Corrective tax design and market power

Martin O'Connell and Kate Smith*

September, 2020

Abstract

We study the design of corrective taxes when firms have market power, showing how optimal policy depends on the *relative* size of price-cost margins among externality generating goods and alternative products, and the degree of consumer switching across these product sets. We consider an application to sugar sweetened beverage taxation. We embed an empirical demand and supply model into an optimal tax framework and show that policy that ignores distortions from the exercise of market power leads to welfare gains only 30% of those under optimal policy; and policy that accounts for distortions from market power on externality generating goods, but ignores them for alternatives, results in only 70% of the welfare gains under optimal policy.

Keywords: externality, corrective tax, market power, oligopoly
JEL classification: D12, D43, D62, H21, H23, L13
Acknowledgements: The authors would like to thank participants at seminars at Chicago Booth, Michigan, Oxford, Warwick, UCL, UIUC, WUSTL and Wisconsin-Madison for many very helpful discussions. The authors would like to gratefully acknowledge financial support from the European Research Council (ERC) under ERC-2015-AdG-694822, the Economic and Social Research Council (ESRC) under the Centre for the Microeconomic Analysis of Public Policy (CPP), grant number RES-544-28-0001, and the British Academy under pf160093. Data supplied by TNS UK Limited. The use of TNS UK Ltd. data in this work does not imply the endorsement of TNS UK Ltd. in relation to the interpretation or analysis of the data. All errors and omissions remain the responsibility of the authors.

*Institute for Fiscal Studies and University College London. **Correspondence**: martin_o@ifs.org.uk and kate.smith@ifs.org.uk

1 Introduction

There is widespread and growing interest in the extent of market power across the economy, and the implications that this may have for policy design. Policymakers are also increasingly dealing with the challenges posed by externalities, for example, due to carbon emissions or "unhealthy" consumer behaviors. Buchanan (1969) first addressed the importance of market power in corrective tax design, arguing that, when output is already below the competitive level, policy aimed at full internalization of externalities is not appropriate. A small empirical literature investigates the trade-off between correcting for externalities and exacerbating distortions associated with the exercise of market power among externality generating products. However, corrective taxes often lead consumers to substitute towards alternative products that generate lower or zero externalities and are supplied non-competitively, e.g., diet drinks, e-cigarettes, clean(er) energy sources. Little attention has been paid to the importance of market power among alternative goods for tax policy.

The contribution of this paper is to detail the implications for corrective tax design of distortions from the exercise of market power across *all* (both externality generating and alternative) products, and undertake an empirically rich and topical application to quantify the welfare losses of failing to take these distortions into account. We use a simple tax design framework to characterize optimal policy, and show how relative market power across externality generating products and alternatives drives optimal policy. We combine this framework with a detailed empirical model of consumer and firm behavior to study sugar sweetened beverage taxation. We show that tax policy that completely ignores distortions from the exercise of market power leads to only 30% of the welfare gains of optimal policy. Policy that accounts for distortions from market power on externality generating goods, but ignores them for alternative products, results in only 70% of the gains under optimal policy.

Our tax design framework consists of a planner that sets a tax rate on a particular product attribute associated with externalities (e.g. sugar, pure alcohol, tobacco or carbon) to maximize social welfare. The attribute is available from a set of differentiated products, which are supplied non-competitively – for expositional ease, we refer to these as "sin" products. There is also a set of alternative products that do not create externalities, but are possible substitutes for the sin products and are also supplied non-competitively.

The optimal tax rate trades off distortions from externalities with those arising from the exercise of market power. All else equal, the higher are price-cost margins of the sin products the lower is the optimal tax rate (echoing Buchanan (1969)).

Importantly, however, positive margins for alternative products also impact the optimal tax rate; the higher these are, and the more willing consumers are to switch towards these goods in response to the tax, the higher is the optimal tax rate. Therefore, the market power adjustment to externality correcting tax policy depends on the relative margins of the sin and alternative products. If there exist distortions from the exercise of market power among sin products but these products are supplied relatively competitively compared to alternatives, and if switching between the two product sets is strong, then the optimal tax rate accounting for market power can *exceed* the rate that would be optimal under perfect competition. In addition, the welfare loss from ignoring market power only among alternative goods can be greater than the loss from ignoring it altogether.

