymbol{p}_m) + \sum_{j' \in \mathcal{J}_f} (p_{j'm} - c_{j'm}) \frac{\partial q_{j'm}(\boldsymbol{p}_m)}{\partial p_{jm}} = 0. \tag{4.5}$$

From this system of equations we can solve for the implied marginal cost, c_{jm}, and hence the equilibrium price-cost margin, $\mu_{jm} = p_{jm} - c_{jm}$, for each product in each market. For any set of product taxes (most generally $(\tau_1, \ldots, \tau_J)$) we can use the system of equations (4.5), replacing c_{jm} with $c_{jm} + \tau_j$, to solve for counterfactual equilibrium prices p'_{jm} and margins $\mu'_{jm} = p'_{jm} - \tau_j - c_{jm}$.

4.3 Identification

We begin by discussing identification of the baseline price parameters ($\bar{\alpha}_{d(i)0}$), before discussing the other preference parameters.

We assume that the detailed fixed effects that we include in the model (in addition to the advertising control) absorb taste variation that is relevant for price-setting. The brand-size effects, $\xi^{(1)}_{d(i)b(j)s(j)}$, absorb the influence of unobserved product attributes that are not captured by the included observable product attributes. Variation in taste for brands or particular sizes over time, due, for instance, to seasonal patterns, are captured by the brand-time, $\xi^{(2)}_{d(i)b(j)t}$, and size-time, $\xi^{(3)}_{d(i)s(j)t}$, effects. In addition, we control separately for product level advertising, which will capture the effect on demand of the (overwhelmingly national) advertising in the

[41] M_m is the potential number of non-alcoholic drinks transactions in market m, it differs from the true market size due to inclusion in the demand model of the option to purchase no drinks.

[42] Note, in Section 2 we express quantity in terms of unit volume, say liters, and prices and marginal costs per liter. Here we express quantity as number of transactions and price and marginal cost per transaction. The difference is one of convenience rather than substance, multiplying q_{jm} by the size of the product and dividing p_{jm} and c_{jm} by the size of the product transforms the variables into their analogs in Section 2 without changing the nature of the firms' problem.

UK drinks market.[43] Finally, tastes for brands or sizes may vary across retailers, which is captured by the brand-retailer, $\xi^{(4)}_{d(i)b(j)r}$, and size-retailer, $\xi^{(5)}_{d(i)s(j)r}$, effects.

The price variation that we exploit is *product level time series variation that is differential across retailers*. In particular, while differences in the average price (over time) that different retailers set for a given product are small, the degree of co-movement in prices for the same product in different retail chains over time is low. The differential price movements are generally driven by time-limited price reductions. These price reduction strategies, which are agreed in advance at bilateral meetings between drinks firms and retailers (see Section 3.4), creates, from consumers' perspective, randomness in the prices they face.

In Figure 4.1 we depict graphically some of this 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. The rationale provided by firms for these long running quantity discounts are economies of scale associated with running their plants at or near full capacity (Competition Commission (2013)). 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.

A threat to our identification strategy is that consumers respond to promotions by intertemporally switching their purchases i.e., stocking up when the price is low and consuming from this stock when the price is high. As shown by Hendel and Nevo (2006a) this will lead to over-estimates of own price elasticities (and typically to under-estimates of cross price elasticities). In Appendix C we check for evidence of stockpiling behavior in the UK market, by running the within-consumer tests suggested in Hendel and Nevo (2006b). We find very little evidence of stockpiling – for instance, when a UK household purchases a drink on sale there is no meaningful change in the timing of purchases. In contrast, in the US, buying a soft drinks on sale is associated with an average reduction in the time from previous purchase of 3 days, and an increase to the next purchase of 2.5 days (Hendel and Nevo (2006b)). Instead, we show that sales in the UK lead to *intra*-temporal substitution, with

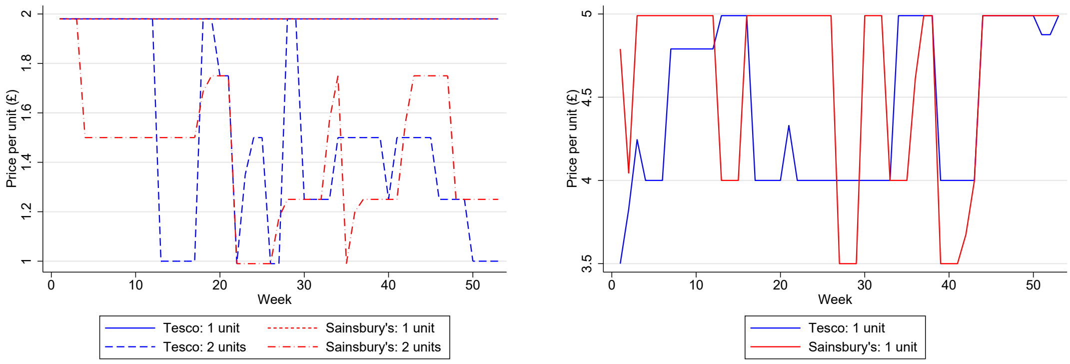
[43] Note targeted price discounts through use of coupons – common in the US (see Nevo and Wolfram (2002)) – is not a feature of the UK market.

sales leading to a 12.5% increase in the probability of switching from bottles to cans (or vice versa), and a 3.3% and 4.5% in the probability of switching between brands and sizes.[44] We also note that UK consumers purchase drinks every two weeks on average, which is around twice as often as their counterparts in the US, and promotions are much more long running in nature, compared with week to week price fluctuations in the US (see an archetypal example in Figure 1 of Hendel and Nevo (2013)).[45]

Figure 4.1: *Examples of price variation for Coke options*

(a) 2l bottle

(b) 12×330ml cans



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.

For the non-price product attributes the baseline preference is identified as long as the attribute exhibits within brand-size variation (otherwise it is absorbed by the brand-size fixed effects). For instance, we estimate the consumers' baseline preference for products that contain sugar. This relies on the brand-size effects being common across diet and sugar variants. A given demographic group will have a stronger estimated preference for sugar if, conditional on all the controls in the model (including the brand-size effects), this group purchases products containing sugar more often than other groups.

We allow all preference parameters to vary by the demographic groups. Incorporating this flexible preference heterogeneity across observed dimensions is possible because we observe many consumers within each group making decisions, while

[44] Buying on sale is associated with a less than 1% change in the probability of switching retailer.

[45] An interesting avenue for future work is to explore why there is such differences between the two countries. Part of the answer is likely to lie with higher transport and storage costs in the UK: the average size of UK homes is around half of those in US – the mean floor space of UK homes in 2008 was 85m^2, while in 2009 in the US it was 152m^2 (UK Government (2018)) – and vehicle ownership rates are 25% lower – in 2014 the US had 816 vehicles per capita (U.S. Department of Energy (2019)), in 2017 the UK had 616 (ACEA (2019)).

facing different price vectors. We also allow preference heterogeneity in price sensitivity and tastes for sugar and certain sets of product (e.g., branded soft drinks, fruit juices) by equivalized household income, using the fact that we observe consumers at different points of the income distribution making choices. The main source of identifying variation for the parameters governing unobserved preference heterogeneity comes from the panel structure of our micro data. For instance, if the within-consumer covariance in the price of chosen options across choice sets is high (relative to the covariance across all choice occasions in the relevant demographic group), this indicates significant consumer specific preference heterogeneity in price sensitivity, and acts to increase the spread parameter, $\sigma^{\alpha}_{d(i)0}$, on price preferences.[46]

5 Demand and supply estimates

5.1 Consumer substitution patterns

We estimate the demand model outlined in Section 4.1 using simulated maximum likelihood,[47] and report the coefficient estimates in Appendix D. The estimated coefficients exhibit some intuitive patterns: those with relatively high overall added sugar in their diets have stronger preferences for sugary drinks products, and those with lower incomes are more sensitive to price, have stronger preferences for soft drinks and weaker preferences for pure fruit juice. The variance parameters of the random coefficients are significant both statistically and in size, indicating an important role for unobserved preference heterogeneity.

Product level elasticities

The estimated preference parameters determine our demand model predictions of how consumers switch across products as prices change. The model generates a large matrix of product level own- and cross-price demand elasticities. The mean own-price elasticity is around -2.1 (in both the at-home and on-the-go segments), though with significant variation around this: 25% of products have own-price elasticities with magnitude greater than 2.5, a further 25% of products have own-price elasticities with magnitude less than 1.6. The distribution of the cross-price

[46] See Berry and Haile (2020) for a formal treatment of identification of choice models with micro data.

[47] 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.

elasticities exhibits a high degree of skewness, with the mean close to the 75^{th} percentile. This reflects consumers' willingness to switch between products close together in product characteristic space.

To illustrate this, Table 5.1 shows product level elasticities associated with a price change for two popular sizes – a 2l bottle and a 10 pack of 330ml cans – of Coke Regular and Diet Coke. It shows the impact on demand for each of the 2l bottle and 10×330ml packs of Coke and Pepsi, and the mean elasticities for other (non-cola) sugar-sweetened and diet beverages, and for pure fruit juice. The table highlights a number of intuitive patterns: (i) consumers are more willing to switch across cola products of the same variety (e.g. within Regular) than they are to other varieties (e.g. Diet or Zero) or to non-cola drinks; (ii) consumers are more willing to switch between products of the same size/pack type than they are to different sizes; (iii) consumer substitution from sugary varieties of Coke to sugary non-cola drinks (both sugar-sweetened beverages and fruit juice) is stronger than it is from Diet Coke. In Appendix D we report further details of product level elasticities.

Table 5.1: *Selected elasticities for cola products*

	Coke						Pepsi				Non-colas		
	Regular		Diet		Zero		Regular		Max		SSBs	Diet	Fruit juice
	2l b.	10 pk.	2l b.	10 pk.	2l b.	10 pk.	2l b.	10 pk.	2l b.	10 pk.			
Regular													
2l bottle	-2.204	0.018	0.011	0.009	0.011	0.009	0.023	0.017	0.012	0.009	0.007	0.003	0.005
10x330ml can	0.036	-2.832	0.017	0.022	0.017	0.021	0.035	0.042	0.017	0.021	0.013	0.006	0.008
Diet													
2l bottle	0.010	0.007	-2.185	0.014	0.019	0.014	0.010	0.007	0.020	0.014	0.003	0.006	0.003
10x330ml can	0.014	0.018	0.026	-2.777	0.027	0.032	0.014	0.017	0.026	0.031	0.006	0.011	0.005

Notes: Numbers show price elasticities of market demand (for products listed in top row) in the most recent year covered by our data (2012) with respect to price changes for two specific pack sizes of Coke Regular and Diet Coke (shown in first column). "Non-colas" exclude Coke and Pepsi and are means over products belonging to each of the sets, sugar-sweetened beverages (SSBs), diet drinks and fruit juices.

Switching between sets of products

In Table 5.2 we summarize the effects of increasing the price of all sugar-sweetened beverages by 1%.[48] This leads to a 1.41% fall in liters demanded of sugar-sweetened beverages. Around 33% of the reduction in demand for sugar-sweetened beverages is diverted to alternative drinks. As we discuss in Section 2, if alternative products are supplied non-competitively the degree of switching to them in response to a marginal

[48] To calculate the confidence intervals, we obtain the variance-covariance matrix for the parameter vector estimates using standard asymptotic results. We then take 100 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).

tax rise is an important determinant of the optimal policy. The diversion ratio in Table 5.2 does not directly tell us this since (i) it reflects only demand responses to a price change, but not supply side pricing responses to a tax change and (ii) it is evaluated at observed prices and not at the optimal tax rate. Nonetheless, as we show in Section 6, the relatively high degree of substitution between the two product sets indicated by the diversion ratio plays an important role in determining optimal sugar-sweetened beverage taxation.[49] The 1% increase in the price of sugar-sweetened beverages (at observed prices) leads to essentially no change in overall drinks expenditure.

Table 5.2: *Switching due to an increase in the price of sugar-sweetened beverages*

Own price elasticity for sugar sweetened beverages	-1.41
	[-1.46, -1.37]
% lost demand diverted to substitute drinks	32.8
	[32.2, 33.9]
% change in overall drinks expenditure	0.047
	[0.030, 0.062]

Notes: We simulate the effect of a 1% price increase for all sugar-sweetened beverage products. The first row shows the % reduction in volume demanded of sugar-sweetened beverages, the second row shows how much of the volume reduction is diverted to substitute drinks products, and the third row shows the percent change in total drinks expenditure. Numbers are for the most recent year covered by our data (2012). 95% confidence intervals are given in square brackets.

5.2 Estimated costs and margins

We use the first order conditions of the firms' profit maximization problem (equation (4.5)) to solve for product marginal costs, and hence the price-cost margins and Lerner indexes (margin over price) at observed prices. In Table 5.3 we show the averages of these for sugar-sweetened beverages and alternative products.

The average Lerner index is 0.57 for sugar-sweetened beverages and 0.55 for alternative products. This indicates that firms exercise a significant degree of market power when setting the prices both of sugar-sweetened beverages and alternative drinks. As we illustrate in Section 6, failing to account for distortions from the exercise of this market power leads to substantial unrealized welfare gains when setting tax policy. In Appendix D we show that there is substantial variation in equilibrium margins across brand and that average margins are declining in product size (since, on average, price per liter is strongly declining in size, while marginal cost per liter is flatter across the size distribution).

[49] There is substantial switching between these broad sets of products despite product level cross price elasticities that are relatively low in magnitude compared to product level own price elasticities. This is because we model choice over a large number of very disaggregate products.

Table 5.3: *Summary of costs and margins*

	Sugar sweetened beverages	Alternative products
Price (£/l)	1.09	1.07
Marginal cost (£/l)	0.42	0.44
Price-cost margin (£/l)	0.67	0.63
Lerner index (margin/price)	0.57	0.55

Notes: We recover marginal costs for each product in each market. The table shows the average price, marginal cost, price-cost margin (all expressed in per liter terms) and Lerner index among sugar-sweetened beverages and substitute products, constructed using quantity weights. We report the values for the most recent year covered by our data (2012).

It is important to emphasize that distortions from the exercise of marker power are endogenous to tax policy. For instance, in the case of a single tax rate applied to sugar-sweetened beverages; a marginal tax rise may exacerbate distortions from market power on these product if firms respond by raising their tax-exclusive prices (and hence product level margins) and/or consumers respond by downsizing to smaller high margin products.

5.3 Discussion of demand and supply estimates

Our demand and supply estimates enable us to capture rich consumer level substitution patterns across products, as well as product level price-cost margins. These are key inputs into our empirical study of optimal taxation with market power. However, this necessarily entails making identifying and functional form assumptions about the nature of demand and firm competition. Here we compare our estimates with those in the existing literature, and to alternative information on price-cost margins. We also provide validation of our model using data on price changes following the introduction of the UK's soft drinks tax.

We estimate an own-price elasticity for sugar-sweetened beverages of 1.41. We calculate this by simulating an increase in the prices of all sugar-sweetened beverages by 1% and recovering the change in demand for those products, allowing for substitution between sugary beverages, to alternative drinks and to not buying drinks. Allcott et al. (2019a) employ an alternative approach, using US scanner data and an instrumental variable methodology applied to quarterly purchases of sugar-sweetened beverages. They estimate an own-price elasticity for sugar-sweetened drinks in line with ours (between -1.37 and -1.48, depending on the specification).

Our model allows us to simulate how firms choose to adjust their margins when a tax is introduced. A potential concern with simulated pass-through of a hypothetical tax is that it can be influenced by functional form assumptions. We seek to alleviate this concern through specifying a rich demand model that we estimate using micro level data.[50] We also provide direct evidence that our model succeeds in generating realistic tax pass-through predictions by simulating the introduction of the UK's Soft Drinks Industry Levy (SDIL) in 2018 and comparing this with what happened in practice – see Appendix E for full details. This tax was introduced more recently than the period covered by our data, so we use a weekly database of UPC level prices, collected from the websites of six major UK supermarkets, that cover the period 12 weeks before and 18 weeks after the introduction of the tax.[51] We find evidence that the tax was slightly overshifted, with average pass-through rates of 105-108% and no change in the price of untaxed products. The price changes predicted by our model are very close to the observed price changes.

These patterns are broadly consistent with the literature that conduct ex post evaluations of the effects of sugar-sweetened beverage taxes on prices. For example, the Philadelphian tax was found to be fully passed through to prices (Seiler et al. (2021), Cawley et al. (2018)), and in Mexico the tax was fully to slightly more than fully passed through to prices (Grogger (2017), Colchero et al. (2015)). An exception is Berkeley, where pass-through of the tax is estimated to be statistically insignificant or low (e.g. Rojas and Wang (2017), Bollinger and Sexton (2018)).

Finally, we note that our estimates of price-cost margins are consistent with those in accounting data. Our estimate of the average Lerner index in the market is 0.56; gross margins reported in accounting data in this market are between 35-70% (see Competition Commission (2013)).

6 Optimal sin tax results

In this section we combine our empirical model of the UK drinks market with the tax design framework that we outline in Section 2; see Appendix F for the solution algorithm. We begin by considering the efficiency maximizing sin tax rate, where the government seeks to minimize allocative distortions, but is indifferent to the burden of the tax across individuals. Next we consider the optimal sin tax rate when the government has distributional concerns, highlighting the importance of inequality in consumption and profit holdings across individuals in determining

[50] An important determinant for tax pass-through is the curvature of market demands (for instance, see Weyl and Fabinger (2013)). Structural demand models often place restrictions on demand curvature. One way of relaxing restrictions on the curvature of market demands is through incorporating rich individual level preference heterogeneity (see Griffith et al. (2018)).

[51] The supermarkets are the big four – Tesco, Asda, Sainsbury's and Morrisons – as well as smaller national chains Iceland and Ocado. 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.

optimal policy. Finally, we consider the potential welfare gains from more flexible policy instruments (e.g., multi-rate taxation), and from levying tax directly on product sugar content rather than volumetrically.

The welfare effect of tax policy depends on the magnitude of externalities from sugar-sweetened beverages. Consumption of these products can increase health care costs, the bulk of which are not borne by the individual.[52] We measure the monetary value of externalities using epidemiological evidence from Wang et al. (2012), who estimate the impact of a reduction in sugar-sweetened beverage consumption on health care costs, and the World Health Organization (2003), who, based on a review of the medical consequences of added sugar intake, recommend it should make up less than 10% of dietary calories. Using this information, we approximate the externality per 10g of sugar consumption as 5.3 pence for the 80% of consumers with dietary sugar above 10% of their calorie intake. This translates into an average externality for this group of consumers of 14 pence per 10 oz of sugar-sweetened beverage,[53] and an average across all consumer of 11 pence. We provide further details in Appendix F, and below we show how our results vary with alternative assumptions about the nature of externalities.

If a marginal change in drinks tax policy causes consumption changes outside the drinks market, distortions from the exercise of market power in the supply of non-drinks products will impact optimal policy. In our baseline results we assume that the numeraire good is competitively supplied. However we show below that, since the impact of a marginal tax change on total drinks spending (and hence numeraire good consumption) is small, our results are numerically insensitive to this assumption.

6.1 Efficiency maximizing policy

Efficiency maximizing tax policy minimizes the allocative distortions resulting from governmental (distortionary tax) and non-governmental (externality and market power) distortions. It is indifferent to the distribution of welfare gains across individuals. The efficiency maximizing policy corresponds to the maximum of the

[52]The Scientific Advisory Committee on Nutrition (2015) reviews the medical literature. The liquid nature of sugar in drinks means that it is digested quickly, leading to insulin spikes and a higher propensity to develop type II diabetes. It also means that they are less likely than solid sources of calories to sate appetites, and therefore are associated with weight gain. In the UK it is estimated that the costs of treating obesity and related conditions added £5.8 billion in 2006-07 to the costs of public health care provision (Scarborough et al. (2011)). Cawley and Meyerhoefer (2012) estimate 88% of the US medical costs of treating obesity are borne by third parties.

[53]The average sugar-sweetened beverage has 26g of sugar per 10 oz. For those consumers with excess added sugar in their diets, product-level externalities from sugar-sweetened beverages range from 5 to 18 pence per 10 oz.

social welfare function (equation (2.1)) when all individuals have social marginal welfare weights equal to one.

To illustrate the impact of market power on the efficiency properties of sin taxation, we consider a single tax rate applied to the set of sugar-sweetened beverages (i.e., the sin products, which comprise product set $\mathcal{S}$). We summarize the impact of the tax in the first row of Table 6.1. The efficiency maximizing tax rate is 4.19 pence per 10 oz; at the time of writing, US and UK sweetened beverage taxes range from 7 to 15 pence per 10 oz. The tax leads to an average increase in the price of sugar-sweetened drinks of 19.0% (median tax pass-through is 115%), and little change, on average, in the price of alternative drinks. This results in a 28.7% reduction in consumption of sugar-sweetened beverages, and a 7.0% increase in the consumption of alternative drinks.[54] The tax leads to substantial losses in consumer surplus (£510m, compared to total expenditure in the overall drinks market of approximately £9b) and moderate profit losses (£190m, or around 3.5% fall in market variable profits). However, this is more than made up for by a £386m (25%) fall in externality costs and £409m in excise tax revenue. Overall, economic efficiency increases by £94m.

The efficiency maximizing tax rate can be expressed as a linear combination of three components – terms reflecting distortions from externalities, the exercise of market power for sin products, and the exercise of market power for alternative (substitute) products. To illustrate the role that each plays in determining efficiency maximizing tax policy, it is instructive to consider two naive policies, one that ignores distortions from market power for all goods, and one that ignores them for alternative products.[55]

A government that completely ignores distortions from market power would set a tax rate equal to 11.97 pence per 10 oz. This policy ignores the fact that equilibrium prices are set in excess of marginal costs and results in a sub-optimally high rate. The second row in Table 6.1 shows that this leads to a fall in consumer surplus and profits that is over twice as large as under efficiency maximizing policy, with the combined loss outweighing the fall in externality costs plus tax revenue meaning that overall economic efficiency falls by £71m. A government that takes account of market power distortions but only among the taxed sin products, would choose to set a tax rate that is approximately zero; the positive equilibrium margins

[54] This fall in sugar-sweetened beverage consumption is due to an 14.3% reduction in the probability, on average, a consumer purchases from this product set and, conditional on buying, a reduction in volume of 11.9%.