Whether, and to what extent, these considerations matter in practice is an empirical question.[1] We explore their relevance 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 (Allcott et al. (2019b)). In recent years, motivated by public health concerns, a number of countries and localities have introduced taxes on these products.[2] The drinks market is characterized by large multi-product firms that offer strongly branded products and are likely to enjoy market power. In addition, a substantial part of the market comprises zero sugar products, which are not associated with sugar-based externalities, but are substitutes for sugary drinks.

We specify a model of demand for the differentiated products in the market and firm price competition, in the tradition of Berry et al. (1995) and Nevo (2001). This allows us to estimate the consumer switching patterns and price-cost margins necessary to compute the optimal tax rate. We exploit micro level data from the UK on purchases of non-alcoholic drinks. Our data set has a number of strengths that help us obtain credible estimates. First, it is longitudinal, which helps us identify a rich distribution of consumer preferences (see Berry and Haile (2010)) and to capture how preferences vary across a long-run measure of the total sugar in a consumers' diet, an important dimension over which externalities are likely to vary. Second, it includes information on purchases made for at-home and on-the-go

[1] Oates and Strassmann (1984), for instance, argue that in the case of pollution control, Buchanan's (1969) critique of Pigouvian policy is not of great practical importance because the distortions from externalities are much larger than those arising from non-competitive behavior.

[2] As of December 2019, 44 countries and 8 US cities had some form of sugar sweetened beverage tax (GFRP (2019)). Griffith et al. (2019) surveys the literature evaluating the effects of these taxes.

consumption. Most empirical studies of sugar sweetened beverage demand do not cover purchases made on-the-go, and yet it is an important source of dietary sugar.[3]

Our demand estimates show that consumers respond to product level price changes by switching most strongly to similar products, e.g., from one sugary drink to another of the same size and flavor. However, when the price of all sugar sweetened beverages rise, there is considerable switching to alternatives: in response to a 1% rise in price, 33% of the reduction in volume of sugar sweetened drinks is diverted to alternative drinks. The supply estimates indicate a substantial degree of market power, both for sugar sweetened beverages and for alternative drinks; among both sets of products, price is, on average, around double marginal cost.

We use the demand and supply estimates to solve for the optimal tax rate on the sugar in sweetened beverages, incorporating the equilibrium pricing response of firms. We find that firms increase prices by slightly more than the tax. To validate these estimates, we compare the model's out-of-sample prediction of pass-through of the UK's Soft Drinks Industry Levy with data on how prices changed when this policy was introduced. The model's prediction of slightly more than 100% pass-through is in line with the observed price changes, and with findings in other jurisdictions.[4] This margin adjustment affects both the extent to which tax policy lowers externalities and its impact on the distortions from market power.[5]

We find that the optimal tax rate is less than half the rate a planner would set if it ignored distortions from market power completely. Conversely, the optimal tax rate is over double the rate a planner would set if it accounts for distortions from market power only for the sin goods. This undershooting is due to the significant degree of market power for alternative products coupled with a considerable willingness of consumers to switch towards them when the prices of the sugar sweetened beverages rise. Ignoring all distortions from market power results in substantial unrealized welfare gains. However, accounting for the market power of sin goods, but ignoring it for alternative products, also leads to smaller welfare gains, which are only 70% of those of the optimal policy. These results add to the growing body of evidence of the significant costs of policy naivety in imperfectly competitive markets.[6]

Our precise quantitative results depend on the magnitude and nature of externalities. We calibrate the externality function using evidence from the epidemiological

[3] An exception is Dubois et al. (2020), who focus on modeling on-the-go demand for drinks in the UK market and consider which consumer groups would change their consumption most strongly in the face of price rises.

[4] See, for example, Seiler et al. (2019) for Philadelphia and Grogger (2017) for Mexico.

[5] This point is also made by Preonas (2019) who estimates that mark-up adjustment by coal plants would lead to undershifting of a carbon tax.