[55] In the first case the government sets τ that satisfies $\tau = \bar{\phi} + \frac{n(\mathcal{S})}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}\text{cov}(\phi^i_j, dx^i_j/d\tau_{\mathcal{S}})$, in the second case it sets the rate that satisfies $\tau = \bar{\phi} + \frac{n(\mathcal{S})}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}\text{cov}(\phi^i_j, dx^i_j/d\tau_{\mathcal{S}}) - \bar{\mu}_{\mathcal{S}}$.

for the sin goods offset the externality associated with their consumption. However, this policy results in a sub-optimally low tax rate. It ignores the fact that other products (substitutes to sin products) are supplied non-competitively. As we show in Tables 5.2 and 5.3, substitution to alternative drinks products is substantial, and these products (like sin products) have substantial mark-ups. Efficiency maximizing policy depends on the average margins for sin goods relative to those on alternatives.

Table 6.1: *Efficiency maximizing single rate sin taxation*

				Change (relative to zero tax) in: (£m)				
	Tax rate (p/10oz)	% change in Price	% change in Cons.	Consumer surplus	Total profits	Excise tax revenue	External cost savings	Total efficiency
Optimal	4.19	19.0%	-28.7%	-510	-190	386	409	94
Pigouvian	11.97	55.8%	-59.2%	-1199	-429	808	749	-71

Notes: Optimal refers to efficiency maximizing policy. Pigouvian refers to policy set by a government that ignores distortions from the exercise of market power. Price and consumption changes are for sugar-sweetened beverages. Consumer surplus, Total profits, External costs, Excise tax revenue and Total efficiency numbers are per annum and report the change relative to no drinks taxation. Total profits are inclusive tax revenues from taxation of corporate profits. Total efficiency = Consumer surplus+Total profits+Excise tax revenue+External cost savings.

6.2 Optimal policy with distributional concerns

When the government has distributional concerns, it must balance efficiency with equity considerations. In this case, the distribution of the effects of tax policy across individuals, which depends both on the distribution of consumption and profit holdings, matters for optimal policy.

Profits flow to the government (via corporate and dividend taxes), and to domestic and overseas residents. Measuring stock ownership across the income distribution is challenging and remains a topic of considerable debate. Recent papers, including Saez and Zucman (2016) and Smith et al. (2020), use a combination of dividend income and realized capital gains to estimate wealth in publicly traded stocks. In this spirit, we use information from the UK national accounts and the distribution of dividend income to allocate profits to different groups. The effective average corporate (see Bilicka and Devereux (2012)) and dividend tax rates leads to the government collecting 29% of profits. Using data from the national accounts, we set the fraction flowing overseas to 30%. We assume that the remaining 41% is distributed to UK residents in proportion to the share of (net-of-tax) dividend income received by households in equivalized income bands. The net-of-tax profit holdings of domestic residents is concentrated among the relatively wealthy; households with equivalized income below £10k make up around 25% of the population, but receive less than 3% of post-tax domestic dividend income; households with equivalized income above £45k comprise 5% of the population, but receive more than 21%. Sugar sweetened beverage taxation will mainly affect the profits of drinks firms. We assume that profit holdings in these firms are approximated by profit holding in the economy more generally (which is reasonable based on diversified investment portfolios). However, as a robustness exercise we also show results in the case when post-tax profits flow to individuals with zero marginal social welfare weights. In Appendix F, we provide full details of these calculations.

We parameterize the social marginal welfare weights as $g^i = (\tilde{y}_i)^{-\vartheta}$, where $\tilde{y}_i$ is equivalized household income and ϑ captures the degree of inequality aversion in government preferences.[56] As our baseline we set $\vartheta = 1$. In all calculations, we assume the government places a social marginal welfare weight of zero on the portion of profits that flows to overseas individuals.

We first illustrate how the distribution of consumption and profit holdings affect policy for a fixed set of government preferences for equity, before showing how the strength of these preferences affect the optimal rate. In the final part of this section, we show how the market structure and nature of externalities affects optimal policy.

How the distribution of consumption and profits affects policy

In Table 6.2 we summarize the impact of distributional concerns on the optimal sin tax rate. The first row of the table shows the optimal rate, and its impact on welfare, under our measure of the true distribution of profit holdings (as described above). The optimal tax rate in this case is 5.97 pence per 10 oz – over 40% higher than the efficiency maximizing rate (4.19p/10oz) – and it achieves an increase in social welfare of £167m. There are three channels through which distributional concerns affects the optimal rate. We conduct two thought experiments, based on counterfactual distributions of profits, to highlight the relative importance of these different channels.

First, we consider what the optimal rate would be if the government were to collect all profits as tax revenue (row two of the table). This isolates the impact of distributional concerns arising purely from consumption patterns. Since sugar-sweetened beverages are more popular with low income households, the optimal tax in this case lies *below* the efficiency maximizing rate (it is 3.23 pence per 10 oz). A similar effect is highlighted by Allcott et al. (2019a). Second, we consider optimal

[56] This is a common approach in the optimal taxation literature, see, for example, Saez (2002) and Allcott et al. (2019a).

policy if the share of profits flowing to overseas individuals equals the true share, but the government collects all domestically owned profits as tax revenue (row three). This introduces a second channel through which distributional concerns impact optimal policy; since a share of profits are owned by foreigners, who, under our social preference specification, are assigned social marginal welfare weights of zero, the government places less weight on the fall in profits resulting from tax policy. This leads the optimal tax rate to increase from 3.23 (when government is assumed to tax fully foreign as well as domestic profits) to 4.71 pence per 10 oz.

Comparing the final row with the first illustrates the importance of the third channel – the impact of the unequal distribution of domestic profits in combination with the government's preference for equity. This leads the optimal tax rate to increase further (from 4.71 to 5.97 pence per 10 oz). Domestic profits are disproportionately in the hands of high income households, meaning the incidence of profit losses associated with the sin tax is mainly on those with relatively low social marginal welfare weights. All else equal, this increases the progressivity of the tax, raising the optimal rate and the size of associated welfare gains.

Table 6.2: *Impact of distributional concerns on optimal sin tax policy*

		Change (relative to zero tax) in:					
		Welfare components (£m)					
	Tax rate	Private welfare, from:		Tax revenue:		Ext. cost	Total
	(p/10oz)	Cons.	Profits	Sin tax	Profit tax	savings	welfare
True profit distribution	5.97	-747	-43	522	-74	509	167
All profits taxed at 100%	3.23	-439	0	336	-152	311	56
Domestic profits taxed at 100%	4.71	-611	0	445	-147	424	111

Notes: Numbers summarize the effect of policy when the social marginal welfare weight on foreign individuals is 0 and on domestic individuals is $1/\tilde{y}_i$, showing effects under the true distribution of profit holdings (row (1)) and counterfactual distributions (rows (2) and (3)). Welfare numbers are per annum and report the change relative to no drinks taxation. Total welfare = Private welfare+Tax revenue+External cost savings.

In summary, relative to an efficiency maximizing government, distributional concerns lead to an increase in the optimal tax rate. This is because post-tax profits are mainly in the hands of foreign and relatively wealthy domestic residents, who have relatively low or zero social marginal welfare weights. This effect more than offsets the influence of distributional concerns over consumption patterns, which, all else equal, act to lower the optimal rate since sugar-sweetened beverages are disproportionately consumed by relatively low income (high social marginal welfare weight) consumers.

How the strength of preferences for equity affects policy

In Figure 6.1 we summarize how differences in the strength of social preferences for equity impact the optimal tax rate (panel (a)), the associated welfare gain (panel (b)), and the fraction of possible welfare gains forgone if the government ignores distortions from the exercise of market power when setting policy[57] (panel (c)). On the horizontal axis we plot efficiency maximizing policy, and policy when the social marginal welfare weight on foreign profits is 0 and the weights on domestic consumers are $(\tilde{y}_i)^{-\vartheta}$ for $\vartheta = \{0, 1, 2, 3, 4\}$.

When $\vartheta = 0$ the government places a social marginal welfare weight of zero on foreign profits, which, relative to efficiency maximizing policy, results in a higher optimal tax rate and associated welfare gain. With $\vartheta = 0$ the government is indifferent to inequality across domestic individuals. When $\vartheta > 0$ the fact that consumption is concentrated among those with low incomes, all else equal, acts to lower the optimal rate. However, this is offset by post-tax domestic profit holdings being concentrated among those with high incomes (meaning they bear a disproportionately large share of profit reductions from sin taxation), which acts to raise the optimal rate. ϑ controls the strength these effects exert on optimal policy. The optimal rate (and associated welfare gain) peaks at $\vartheta = 1$. When $\vartheta > 1$ the impact of consumption inequality on the optimal rate dominates the effect of inequality in domestic profits holding. Overall, the optimal tax rate and associated welfare gain are relatively insensitive to the inequality aversion parameter, and for all values of ϑ, the optimal tax rate and associated welfare gain are larger than under efficiency maximizing policy.

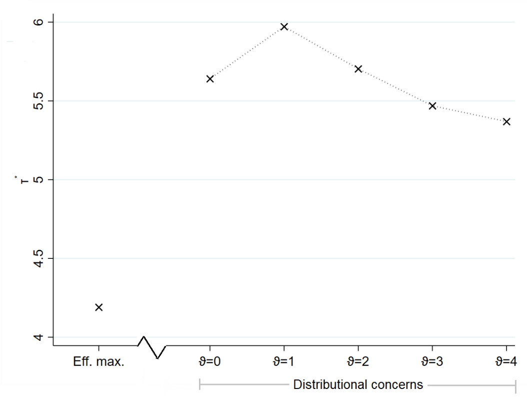
Panel (c) shows that the welfare losses from ignoring distortions due to the exercise of market power are mitigated under social preferences for equity, but they remain substantial. Policy set by an efficiency maximizing planner that ignores market power distortions leads to forgone welfare gains of more than 150%. The forgone welfare gains from this form of naive policy when there are distributional concerns are between 64% (when $\vartheta = 0$) and 35% (when $\vartheta \geq 2$). With stronger preferences for equity, since post-tax profits are disproportionately in the hands of those with relatively low or zero social marginal welfare weights, the government's welfare function places less weight on profits (than under efficiency maximization), and therefore ignoring market power is less costly. However, even with strongly inequality averse preferences there remains a substantial loss from ignoring market power – this is because some profits are collected by the government as tax revenue,

[57]Solving for the τ that satisfies $\tau = \bar{\phi} + \frac{n(\mathcal{S})}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}}\text{cov}(\phi^i_j, dx^i_j/d\tau_{\mathcal{S}}) + \frac{1}{d\mathbb{X}_{\mathcal{S}}/d\tau}\text{Cov}\left(g^i, \sum_{j\in\mathcal{S}} x^i_j\right)$.

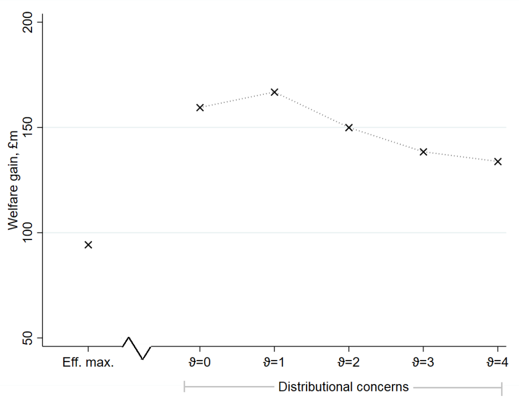
and because low and moderate income consumers have small, but non-zero, profit holdings.

Figure 6.1: *Impact of different preferences for equity on optimal sin tax policy*

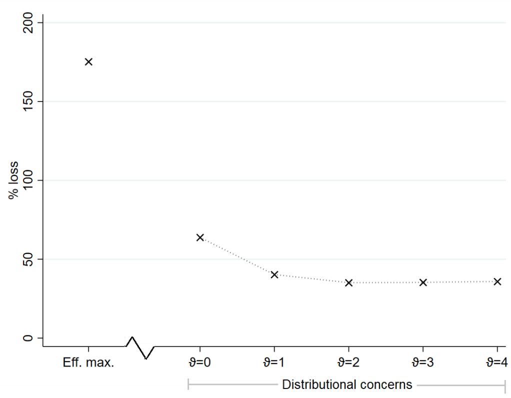
(a) Optimal tax rate



(b) Welfare gain



(c) Loss from ignoring market power



Notes: The horizontal axis plots social preferences. Eff. max. corresponds to an efficiency maximizing government. Under distributional concerns the social marginal welfare weights on overseas individuals are 0 and the weights on domestic individuals are $(\tilde{y}_i)^{-\vartheta}$*. Panel (a) shows how the optimal tax rate (p/10oz) varies, panel (b) shows how the gain in welfare (per annum) relative to no drinks tax varies and panel (c) shows how the % welfare loss from policymaking that ignores market power varies.*

Another possibility is that, while the government has inequality averse preferences with respect to consumption, it places a social marginal welfare weight of zero on all *post*-tax profits, perhaps due to uncertainty in measuring the distribution of post-tax profits across individuals. Even in this case market power and the realization of profits has an important bearing on optimal sin tax policy, because sin taxation leads to spillovers to the profit (i.e., corporate and dividend) tax bases. The optimal sin tax rate in this case is 6.90 pence per 10 oz. If the government ignores market power when setting the tax rate (and hence the spillovers to the corporate and dividend tax bases) it would set a rate of 11.26, which would result in unrealized welfare gains of 26% (see row (2) in Table 6.3). This illustrates that, even if post-tax profits holdings are more concentrated in the hands of the rich than our measure of the profit distribution suggests (which is based on post-tax dividend income), market power continues to have an important bearing on optimal tax policy.

How market structure affects policy

Our baseline results assume that the numeraire good is competitively supplied. However, it is possible that when consumers switch away from drinks they switch to other non-competitively supplied goods. This acts to raise the optimal tax rate through an efficiency channel, as it dampens the market power correction element of the optimal tax (which equals the average margin of sin goods relative to alternatives). Yet it acts to lower the optimal rate through an equity channel, as switching to a non-competitively supplied numeraire good leads to off-setting profit gains mainly for the rich. Row (2) in Table 6.3 shows results when the numeraire good is supplied non-competitively, with a margin equal to that implied by the estimate of the UK economy-wide mark-up in De Loecker and Eeckhout (2018). It shows that the optimal tax rate is 6% higher and the associated welfare gain marginally higher, than in our baseline with a competitively supplied numeraire. The modest impact of adding market power for the numeraire on optimal policy is due to the relatively small fall in overall drinks expenditure (and hence rise in numeraire good consumption) induced by a marginal tax rise.

In row (3) of Table 6.3 we consider optimal tax policy under the counterfactual market structure of a perfectly competitive drinks market. Without any tax in place, moving from the true market structure to perfect competition raises welfare by £622m; the gains from eliminating market power distortions dwarf the resulting increase in externality costs. The optimal tax rate under perfect competition is 10.18 pence per 10 oz and the resulting welfare gain is £515 million, which is around 3 times as large as under the true market structure (row (1)). Competition and optimal tax policy therefore exhibit a form of complementarity.[58]

[58]This need not be the case. For instance, consider the textbook setting of a single monopoly seller and homogenous buyers. Suppose the monopolist sets a fixed margin on its product given by $\mu \leq \phi$, where ϕ is a per-unit externality. When $\mu = \phi$ we have the first best. A more competitive seller (that sets a lower μ) will reduce welfare. However, the optimal tax rate, $\tau = \phi - \mu$ will exactly offset this, bringing the market back to the first best. In this case competition and tax policy are perfect substitutes.

How the nature of externalities affects policy

Rows (5)-(10) of Table 6.3 show how changes in the shape of the function mapping sugar-sweetened beverage consumption to externalities impact tax policy. Rows (5) and (6) vary the overall magnitude of external costs, setting them at 25% above and 25% below the baseline value, implied by Wang et al. (2012). Higher consumption externalities lead to an increase in the optimal tax rate and the associated welfare gains. However, even with larger externality costs, the losses from ignoring distortions from the exercise of market power remain substantial (at 23%).

Row (7) and (8) show results when we vary the convexity of the externality function (holding fixed the average marginal externality across consumers). Row (7) assumes marginal externalities are equal for all consumers; row (8) assumes that only those with total dietary sugar above 15% (compared with 10% in the baseline) generate externalities through their consumption. The consumption of people with higher overall dietary sugar is moderately more responsive to tax changes. Therefore, the more concentrated externality generation is among overall high sugar consumers, the more effective the sin tax is at reducing externality distortions. Hence, more convexity in the externality function leads to a higher optimal rate and larger resulting welfare gains.

We also consider the consequence of externality spillovers. In particular, consumers respond to the sugar-sweetened beverage tax by, in part, switching to pure fruit juices and flavored milk. These are typically exempt from drinks taxes, in part because these products contain other (positive) nutrients that may offset negative consequences of sugar intake. However, this is subject to debate. In row (9) we show the consequences for optimal policy when externalities are associated with the untaxed alternative sources of sugar. This type of externality leakage diminishes the effectiveness of tax in reducing externality distortions, leading to a lower optimal rate and smaller welfare gains (relative to the baseline, where consumption of these alternative products is assumed not to generate externalities). Row (10) quantifies the gains of including pure fruit juices and flavored milk in the tax base, in the case where consumption of these products is associated with externalities. This results in a higher optimal rate and welfare gains that are £43m larger – untaxed externalities lead to sizeable forgone welfare gains.

Table 6.3: *Optimal tax policy and the costs of ignoring market power*

		Optimal policy		Ignoring market power		
		Tax rate (p/10oz)	Change in welfare (£m)	Tax rate (p/10oz)	Change in welfare (£m)	% loss
(1)	Baseline	5.97	167	11.26	100	40%
(2)	Zero weight on post-tax profits	6.90	213	11.26	157	26%
Market structure						
(3)	Numeraire good market power	6.34	169	11.26	116	32%
(4)	Perfect competition	10.18	515	10.18	515	0%
Externalities						
(5)	25% larger	8.44	312	14.63	239	23%
(6)	25% smaller	3.49	61	8.01	-2	103%
(7)	Linear	5.35	138	10.33	73	47%
(8)	More convex	7.88	267	13.98	194	27%
(9)	Leakage	5.24	153	10.12	105	31%
(10)	Broader base (+leakage)	6.25	196	12.97	109	44%

Notes: The first two columns summarize optimal policy, the final three summarize policy set by government that ignores distortions from market power (with the final column showing the % of welfare gains from optimal policy forgone). All numbers are based on social marginal welfare weights on foreign individuals of 0 and on domestic individuals of $1/\tilde{y}_i$. *Row (1) repeats numbers under our central calibration of the numeraire good margin and externalities. Row (2) sets social marginal welfare weights on post-tax profits to zero. The remaining rows present results under alternative market structures and externality functions, with details described in the text. Welfare numbers are per annum and report the change relative to no drinks taxation.*

6.3 Optimal policy with alternative tax instruments

Our focus to this point is on a volumetric single tax rate levied on sugar-sweetened beverages. This form of taxation is very common among the taxes that have actually been implemented.[59] In this section we consider alternative tax instruments. In particular, we consider a system of multiple volumetric tax rates applied in the drinks market, and a tax that is levied directly on product sugar content.

Multi-rate system

Levying a single tax rate on sugar-sweetened beverages is simple, but may leave substantial unrealized welfare gains relative to more flexible instruments. It is not feasible for the government to set product specific prices in the unregulated non-alcoholic drinks market. We instead consider the gains from setting separate tax rates on different sets of products (e.g., colas, lemonades etc.). An excise tax

[59] As of May 2021, of the 44 countries and 9 US cities that have implemented sugar-sweetened beverage taxes, only Mauritius, South Africa and Sri Lanka have taxes levied directly on sugar content. In the large majority of cases, sugar-sweetened beverage taxes entail a single rate. Exceptions are the UK and Portugal, which have banded systems entailing two rates.

system of this nature has a precedent in existing alcohol tax systems. Throughout we consider a government with distributional concerns, so our policy simulations optimally balance efficiency and equity considerations.

In total we consider four different tax systems. These systems vary across two dimensions: (a) the number of tax rates available to the government and (b) the constraint placed on the government's ability to subsidize drinks consumption. We consider a two-rate system that sets different rates on sugar-sweetened and alternative drinks, and a multi-rate system that sets different rates on 12 drinks types.[60] We consider two variants of each system. The first requires that the total effect on the government's budget (inclusive of the pecuniary externality) is non-negative; we refer to this as the "no budget deterioration" case. The second requires all tax rates be non-negative; we refer to this as the "no subsidy" case. Table 6.4 summarizes the impact of these alternative tax systems on price, consumption and welfare.

The no budget deterioration variant of the two-rate system (column (1)) sets a tax rate of 4.63 pence per 10 oz and a subsidy of 3.65 pence per 10 oz for alternatives. This leads to price changes of +22.0% and -15.5%, respectively. Consumption of sugar-sweetened beverages falls by 38.5%, with a 37.6% rise in the consumption of alternative drinks. Social welfare rises by £248m, which is 50% higher than under the single rate. Private welfare increases and the government budget improves, with the latter due to a large decrease in external costs offsetting falls in excise tax revenue.[61] Under the no subsidy two-rate system – shown in column (2) – the tax rate on alternative drinks is zero, and the system is identical to the optimal single rate sugar-sweetened beverage tax rate.

Columns (3) and (4) summarize the impact of the no budget deterioration and no subsidy variants of the multi-rate system. The additional flexibility of the multi-rate systems further improve welfare by 20% and 16%, relative to the no budget deterioration and no subsidy two-rate systems, respectively. This is because it can better target both the externality and market power distortions. To quantify the relative importance of these two channels, we compute the gains from moving from a single sin tax rate to the no subsidy multi-rate system under the counterfactual market structure of perfect competition. Under perfect competition, the only gains from increasing the number of rates is to better target the most socially costly consumption, e.g., products with high sugar contents or that are particularly popular

[60]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.).

[61]An alternative policy that requires that excise tax revenue is non-negative leads to similar changes in overall welfare as the no budget deterioration. However, in contrast, this results from a large improvement in the government budget more than offsetting falls in private welfare.

with consumers whose intake creates externalities. Under the true market structure, there are additional gains from accounting for variation in equilibrium margins across different drinks. The welfare gains from moving to a multi-rate system under perfect competition are 1/4 as large as those under the true market structure. Hence, 3/4 of the gains associated with optimal tax rate differentiation are realized due to the tax system being better tailored to combat market power.