[6] For example, in revenue-raising (**?**) and subsidy design (e.g., Polyakova and Ryan (2019)).

literature, but there is considerable uncertainty over this. It is important therefore to consider the sensitivity of our findings to plausible variation in externalities. We show that varying the scale of externalities, their degree of convexity with dietary sugar intake, and whether externalities arise from untaxed sugar sources does not alter our central point: failing to account for distortions associated with the exercise of market power for sin *and alternative* products leads to substantial unrealized welfare gains, which are similar in magnitude to our baseline calibration.

The welfare gains from optimally taxing externalities, as well as the costs of policy naivety, depend on the market structure. We explore the interaction between tax policy and competition by considering a set of counterfactual product ownership structures. When the drinks market comprises single product firms, and is therefore more competitive than in reality, welfare rises even in the absence of tax, the gains from optimal policy are larger, and the costs of ignoring market power either for all products, or only among alternative products, are less. The converse is true under joint profit maximization. Competition and tax policy therefore exhibit a form of complementarity. When sin goods are relatively competitively supplied, but alternative products are comparatively non-competitively supplied, welfare gains from optimal tax policy are large. In this case, the relatively low equilibrium margins for sin products means there is considerable scope for policy to raise prices without exacerbating market power distortions relative to alternative products.

Our work complements several recent papers that consider various aspects of the design of corrective taxes and apply them empirically to topical policy questions. Fowlie et al. (2016) study policy to reduce carbon emissions in the US cement market: in an empirical evaluation of the trade-off highlighted by Buchanan (1969), they show that policy that aims for full carbon abatement is welfare reducing and inferior to policy that accounts for the market power distortions. Relative to their paper, we highlight the important role played by market power among untaxed alternative goods. Like us, Allcott et al. (2019a) study sugar sweetened beverage taxation, but they focus on how a planner's redistributive motives and the existence of internalities affect optimal policy. Griffith et al. (2019) quantify the welfare gains that can be realized from varying tax rates across different sources of an externality when consumers have heterogeneous preferences and marginal externalities. We complement the insights of these papers, both of which assume a perfectly competitive environment, by extending the results on optimal corrective tax design to settings characterized by market power and quantifying their importance for policy.

The rest of this paper is structured as follows. In Section 2 we outline our tax design framework. Section 3 describes the UK market for drinks and the micro data that we use. In Section 4 we present our empirical model of consumer demand and firm pricing competition. Section 5 presents our results. A final section draws together the implications of our results and concludes.

2 Tax design framework

Our aim is to highlight how distortions associated with the exercise of market power influence the welfare maximizing rate of tax on externality generating products. We study a market that comprises a set of differentiated products, a subset of which have externalities associated with their use. The products are provided by firms that set their prices under conditions of imperfect competition. We consider a social planner with access to a linear tax that it can set on the externality generating products,[7] and whose aim is to maximize social welfare.[8]

2.1 Set-up

Consider a market that comprises many products $j = \{1, \ldots, J\}$. A subset of products, $j \in \mathcal{S}$, contain an attribute that is associated with an externality, where we denote by z_j the amount of the attribute in product j, while for the remaining products, $j \notin \mathcal{S}$ (which we denote by the set $j \in \mathcal{N}$), $z_j = 0$. For expositional ease we refer to products in set $\mathcal{S}$ as "sin" products and those in $\mathcal{N}$ as alternative products; it should be understood that sin products simply refer to those with a positive amount of attribute z. We focus on externalities, but our framework can accommodate certain forms of internalities; see Appendix A.

Consumer i, facing prices, $\mathbf{p} = (p_1, \ldots, p_J)$, chooses how to allocate her income, y_i, between the J products and a numeraire good, which represents expenditure outside of the market of interest. We assume consumers have preferences that are quasi-linear and can be represented by the indirect utility function $V_i(\mathbf{p}, y_i) = y_i + v_i(\mathbf{p})$, and denote consumer level demand for product j by $q_{ij}(\mathbf{p})$. The quasi-linear preference structure means that a price change for a product does not induce

[7] Our focus on a linear tax schedule is in keeping with the vast majority of the optimal tax literature, which restricts attention to linear commodity taxes due the practical impediments of implementing non-linear commodity taxes.