Table 6.4: *Multi-rate taxation*

	(1)	(2)	(3)	(4)
	Two-rate		Multi-rate	
	No budget deterioration	No subsidy	No budget deterioration	No subsidy
Tax rate (p/10oz) for:				
Sin products	4.63	5.97	4.28*	5.53*
Alternatives	-3.65	0.00	-4.01*	0.00*
Price change for:				
Sin products	22.0%	27.3%	22.5%	27.1%
Alternatives	-15.5%	-0.7%	-15.8%	-0.6%
Consumption change for:				
Sin products	-38.5%	-37.7%	-37.7%	-35.8%
Alternatives	37.6%	9.4%	35.1%	8.5%
Welfare components (£m):				
Private welfare	-1265	-790	-297	299
Consumption	*-1200*	*-747*	*-287*	*276*
Profit holdings	*-64*	*-43*	*-10*	*23*
Gov. budget	1408	957	532	0
Excise tax rev.	*684*	*522*	*0*	*-572*
Profits tax rev.	*-111*	*-74*	*-18*	*40*
Ext. cost savings	*834*	*509*	*550*	*532*
Total welfare (£m)	143	167	235	299

* *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.

A sugar tax

An alternative to the volumetric sugar-sweetened beverage taxes typically used is to levy a tax directly on the sugar in drinks. This has the advantage of making the tax proportional to the total quantity of the externality generating attribute that it contains. In Table 6.5 we compare the optimal sugar tax 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 higher increase in the average price of sugar-sweetened beverages

(30.3% relative to 27.3% under the volumetric tax). This leads to slightly higher 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 6.5: *Sugar taxation*

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

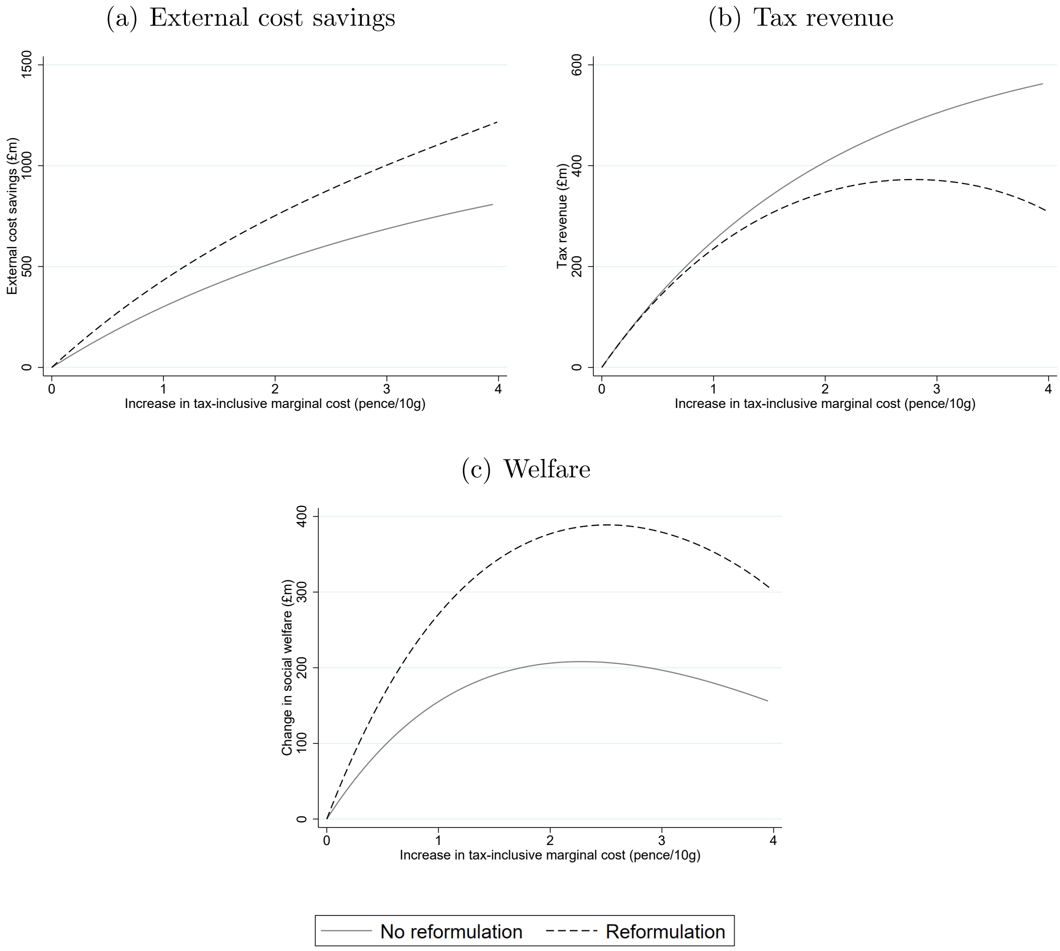
Notes: Row (1) shows the effects of the optimal volumetric tax for reference. It repeats some of the information in row (1) of Table 6.2. Row (2) shows the effects of the optimal tax on the sugar in sweetened beverages under the assumption of no product reformulation. Numbers in the final column show 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.*

Another possible advantage of a sugar tax is that it would incentivize input substitution, since a firm can lower its exposure to the tax by reducing the sugar content of its products. In contrast, under a volumetric tax, firms can avoid the tax only by removing all sugar from products, which, given almost all major sugar-sweetened beverages are already available in a zero sugar/diet version, is unlikely to be appealing to them.[62]

To highlight the consequences of input substitution for optimal sugar taxation we extend our model of firm competition, allowing firms to optimally choose both product prices and sugar contents. We follow Barahona et al. (2021) in assuming that firms can deviate from the cost-minimizing sugar content without altering product taste, but doing so raises the marginal cost of production (quadratically in the amount of sugar removed from the product). Under a sugar tax firms therefore trade-off higher production costs against reducing their tax liability. See Appendix G for full details.

[62] If we simulate the removal of sugar-sweetened products using our estimates we find all firms suffer substantial profit losses.

Figure 6.2: *Impact of sugar tax with input substitution*



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.

In Figure 6.2 we illustrate how input substitution affects sugar taxation by comparing the case when reformulation costs are high (set to be prohibitive, so firms do not choose to reformulate their product) with when they are low (where at the welfare maximizing tax rate firms choose to remove 30% of sugar from sweetened beverages leading to a 9% rise in their average production costs). The horizontal axes plot the increase in the tax-inclusive marginal cost of sugar-sweetened drinks induced by policy (expressed per 10 grams of pre-tax product sugar content). When reformulation costs are high, this equals the sugar tax rate. When they are low, the increase in tax-inclusive marginal costs reflect both the tax on sugar and the higher production costs associated with input substitution. Conditional on the tax-inclusive marginal cost, policy under high and low reformulation cost scenarios results in the same equilibrium (tax-inclusive) prices and quantities. Hence the effect on private welfare (through consumption and profits) is the same.

However, the high and low reformulation costs differ in their impact on the government's budget through their effect on externalities (panel (a)) and on tax revenue (panel (b)). Low reformulation costs lead to larger reductions in externalities, but smaller increases in tax revenue. Lower tax revenue under the input substitution scenario one-for-one reflects higher costs of production. Panel (c) shows that low reformulation costs lead to larger increases in social welfare, as the reductions in externalities more than offset the efficiency loss due to higher production costs. In Appendix G we show that the lower are reformulation costs, the higher are the welfare gains from sugar taxation. Although firms chose how much to reformulate to maximize profits, the fact that externalities from sugar are sufficiently high means that their reformulation decisions improve social welfare.

7 Conclusion

In this paper we show how market power affects the efficiency and redistributive properties of sin taxation. Allocative distortions from the exercise of market power lead optimal sin tax policy to depend on the extent of equilibrium price-cost margins on sin products, relative to alternatives. The relative concentration of profit holdings in the hands of the wealthy, lead policy to be more progressive than if no profits were realized, counteracting the regressive incidence of the tax based on consumption patterns. Even if the government places zero weight on all private profit holdings, spillovers to the corporate and dividend tax bases mean ignoring marking power when setting tax policy remains costly.

We illustrate the empirical importance of these effects with an application to the optimal taxation of sugar-sweetened beverages, showing that market power exerts an economically meaningful impact on optimal policy and ignoring it leads to substantial unrealized welfare gains. We also show the possible gains that arise through more flexible (multi-rate) tax instruments (compared to the single tax rate used in the US and many other countries), which enable policy to better target both externality and market power related distortions. In addition, we show that using an input, rather than volumetric tax, can lead to significantly better performance. Our work adds to a recent strand of literature, which shows empirically the importance of market power for tax policy; including revenue maximizing taxation (Miravete et al. (2018)), efficiency maximizing corrective taxation (Fowlie et al. (2016)), efficiency maximizing pricing in regulated markets (Conlon and Rao (2015), Miravete et al.

(2020)) and demographic targeted subsidy design (Tebaldi (2017) and Polyakova and Ryan (2019)).

We conclude by suggesting two promising avenues for future research. Throughout we have considered a government's choice over taxation in a given market. The government takes account of the competitive environment (accounting for distortions from market power and firms' strategic responses) and spillovers to other tax bases, but treats these other policy instruments as fixed. While this is realistic when considering the introduction of a new tax in single small market, for policy that affects a large swath of the economy, such as that targeted at climate change, the gains from also adjusting competition policy and other aspect of tax policy are likely to be considerable. Several papers highlight the importance of market power in markets associated with greenhouse gas emissions (e.g, see Bushnell et al. (2008) for electricity and Hastings (2004) for gasoline). An interesting avenue for future work is to consider the joint determination of commodity taxation with other aspects of policy, including competition policy and corporate taxation, in tackling systemic externalities, such as climate change.

Allcott et al. (2019a) highlight how the corrective motive arising from consumer misoptimization interacts with government redistributive objectives in sin tax design: if consumers with high social marginal welfare weights are more likely to make mistakes, this acts to make tax policy more progressive. An interaction between consumer misoptimization and market power arises if firms strategically exploit consumer mistakes, for instance, through persuasive advertising, exploiting consumer self-control problems or obfuscating the unhealthy nature of products (e.g., see Spiegler (2010)). A second interesting avenue for future research is to explore the interaction between consumer optimization and market power and its implications for sin tax design.

References

ACEA (European Automobile Manufacturers Association) (2019). Vehicles Per Capita, by Country. https://www.acea.be/statistics/tag/category/vehicles-per-capita-by-country.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019a). Regressive Sin Taxes, with an Application to the Optimal Soda Tax. *Quarterly Journal of Economics 134*(3), 1557–1626.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019b). Should We Tax Soda? An Overview of Theory and Evidence. *Journal of Economic Perspectives 33*(3), 23.

Barahona, N., C. Otero, S. Otero, and J. Kim (2021). Equilibrium Effects of Food Labeling Policies. pp. 118.

Berry, S. and P. Haile (2020). Nonparametric Identification of Differentiated Products Demand Using Micro Data. Technical Report w27704, National Bureau of Economic Research, Cambridge, MA.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics 25*(2), 242–262.