[8] The planner does not have a redistributive *motive* when setting the tax rate, though in the empirical implementation we can describe the redistributive *effect* of the policy, which can be used to help design distribution-neutral adjustments to the income tax system to offset the redistributive effects of the corrective tax across the income distribution (Kaplow (2012)). Under perfect competition, if the planner can set a non-linear labor tax, redistributive motives do not influence optimal commodity taxes as long as differences in consumption patterns across the income distribution are driven purely by income differences (Saez (2002)). Jaravel and Olivi (2019) show that this result extends to an economy characterized by imperfect competition.

any income effects. This assumption is reasonable when focusing on a market that accounts for a modest share of total consumer spending.[9] We denote consumer level demand for the sets of sin and alternative products by $Q_i^{\mathcal{X}}(\mathbf{p}) \equiv \sum_{j\in\mathcal{X}} q_{ij}(\mathbf{p})$ for $\mathcal{X} = \{\mathcal{S}, \mathcal{N}\}$. We denote market level demand for product j by $q_j(\mathbf{p}) \equiv \sum_i q_{ij}(\mathbf{p})$ and for product set $\mathcal{X}$ by $Q^{\mathcal{X}}(\mathbf{p}) \equiv \sum_i Q_i^{\mathcal{X}}(\mathbf{p})$. We denote total consumption by consumer i of the externality generating attribute by $Z_i \equiv \sum_j z_j q_{ij}(\mathbf{p})$.

Consider a tax levied in proportion to z_j. We assume that the products are supplied imperfectly competitively. We do not impose any restriction on the nature of competition at this stage, specifying only that equilibrium prices satisfy:

$$p_j - \tau z_j - c_j = \mu_j \quad \forall\, j, \tag{2.1}$$

where μ_j denotes product j's price-cost margin. Equilibrium prices and margins depend on the tax rate, the marginal costs of all products, as well as the market structure. In our empirical application we specify and estimate a model of firm competition, which we validate using quasi-experimental evidence on how margins change following a tax increase. We assume that the numeraire is competitively supplied, and its consumption does not generate any externalities. It is straightforward to allow for a numeraire good that is non-competitively supplied and/or creates externalities; we do this in Section 5.5.

2.2 Optimal policy

We consider a social planner that chooses the rate of tax to maximize total welfare:

$$\max_{\tau} v(\mathbf{p}(\tau)) - \sum_i \Phi_i(Z_i(\tau)) + \sum_j (p_j(\tau) - c_j) q_j(\tau), \tag{2.2}$$

where $v(\mathbf{p}) \equiv \sum_i v_i(\mathbf{p})$ denotes total consumer surplus from participation in the market, $\sum_i \Phi_i(Z_i)$ denotes the total externality cost, and the final term denotes the tax inclusive profits (tax revenue plus firm profits). To highlight the dependency of the market equilibrium on tax policy, we write equilibrium prices, quantities and margins as a function of the tax rate.

[9]In general, the own price effect on demand for good j follows the Slutsky equation $\epsilon_{ij} = \epsilon_{ij}^h + \frac{p_j q_{ij}}{y} e_{ij}$, where ϵ_{ij} and ϵ_{ij}^h 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 $\epsilon_{ij} \approx \epsilon_{ij}^h$ and preferences are approximately quasi-linear.

Rearranging the associated first order condition gives the following implicit formula for the welfare maximizing tax rate, τ^*:

$$\tau^* = \widetilde{\Phi}(\tau^*) - \widetilde{\mu}^S(\tau^*) + \widetilde{\mu}^N(\tau^*) \times \widetilde{\Theta}(\tau^*), \tag{2.3}$$

where (see Appendix A for the derivation):

$$\widetilde{\Phi}(\tau^*) = \sum_i \phi_i(\tau^*)\omega_i(\tau^*), \qquad \widetilde{\mu}^S(\tau^*) = \tfrac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)} \sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}(\tau^*)\mu_j(\tau^*),$$
$$\widetilde{\Theta}(\tau^*) = -\tfrac{Q'^{\mathcal{N}}(\tau^*)}{Q'^{\mathcal{S}}(\tau^*)}, \qquad \widetilde{\mu}^N(\tau^*) = \tfrac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)} \sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}(\tau^*)\mu_j(\tau^*).$$