Bilicka, K. and M. Devereux (2012). CBT Corporate tax ranking 2012. *Oxford University Centre for Business Taxation*, 37.

Bollinger, B. K. and S. Sexton (2018). Local Excise Taxes, Sticky Prices, and Spillovers: Evidence from Berkeley's Soda Tax. *SSRN Electronic Journal*.

Bonnet, C. and P. Dubois (2010). Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance. *RAND Journal of Economics 41*(1), 139–164.

Bonnet, C. and V. Réquillart (2013). Tax incidence with strategic firms in the soft drink market. *Journal of Public Economics 106*, 77–88.

Bresnahan, T. F. (1989). Empirical studies of industries with market power. In *Handbook of Industrial Organization*, Volume 2, pp. 1011–1057.

Browning, M. and C. Meghir (1991). The Effects of Male and Female Labor Supply on Commodity Demands. *Econometrica 59*(4), 925–951.

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

Buchanan, J. M. (1969). External Diseconomies, Corrective Taxes, and Market Structure. *The American Economic Review 59*(1), 174–177.

Bushnell, J. B., E. T. Mansur, and C. Saravia (2008). Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets. *American Economic Review 98*(1), 237–266.

Cawley, J., D. Frisvold, A. Hill, and D. Jones (2018). The Impact of the Philadelphia Beverage Tax on Prices and Product Availability. *NBER Working Paper 24990*.

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

Chernozhukov, V., J. Hausman, and W. Newey (2019). Demand analysis with many prices. *NBER Working Paper 02138*.

Chetty, R. (2009). Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics 1*, 451–488.

Colchero, M. A., J. C. Salgado, M. Unar-Munguía, M. Molina, S. Ng, and J. A. Rivera-Dommarco (2015). Changes in Prices After an Excise Tax to Sweetened Sugar Beverages Was Implemented in Mexico: Evidence from Urban Areas. *PLOS ONE 10*(12), e0144408.

Competition Commission (2000). *Supermarkets: A Report on the Supply of Groceries from Multiple Stores in the United Kingdom.* London: The Stationary Office.

Competition Commission (2013). *AG Barr / Britvic Merger Inquiry.* London: The Stationary Office.

Conlon, C. T. and N. S. Rao (2015). The Price of Liquor is Too Damn High: Alcohol Taxation and Market Structure. *NYU Wagner Research Paper No. 2610118*.

Cooper, M., J. McClelland, J. Pearce, R. Prisinzano, J. Sullivan, D. Yagan, O. Zidar, and E. Zwick (2016). Business in the United States: Who Owns it and How Much Tax Do They Pay? *Tax Policy and the Economy*, 41.

De Loecker, J. and J. Eeckhout (2018). Global Market Power. *NBER Working Paper 24768*.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The Rise of Market Power and the Macroeconomic Implications. *The Quarterly Journal of Economics 135*(2), 561–644.

Diamond, P. A. (1973). Consumption externalities and corrective imperfect pricing. *The Bell Journal of Economics 4*(2), 526–538.

Dubois, P., R. Griffith, and M. O'Connell (2018). The Effects of Banning Advertising in Junk Food Markets. *Review of Economic Studies 85*(1), 396–436.

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

Einav, L. and J. Levin (2010, May). Empirical Industrial Organization: A Progress Report. *Journal of Economic Perspectives 24*(2), 145–162.

Fowlie, M., M. Reguant, and S. P. Ryan (2016). Market-Based Emissions Regulation and Industry Dynamics. *Journal of Political Economy 124*(1), 55.

Ganapati, S., J. Shapiro, and R. Walker (2020). Energy Cost Pass-Through in U.S. Manufacturing: Estimates and Implications for Carbon Taxes. *American Economic Journal: Applied Economics 12*(2), 303–42.

Global Food Research Program (GFRP) (2021). Sugary drinks taxes around the world. http://globalfoodresearchprogram.web.unc.edu/resources/.

Gorman, W. M. (1980). A Possible Procedure for Analysing Quality Differentials in the Egg Market. *Review of Economic Studies 47*(5), 843–856.

Griffith, R., L. Nesheim, and M. O'Connell (2018). Income effects and the welfare consequences of tax in differentiated product oligopoly. *Quantitative Economics 9*(1), 305–341.

Griffith, R., M. O'Connell, and K. Smith (2019). Tax design in the alcohol market. *Journal of Public Economics 172*, 20–35.

Griffith, R., M. O'Connell, K. Smith, and R. Stroud (2019). The evidence on the effects of soft drinks takes. *IFS Briefing Note 255*.

Grogger, J. (2017). Soda Taxes And The Prices of Sodas And Other Drinks: Evidence From Mexico. *American Journal of Agricultural Economics 99*(2), 481–498.

Grummon, A. H., B. B. Lockwood, D. Taubinsky, and H. Allcott (2019). Designing better sugary drink taxes. *Science 365*(6457), 989–990.

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

Hastings, J. S. (2004). Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California. *American Economic Review 94*(1), 317–328.

Hendel, I. and A. Nevo (2006a). Measuring the implications of sales and consumer inventory behavior. *Econometrica 74*(6), 1637–1673.

Hendel, I. and A. Nevo (2006b). Sales and Consumer Inventory. *RAND Journal of Economics 37*(3), 543–561.
Hendel, I. and A. Nevo (2013). Intertemporal Price Discrimination in Storable Goods Markets. *American Economic Review 103*(7), 2722–2751.
Hendren, N. (2016). The Policy Elasticity. *Tax Policy and the Economy 30*(1), 51–89.
Jacobsen, M. R., C. R. Knittel, J. M. Sallee, and A. A. van Benthem (2020). The Use of Regression Statistics to Analyze Imperfect Pricing Policies. *Journal of Political Economy 128*(5), 1826–1876.
Joint WHO/FAO Expert Consultation (Ed.) (2003). *Diet, Nutrition, and the Prevention of Chronic Diseases*. Number 916 in World Health Organisation Technical Report Series. Geneva: World Health Organization.
Kaplow, L. (2012). Optimal Control of Externalities in the Presence of Income Taxation. *International Economic Review 53*(2), 487–509.
Kleven, H. (2020, May). Sufficient Statistics Revisited. Technical Report w27242, National Bureau of Economic Research, Cambridge, MA.
Kopczuk, W. (2003). A note on optimal taxation in the presence of externalities. *Economics Letters 80*(1), 81–86.
Lancaster, K. (1971). *Consumer Demand: A New Approach*. Columbia University Press New York.
Luco, F. and G. Marshall (2020). The Competitive Impact of Vertical Integration by Multiproduct Firms. *American Economic Review 110*(7), 2041–2064.
McFadden, D. and K. Train (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics 15*(November 1998), 447–470.
Miravete, E. J., K. Seim, and J. Thurk (2018). Market Power and the Laffer Curve. *Econometrica 86*(5), 1651–1687.
Miravete, E. J., K. Seim, and J. Thurk (2020). One Markup to Rule Them All: Taxation by Liquor Pricing Regulation. *American Economic Journal: Microeconomics 12*(1), 1–41.
Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica 69*(2), 307–342.
Nevo, A. and C. Wolfram (2002). Why Do Manufacturers Issue Coupons? An Empirical Analysis of Breakfast Cereals. *The RAND Journal of Economics 33*(2), 319–339.
OECD (2019). Prices in the health sector. OECD Indicators.
Pigou, A. C. (1920). *The Economics of Welfare*. McMillan&Co., London.
Piketty, T., E. Saez, and G. Zucman (2018). Distributional National Accounts: Methods and Estimates for the United States. *The Quarterly Journal of Economics 133*(2), 553–609.
Polyakova, M. and S. Ryan (2019). Subsidy Targeting with Market Power. *NBER Working Paper 26367*.
Roberto, C. A., H. G. Lawman, M. T. LeVasseur, N. Mitra, A. Peterhans, B. Herring, and S. N. Bleich (2019). Association of a Beverage Tax on Sugar-Sweetened and Artificially Sweetened Beverages With Changes in Beverage Prices and Sales at Chain Retailers in a Large Urban Setting. *JAMA 321*(18), 1799.

Rojas, C. and E. Y. Wang (2017). Do Taxes for Soda and Sugary Drinks Work? Scanner Data Evidence from Berkeley and Washington. *SSRN Electronic Journal*.

Saez, E. (2002). Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses. *Quarterly Journal of Economics 117*(3), 1039–1073.

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.

Sandmo, A. (1975). Optimal Taxation in the Presence of Externalities. *Swedish Journal of Economics 77*(1), 86–98.

Scarborough, P., P. Bhatnagar, K. K. Wickramasinghe, S. Allender, C. Foster, and M. Rayner (2011). The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: An update to 2006–07 NHS costs. *Journal of Public Health 33*(4), 527–535.

Scientific Advisory Committee on Nutrition (2015). SACN Carbohydrates and Health Report. https://www.gov.uk/government/publications/sacn-carbohydrates-and-health-report.

Seiler, S., A. Tuchman, and S. Yao (2021). The Impact of Soda Taxes: Pass-through, Tax Avoidance, and Nutritional Effects. *Journal of Marketing Research 58*(1), 24–49.

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

Spiegler, R. (2010). *Bounded Rationality and Industrial Organization.* Oxford University Press.

Syverson, C. (2019). Macroeconomics and Market Power: Context, Implications, and Open Questions. *Journal of Economic Perspectives 33*(3), 23–43.

Tebaldi, P. (2017). Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA. *University of Chicago Working Paper*.

UK Government (2018). Floor Space in English Homes. *Report by the Ministry of Housing, Communities & Local Government*, 46.

U.S. Department of Energy (2019). Vehicles per Capita: Other Regions/Countries Compared to the United States. https://www.energy.gov/sites/prod/files/2017/01/f34/fotw%23962_web.xlsx.

Villas-Boas, S. B. (2007). Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data. *Review of Economic Studies 74*(2), 625–652.

Wang, E. Y. (2015). The impact of soda taxes on consumer welfare: Implications of storability and taste heterogeneity. *RAND Journal of Economics 46*(2), 409–441.

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.

Weyl, E. G. and M. Fabinger (2013). Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition. *Journal of Political Economy 121*(3), 528–583.

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

Appendix

Optimal sin taxation and market power

Martin O'Connell and Kate Smith
September, 2021

A Optimal tax formulae

A.1 Optimal policy

The welfare function (equation (2.1)) is:

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

The first order condition, which we write in terms of the policy parameter θ – $dW/d\theta = 0$ – implies:

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

where $g^i \equiv \omega^i \alpha^i / \lambda$. Using

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

we can re-write this as in equation (2.2):

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

which we can re-write as:

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

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

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

A.2 Single sin tax rate

In the special case where a single tax rate is applied to the set of sin products, $\mathcal{S}$, condition (A.1) becomes:

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

or after rearranging:

$$\tau_{\mathcal{S}} = \frac{1}{d\mathbb{X}_{\mathcal{S}}/d\tau_{\mathcal{S}}} \left(\int_i \sum_{j\in\mathcal{S}} \phi_j^i \frac{dx_j^i}{d\tau_{\mathcal{S}}} di - \sum_{j\in\mathcal{M}\bigcup O} \mu_j \frac{dX_j}{d\theta} - \int_i \tau_z^i \frac{dz^i}{d\tau_{\mathcal{S}}} di + \right.$$
$$\left. \int_i (g^i - 1)\left(\sum_{j\in\mathcal{M}} x_j^i \frac{dp_j}{d\tau_{\mathcal{S}}} - \delta^i (1 - \tau_\Pi)\frac{d\Pi}{d\tau_{\mathcal{S}}}\right) di \right),$$

where $d\mathbb{X}_\mathcal{S}/d\tau_\mathcal{S} = \sum_{j\in\mathcal{S}} dX_j/d\tau_\mathcal{S}$. Re-writing:

$$\int_i \sum_{j\in\mathcal{S}} \phi_j^i \frac{dx_j^i}{d\tau_\mathcal{S}} di = \bar{\phi} \times \frac{d\mathbb{X}_\mathcal{S}}{d\tau_\mathcal{S}} + n(\mathcal{S}) \times \text{cov}(\phi_j^i, dx_j^i/d\tau_\mathcal{S}),$$

where $\bar{\phi} \equiv \int_i \frac{1}{n(\mathcal{S})} \sum_{j\in\mathcal{S}} \phi_j^i di$, and defining:

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

we obtain equation (2.3):

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

A.3 Characterization of the tax base erosion component

The base erosion term in equation (2.3) can be expressed in terms of income and price elasticities. To see this first note that the base erosion term can be written:

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

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 $\frac{\partial z^i}{\partial p_j}$:

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

where $\tilde{x}_j^i$ 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 $(1-\tau^i_z)$:

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

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:

$$\frac{\partial \tilde{f}^i_j}{\partial \tilde{Y}^i} = \frac{\partial f^i_j}{\partial Y^i} \tag{A.4}$$

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 $\xi_j^i \equiv \frac{\partial f_j^i}{\partial z^i}\frac{z^i}{x_j^i}$ is the elasticity of good j with respect to labor earnings and $\zeta_z^i \equiv \frac{\partial z^i}{\partial(1-\tau_z^i)}\frac{(1-\tau_z^i)}{z^i}$ is the elasticity of taxable earnings.

Hence the tax base erosion terms can be written:

$$\frac{d(\int_i \mathcal{T}(z^i)di)}{d\tau_{\mathcal{S}}} = \int_i \frac{\tau_z^i}{1-\tau_z^i}\zeta_z^i \sum_{j\in\mathcal{M}} \xi_j^i x_j^i \rho_j di.$$

B Additional data tables

Table B.1 compares the demographic composition of the Kantar Worldpanel with the nationally representative Living Costs and Food Survey.

In Figure 3.1(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_{iqy} denotes the volume of sugar-sweetened beverages purchased by household i in year-quarter (y,q) for at-home consumption. $\text{income quintile}_{iy}^k$ is an indicator variable equal to 1 if household i is in income quintile k in year y, and μ_i are household fixed effect. We estimate this over the period 2008 to 2012.

Figure B.1 plots the estimated $\hat{\beta}_k^{NOFE}$ and $\hat{\beta^{FE}}_k$. 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 goes to zero 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.

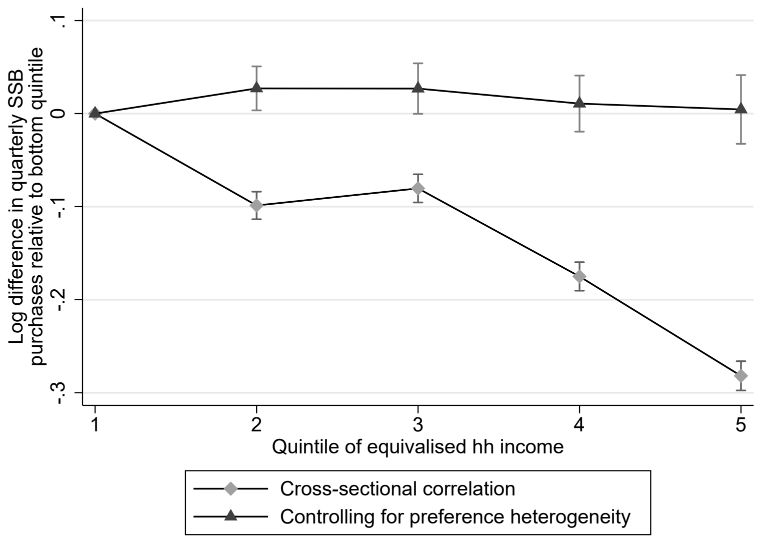
Figure B.2 replicates Figure 3.1 in the paper, but for the on-the-go segment.

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]

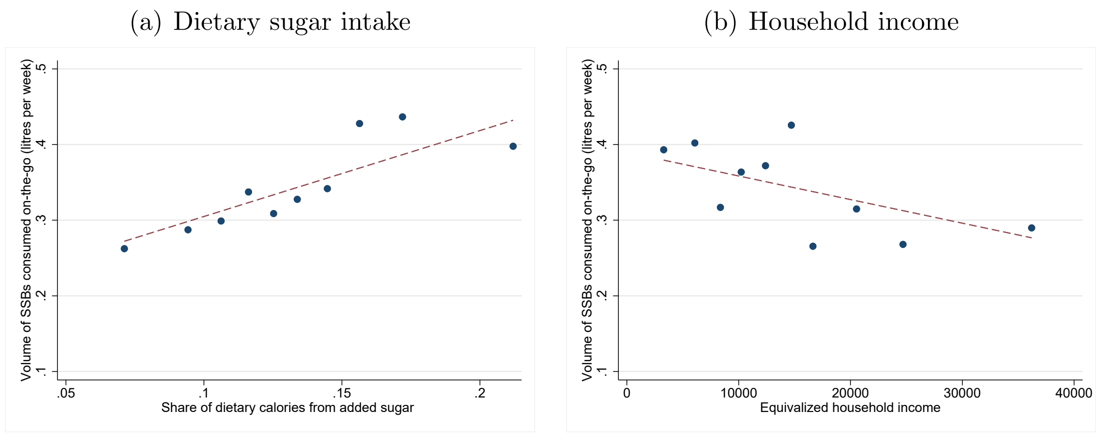
Notes: Table shows the share of households in the Kantar Worldpanel and Living Costs and Food Survey in 2012 by various demographic groups. Socioeconomic status is based on the occupation of the head of the household and is shown for the set of non-pensioner households. 95% confidence intervals are shown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

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 at home by deciles of equivalized (using the OECD-equivalence scale) household income.

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 refer to a brand-variant-size combination as a product.

Table B.2: *Firms and brands*

Firm	Brand	Type	Market share (%) At-home	Market share (%) On-the-go	Price (£/l) At-home	Price (£/l) On-the-go
Coca Cola Enterprises			*33.0*	*59.0*		
	Coke	Soft	20.4	36.3	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
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	3.4	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
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	9.1
Convenience stores	2.3	54.5
in region:		
South	–	13.6
Central	–	15.5
North	–	25.4

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

We estimate our demand model allowing all preference parameters to vary by the demographic groups shown in Table B.5. 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 B.5: *Demographic 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 Non-separabilities

We investigate whether there is evidence of two types of intertemporal non-separabilities that could invalidate our empirical approach. First, whether recent at-home purchases influence individuals' demand in the on-the-go segment of the market, and second, whether consumers stockpile in response to sales.

C.1 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.1 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 our modeling of the at-home and on-the-go segments as separate parts of the market. They also are consistent with formal test of non-separability between the segments conducted in Dubois et al. (2020).

Table C.1: *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.2 Stockpiling

We consider whether there is evidence of households in the at-home segment stockpiling drinks by conducting a number of checks based on implications of stockpiling behavior highlighted by Hendel and Nevo (2006b). Hendel and Nevo (2006b) 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 households. These are periods 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 (2006b) 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 (2006b) 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:

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

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.2 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 leads to 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 (2006b) if stockpiling behavior was present.

Time between purchases. The second and third implications of stockpiling behavior highlighted in Hendel and Nevo (2006b) 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.2 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 (2006b) 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 (2006b) 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.2: *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?			0.1451***	-0.2263***	-0.0892***
			(0.0198)	(0.0198)	(0.0016)
Mean of dependent variable	0.2271	3.9250	12.1625	12.1625	0.4638
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.3 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.3: *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 tables of 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 demographic 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.

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	× Price	-0.008	-0.009	-0.011	-0.010	-0.012	-0.010
with income		(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	1.070	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	2.639	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	0.627	1.230	1.215	1.566	0.821	0.064
		(0.205)	(0.104)	(0.130)	(0.119)	(0.095)	(0.090)
	Advertising	0.787	0.666	0.545	0.553	0.457	0.603
		(0.077)	(0.045)	(0.060)	(0.046)	(0.031)	(0.046)
Interaction	× Price	0.022	-0.013	0.013	-0.038	-0.076	-0.081
with income		(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	-0.091	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	0.038	-0.037	-0.011	-0.080	-0.036	0.082
		(0.009)	(0.005)	(0.007)	(0.006)	(0.004)	(0.005)
Variance	Price	0.530	0.083	0.030	0.273	0.120	0.115
		(0.117)	(0.013)	(0.009)	(0.049)	(0.011)	(0.018)
	Sugary	8.767	4.380	7.230	8.576	7.970	6.333
		(0.724)	(0.235)	(0.534)	(0.423)	(0.320)	(0.354)
	Drinks	3.407	5.532	3.495	5.551	2.968	3.300
		(0.395)	(0.282)	(0.320)	(0.301)	(0.168)	(0.203)
	Large	4.841	4.985	3.630	7.787	4.157	5.667
		(0.396)	(0.299)	(0.248)	(0.412)	(0.171)	(0.304)
	Cola	4.931	5.358	3.660	7.535	7.191	7.472
		(0.398)	(0.294)	(0.300)	(0.418)	(0.284)	(0.346)
	Lemonade	3.377	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	17.282	3.295	4.448	8.997	3.006	2.728
		(2.497)	(0.513)	(0.613)	(0.776)	(0.304)	(0.381)
	Flavored milk	5.673	2.251	9.466	4.637	4.140	2.727
		(1.017)	(0.485)	(1.097)	(0.919)	(0.556)	(0.503)
Covariance	Sugary-Drinks	-4.420	-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.

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	-2.164	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

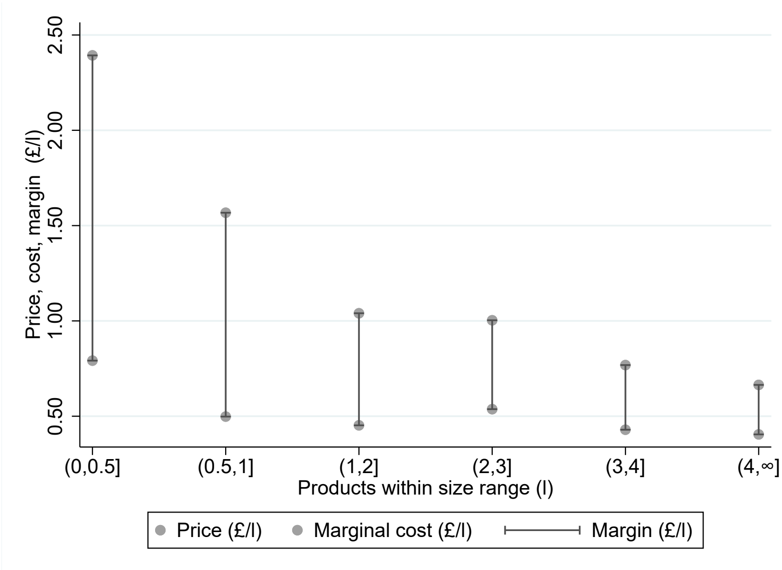
Notes: Numbers show the mean price elasticities of market demand in the most recent year covered by our data (2012). Number shows price elasticity of demand for option in column 1 with respect to the price of option in row 1.