The first component, $\widetilde{\Phi}(\tau)$, reflects distortions from externalities; it depends on consumer-specific marginal externalities, $\phi_i(\tau) \equiv \Phi_i'(Z_i(\tau))$, and each consumer's contribution to the marginal impact of tax on equilibrium consumption of attribute z, $\omega_i(\tau) \equiv \frac{Z_i'(\tau)}{\sum_i Z_i'(\tau)}$. The second component, $\widetilde{\mu}^S(\tau)$, reflects distortions due to market power among the set of sin products. It depends on the weighted average equilibrium price-cost margins of the sin products, where the weights are given by $w_j^{\mathcal{S}}(\tau) \equiv \frac{q_j'(\tau)}{Q'^{\mathcal{S}}(\tau)}$, adjusted for the marginal impact of tax on the equilibrium quantity of sin product relative to its impact on the equilibrium quantity of attribute z, $\frac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)}$. The third component is the product of two terms. $\widetilde{\mu}^N(\tau)$ reflects distortions due to market power for the set of alternative products. It takes a similar form to the second term, depending on the weighted average equilibrium margins among the set of alternative products, with the same adjustment for the impact of a marginal tax change on the quantity of sin products demanded relative to the equilibrium quantity of attribute z. $\widetilde{\Theta}(\tau)$ captures how strongly, in equilibrium, consumption switches from the sin to alternative products in response to a marginal increase in the tax rate.

Intuition in a simplified two-product case

To highlight the intuition behind the various forces that determine the optimal rate, τ^*, it is useful to consider a simplified two-good case, in which there is one sin product and one alternative product (specifically, $\mathcal{S} = \{s\}$ and $\mathcal{N} = \{n\}$). Suppose also that externalities arise directly through consumption of product s, so externalities for consumer i are given by $\Phi_i(q_{is}(\tau))$, and that the externality function is common across consumers and linear, so we can write the marginal externality as $\phi_i(\tau) = \phi$. In this case the formula for the optimal rate of tax on the externality generating good simplifies to:

$$\tau^* = \phi - \mu_s(\tau^*) + \mu_n(\tau^*) \times \tfrac{q_n'(\tau^*)}{-q_s'(\tau^*)}. \tag{2.4}$$

When the two products are supplied competitively (so $\mu_j(\tau) = 0$ for $j = \{s, n\}$ regardless of the level of τ) the optimal policy is a Pigouvian tax ($\tau^* = \phi$) and the first best is achieved (Pigou (1920)).

If the sin good is non-competitively supplied, but either the alternative good is competitively supplied (so $\mu_n(\tau^*) = 0$) or there is zero switching in equilibrium between the two-goods (so $q_n'(\tau^*) = 0$) the optimal tax is implicitly defined by $\tau^* = \phi - \mu_s(\tau^*)$. The tax rate takes account of two off-setting distortions: excess consumption due to externalities and insufficient consumption associated with the exercise of market power for the sin product. This is what is highlighted in Buchanan (1969) and the literature that has followed. In this case product s is priced at the efficient level, $p_s = c_s + \phi$, and the equilibrium price of product n is left unaffected by the tax.

However, if taxing product s also leads consumers to adjust their consumption of some other non-competitively supplied product, then distortions arising from the exercise of market power on this alternative product impact the welfare maximizing tax rate. This is captured by the third term in equation (2.4). If, in equilibrium, the two goods are substitutes then $\frac{q_n'(\tau^*)}{-q_s'(\tau^*)} > 0$. This means that a positive margin for good n, all else equal, raises the optimal tax rate. Conversely, if the two goods are complements, a higher margin for product n acts to lower the optimal tax rate.

In a hypothetical case in which total consumption across the two products is fixed i.e. in response to price changes consumers only reallocate their demand between the two products, so $\frac{q_n'(\tau^*)}{-q_s'(\tau^*)} = 1$, the optimal tax rate is $\tau^* = \phi - (\mu_s(\tau^*) - \mu_n(\tau^*))$. In this case, the Pigouvian policy prescription is adjusted by the difference in equilibrium margins between the two goods. The difference in equilibrium prices of the two products is $p_s - p_n = (c_s - c_n) + \phi$ and the tax achieves an efficient allocation (of the fixed consumption level) across the two products. If the difference in margins across the products is small, then the optimal rate will lie much closer to the Pigouvian rate than the rate implied by policy that accounts for the margin only on the sin good, but that ignores the distortion from market power associated with the alternative good.

More generally, when products s and n are substitutes, the stronger in equilibrium that consumers switch between the two goods, the more weight that the optimal tax formula places on the margin of the alternative product and, all else equal, the higher will be the optimal tax rate.

Generalization to the multi-product setting

The intuition from the two-product case generalizes naturally to the many product case. With many products, the two components of the optimal tax formula that reflect distortions arising for the exercise of market power, $\widetilde{\mu}^S(\tau)$ and $\widetilde{\mu}^N(\tau)$, depend on the *weighted average* price-cost margin among the *sets of* sin and alternative products. As the tax rate varies, the average margin terms may vary for two reasons – (i) firms may reoptimize their prices, changing product level price-cost margins, and (ii) consumers, in equilibrium, may switch differentially away from or towards products with different equilibrium margins. In addition, with many products, the degree to which margins on alternative products raise the optimal rate (i.e. $\widetilde{\Theta}(\tau)$) depends on the extent to which any reduction in consumption of the *set of* sin products induced by a marginal increase in the tax rate is redirected towards the *set of* alternative goods.

Generalization to heterogeneous externalities

In the two-product example we assume that externalities are homogeneous across consumers and linear, which leads to the externality correcting component of the optimal tax formula being ϕ. In contrast, in equation (2.3) we allow for heterogeneous and non-linear externalities. This results in the externality correcting component, $\widetilde{\Phi}(\tau)$, being equal to a weighted average of consumer specific marginal externalities where the weight, $\omega_i(\tau)$, is the contribution of individual i to the marginal change in the equilibrium quantity of the externality generating attribute with respect to a marginal change in the tax rate. The more strongly those whose marginal consumption is most socially costly respond to the tax, then the more effective is the tax in correcting for externalities and, all else equal, the higher is the optimal rate.

The term $\widetilde{\Phi}(\tau)$ takes a similar form to the optimal externality correcting tax with heterogeneous externalities in a perfectly competitive market, derived in Diamond (1973). However, in an imperfectly competitive environment, the weights incorporate the equilibrium pricing response of firms in the market. This highlights that the apparent additive separability between externality and market power correction in equation (2.3) is misleading; a change in market power (for instance, through two firms merging) may impact both equilibrium margins and the externality correcting part of the formula through changing the extent to which tax will be passed through to consumer prices (and hence $Z_i'(\tau)$).

The quantitative importance of market power for corrective policy is ultimately an empirical question. We apply our framework to the taxation of sugar sweetened beverages. We estimate an equilibrium model of consumer demand and firm competition in the market for drinks, which allows us to simulate equilibrium price-cost margins and consumer switching, and we embed it into our tax design framework. We use evidence from the epidemiological literature to calibrate the externalities from sugar sweetened beverage consumption and show sensitivity of our results to this calibration. In the next section we describe the drinks market.

3 The drinks market

Sugar sweetened beverage taxation is a natural setting in which to study the interaction of corrective taxation and 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 – Allcott et al. (2019b) provides a useful survey of the evidence, which we discuss further in Section 5.4. It is also the case that the market is dominated by two large firms, Coca Cola Enterprises and PepsiCo, which are likely to exercise considerable market power. In addition, these firms produce zero- or low-sugar drinks that are likely to both be close substitutes for sugary varieties and to have positive price-cost margins.