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
			[0.30, 0.39]	[0.74, 0.83]	[0.57, 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	1.79	0.56
			[1.42, 1.69]	[1.65, 1.92]	[0.54, 0.59]
	Schweppes Lemonade	0.52	0.14	0.38	0.71
			[0.13, 0.16]	[0.36, 0.39]	[0.69, 0.73]
	Fanta	1.44	0.41	1.03	0.70
			[0.34, 0.49]	[0.95, 1.10]	[0.66, 0.73]
	Dr Pepper	1.33	0.42	0.91	0.66
			[0.38, 0.49]	[0.84, 0.96]	[0.63, 0.69]
	Schweppes Tonic	1.65	0.72	0.93	0.64
			[0.66, 0.77]	[0.88, 0.99]	[0.61, 0.68]
	Sprite	1.26	0.34	0.92	0.72
			[0.28, 0.39]	[0.87, 0.98]	[0.69, 0.75]
	Cherry Coke	1.53	0.50	1.04	0.62
			[0.43, 0.58]	[0.95, 1.11]	[0.58, 0.66]
	Oasis	2.31	0.53	1.79	0.77
			[0.36, 0.73]	[1.58, 1.95]	[0.69, 0.84]
Pepsico/Britvic	Robinsons	1.20	0.32	0.88	0.74
			[0.30, 0.34]	[0.85, 0.90]	[0.72, 0.76]
	Pepsi	1.02	0.40	0.62	0.61
			[0.37, 0.43]	[0.58, 0.65]	[0.59, 0.64]
	Tropicana fruit juice	2.20	0.94	1.25	0.56
			[0.86, 1.02]	[1.18, 1.34]	[0.53, 0.60]
	Robinsons Fruit Shoot	1.81	0.63	1.18	0.64
			[0.57, 0.68]	[1.12, 1.23]	[0.62, 0.66]
	Britvic fruit juice	2.06	0.90	1.16	0.56
			[0.85, 0.95]	[1.11, 1.21]	[0.54, 0.59]
	7 Up	1.22	0.45	0.77	0.68
			[0.40, 0.50]	[0.72, 0.82]	[0.64, 0.70]
	Copella fruit juice	1.40	0.23	1.17	0.83
			[0.19, 0.27]	[1.13, 1.21]	[0.80, 0.85]
	Tango	1.13	0.34	0.79	0.74
			[0.30, 0.39]	[0.74, 0.83]	[0.71, 0.77]
GSK	Ribena	1.77	0.89	0.88	0.50
			[0.85, 0.94]	[0.83, 0.92]	[0.47, 0.52]
	Lucozade	1.62	0.77	0.85	0.53
			[0.72, 0.84]	[0.78, 0.89]	[0.50, 0.55]
	Lucozade Sport	1.49	0.83	0.65	0.44
			[0.80, 0.88]	[0.61, 0.69]	[0.42, 0.46]
	Vimto	1.09	0.50	0.59	0.57
			[0.48, 0.51]	[0.58, 0.61]	[0.55, 0.58]
	Irn Bru	1.56	0.63	0.93	0.61
			[0.54, 0.72]	[0.84, 1.02]	[0.57, 0.66]
Merrydown	Shloer	1.59	0.71	0.88	0.55
			[0.68, 0.74]	[0.85, 0.91]	[0.54, 0.57]
Red Bull	Red Bull	4.74	2.60	2.15	0.44
			[2.31, 2.88]	[1.87, 2.43]	[0.38, 0.49]
Total		1.44	0.55	0.89	0.62
			[0.51, 0.61]	[0.83, 0.93]	[0.59, 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).

E 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).[63] 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

[63]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.

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 was some reformulation in anticipation of the introduction of the SDIL. We categorize products based on the post reformulation sugar contents.[64]

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:

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

where p_{jrt} denotes the price per liter of product j in retailer r in week t,[65] τ_t are week effects, ξ_j are product fixed effects, and ρ_r are retailer fixed effects.

Figure E.1(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 E.1(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 E.1(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.

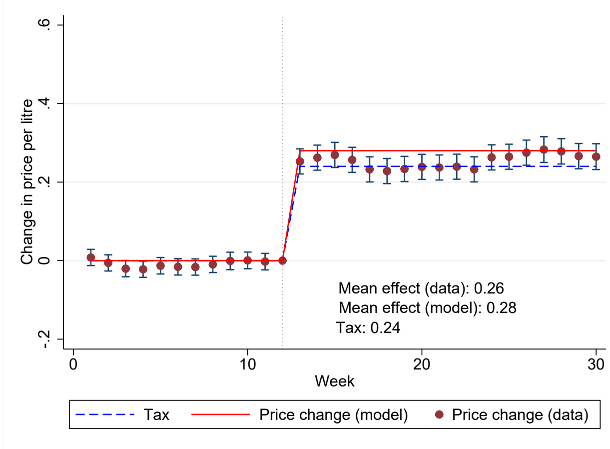
We simulate the introduction of the SDIL using our estimated model of demand and supply in the non-alcoholic drinks market (based on product sugar contents when the SDIL was implemented). The red lines plot the average price increase for each of the three group predicted by our model. These match very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

[64] 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.

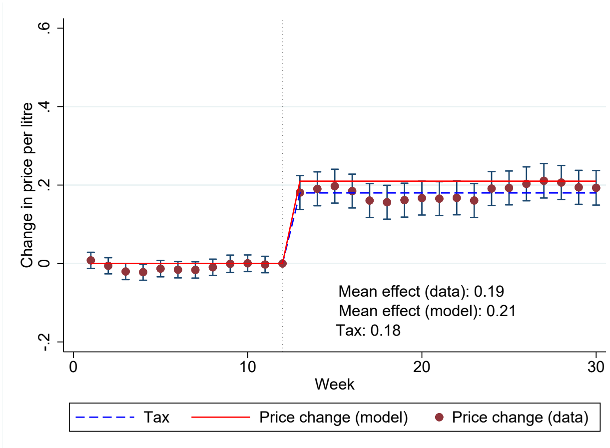
[65] This is the VAT-exclusive price per liter.

Figure E.1: *Out of sample model validation: UK Soft Drinks Industry Levy*

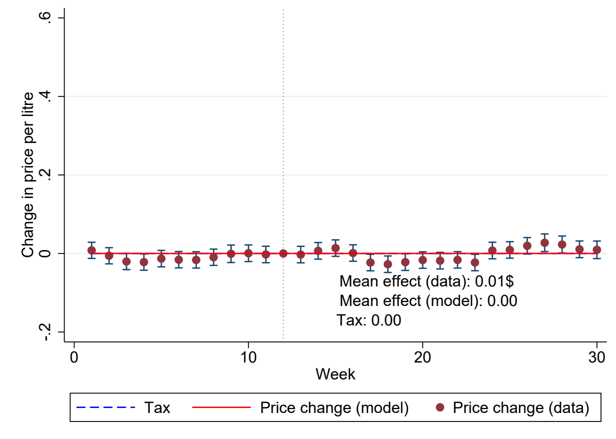
(a) High treatment group



(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 (E.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.*

F Implementation of optimal tax problem

F.1 Externality calibration

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.[66] 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 assume this represents an externality.

Finally, 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.

F.2 Distribution of profits

In Section 6.2 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

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

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 effective average corporate tax rate in 2012 (estimated by the University of Oxford's Centre for Business Taxation (Bilicka and Devereux (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), which records dividend income received by individuals to estimate the relationship between individual's total income and the amount they receive from dividends. Table F.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 F.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 F.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 F.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 F.2: *Distribution of dividends across household equivalized income distribution*

(1)	(2)	(3)	(4)	(5)	(6)
Equivalized		Mean	ATR	% dividends	
hh income	% hh	div income	divs	Pre-tax	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 F.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 F.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.

F.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 (2.3), 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 multi-tax rate system subject to constraints (see Section 6.3 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 solution of 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 6.2.

G A sugar tax with input substitution

In Section 6.3 we 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. (2021). In their model a sugary product's marginal cost comprises: the cost of the sugar in the product, the cost of a substitute input for sugar, and other components. The firm chooses the cost minimizing mix of sugar and the substitute input subject to keeping the taste of the product unchanged. Changing the product's sugar content therefore alters the cost of production and the firm's tax liability (if a sugar tax is in place), but it does not change consumers' valuation of the product.

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). The firm can choose to deviate from z_j^* while keeping the taste of product j unchanged, but this is entails increasing the product's production cost.

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

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

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

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

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

By definition, the sugar contents that satisfy these conditions are $z_j = z_j^*$ for all $j \in \mathcal{J}_f^{\mathcal{S}}$ 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 $j \in \mathcal{J}_f^{\mathcal{S}}$ and f. The first order conditions that characterize the firms' optimal choices are then: for $f = 1, \ldots, F$

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

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

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 firms' 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:

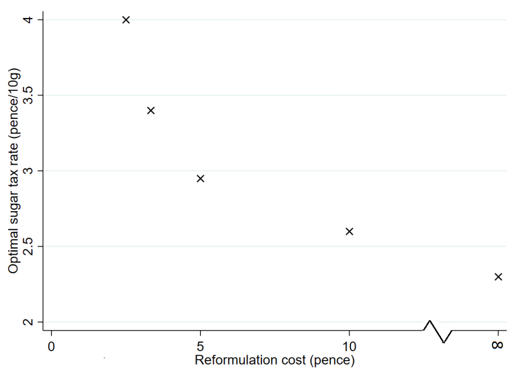
$$\frac{\Delta C_k(\nu, \tau)}{\Delta C_j(\nu, \tau)} = \frac{z_k^*}{z_j^*}.$$

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.

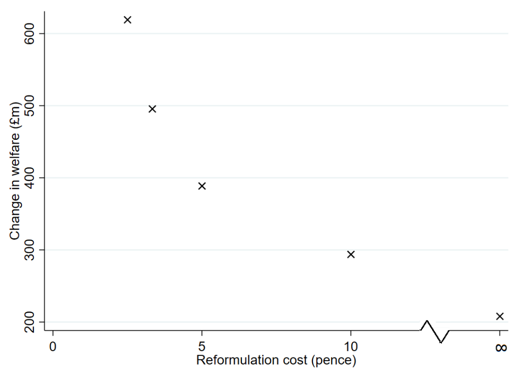
In Figure 6.2 in the paper we compare the implications of a sugar tax when reformulation costs are prohibitive and when they are relatively low (ν=5 pence). Instead of plotting how welfare varies with the tax rate, τ, we plot how it varies with $\frac{\Delta C}{z^*}$ (which is a monotonically increasing 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 G.1 we plot how the optimal sugar tax rate and associated welfare gain vary with the reformulation costs. As the cost of reformulation falls, firms chose 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 G.1: *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.