We model behavior in the UK market for drinks. Our market definition includes a set of drinks typically subject to beverage taxes. These include carbonated drinks, fruit concentrates, sports and energy drinks, to which we refer collectively as soft drinks. Soft drinks can be divided into sugar sweetened beverages and diet (or artificially sweetened) beverages. We consider a sugar sweetened beverage tax, which applies to sugar, but not to artificially, sweetened beverages. Our market definition also includes pure fruit juices and sweetened milk. These are outside the scope of sugar sweetened beverage taxes but, like diet drinks, are likely to be substitutes for sugar sweetened beverages.

3.1 Purchase data

We use micro data on the drinks purchases of a sample of consumers living in Great Britain. The data contain information on household level purchases for "at-home" consumption, as well as purchases made by individuals for "on-the-go" consumption. Our data covers over 90% of drinks consumption by volume; the remainder occurs in restaurants and bars, which are not covered by our data. On-the-go consumption is an important part of the drinks market, accounting for around a third of total sugar consumption from drinks.[10] Our data are collected by the market research firm Kantar and comprise two parts: the Kantar Worldpanel covers the at-home segment and the Kantar On-The-Go Survey covers the on-the-go segment.

Our Kantar Worldpanel dataset contains details of all the grocery purchases (including food, drink, alcohol, toiletries, cleaning products and pet foods) that are made and brought into the home by a representative sample of British households from January 2008 to December 2012.[11] Participating households use a hand held scanner to record all purchases at the UPC (barcode) level. Households participate in the survey for several months, and the data contain detailed information on the UPCs they buy (including brand, flavor, size and nutrient composition), the store in which the purchase took place, and transaction level prices.

The Kantar On-The-Go Survey is based on a random sample of individuals drawn from the Worldpanel households. Using a cell phone app, individuals record purchases of food and drinks at the UPC level made on-the-go from shops and vending machines. The data contain details of the items they purchased, as well as transaction store and price, from June 2009 to December 2012. Individuals aged 13 and upwards are included in the sample.

3.2 Consumers

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.[12]

In our empirical demand model we incorporate observed and unobserved heterogeneity in consumer preferences. We allow observed heterogeneity across the at-home or on-the-go segments, as well as allowing preferences to vary by consumer groups, which are based on consumer age and a measure of the total sugar in the consumer's diet in the preceding year. This allows us to capture differences in demand behavior along important dimensions over which marginal externalities from sugar sweetened beverage intake might vary.

Table 3.1 shows the consumer groups. In the at-home segment we split households based on whether there are any children in the household. In the on-the-go

[10] Statistics are based on our calculations using the National Diet and Nutrition Survey, an individual level dietary intake survey representative of the UK population.

[11] In Appendix B we compare the Kantar Worldpanel across key demographic dimensions with the nationally representative Living Cost and Food Survey.

[12] 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.

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 3.1: *Consumer groups*

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

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

3.3 Products

Firms and brands. In Table 3.2 we list the main firms that operate in the drinks market and the brands that they own. We focus on the principal brands in the market; these comprise over 75% of total spending on non-alcoholic drinks in both the at-home and the on-the-go segments.[13] The firms Coca Cola Enterprises and PepsiCo/Britvic dominate the market, having a combined market share exceeding

[13]This include all soft drinks brands with more than 1% market share in either segment, as well as the main fruit juice and flavored milk brands. For some brands, there are only a very small

65% in the at-home segment and close to 80% in the on-the-go segment. Each of these firms owns several well recognized and long established brands, including some soft drinks and fruit juice brands. The most popular single brand is Coke (or Coca Cola), which accounts for over 20% of the at-home and 36% of the on-the-go market segment. In addition to the main branded products, we include store brands (also known as private labels) in our analysis; these are popular in the at-home segment.

Table 3.2: *Firms and brands*

			Market share (%)		Price (£/l)	
Firm	Brand	Type	At-home	On-the-go	At-home	On-the-go
Coca Cola Enterprises			*33.0*	*59.0*		
	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
Pepsico/Britvic			*33.6*	*20.0*		
	Robinsons	Soft	10.7	–	1.09	–
	Pepsi	Soft	10.1	11.6	0.64	1.93
	Tropicana fruit juice	Fruit	6.1	3.8	1.62	3.63
	Robinsons Fruit Shoot	Soft	2.7	0.8	1.49	2.83
	Britvic fruit juice	Fruit	1.6	–	2.17	–
	7 Up	Soft	0.9	1.7	0.70	1.88
	Copella fruit juice	Fruit	0.8	–	1.68	–
	Tango	Soft	0.8	2.2	0.66	1.73
GSK			*7.6*	*12.7*		
	Ribena	Soft	3.3	3.4	1.69	2.20
	Lucozade	Soft	3.1	6.4	1.11	2.37
	Lucozade Sport	Soft	1.2	2.9	1.15	2.22
JN Nichols	Vimto	Soft	1.6	–	1.06	–
Barrs	Irn Bru	Soft	0.6	2.6	0.61	1.93
Merrydown	Shloer	Soft	2.0	–	1.79	–
Red Bull	Red Bull	Soft	0.2	3.4	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.

number of transactions in one of the two segments of the market; we therefore omit these brands from the choice sets in that segment.

Variants. The majority of soft drink brands are available in sugar sweetened ("regular") and artificially sweetened ("diet" and/or "zero") variants. In Table B.2 in Appendix B we list the variants available for each brand. Among the regular variants there is variation in sugar content across brands – many of the carbonates have around 10g of sugar per 100ml, with some of the fruit flavored soft drinks (such as Oasis and Vimto) having less sugar per 100ml.

Sizes and multi-buy promotions. 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 instance, a consumer may choose to purchase one 2l bottle of Diet Coke, or a pack of 6×330ml cans, or two 2l bottles, and so on. Purchases of multiple packs of the same brand-variant most commonly involve two, or sometimes three, units of the same pack (or UPC) and are typically a consequence of multi-buy offers. Multi-buy promotions in the UK market are long running, so the set of UPCs for which multiple units are popular is broadly stable over time.

Products. We define a product as a brand-variant-size combination. For each brand-variant, the set of possible sizes includes both the available pack sizes (i.e. UPCs) and the most common multiple pack purchases of UPCs.[14] Our product definition embeds some aggregation across very similar UPCs; for instance, across different flavours that have the same sugar contents and prices.[15] In Table B.2 in Appendix B we show, for each brand-variant, the number of sizes available to consumers in the at-home and on-the-go segments. For instance, Diet Coke is available in 10 sizes in the at-home segment, and two sizes in the on-the-go segment.[16] On-the-go sizes are always designed as a single serving, while at-home sizes are typically multi-portion.

3.4 Choice sets and price measurement

The set of products available to a consumer on a particular choice occasion, as well as the price vector they face, depend on the retailer that they visit.

[14] 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%) of transactions. This means that for over 75% of transactions of branded products, we accurately model the choice over number of units to purchase.

[15] 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".

[16] These are, in the at-home segment, 1.25l and 2l bottles, multi-packs of 330ml cans containing 6, 8, 10 and 12 cans, two- and three- unit purchases of 2l bottles, and two-unit purchases of 6-pack and 8-packs of cans; and, in the on-the-go segment, a 500ml bottle and 330ml can.

At-home segment

The median household undertakes a grocery shop once a week. We define a “choice occasion” as any week in which a household purchases groceries, and model what, if any, drink product a household purchases on a choice occasion. We observe households for an average of 36 choice occasions each year, and in total, we have data on 3.3 million at-home choice occasions. On around 42% of choice occasions, a household purchases a drink, with, on average, 12 days between drink purchases. Households select one brand-variant (as defined by columns (2) and (3) of Table B.2) on 60% of choice occasions on which drinks are purchased. On choice occasions in which a household chooses multiple options, we assume that, conditional on household specific preferences, these purchases are independent, e.g., because they are bought for different household members.

For each choice occasion we observe the retailer in which the purchase was made and the exact price paid. Table 3.3 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.

We do not model a consumer’s choice over which retailer to shop in; we assume this decision is driven by factors such as the proximity of nearby stores and overall preferences for grocery outlet (for which we control in demand), and not people shopping around stores to find those with a temporarily low price for a specific drinks product. In our setting this assumption is reasonable. On most choice occasions (i.e. weeks) consumers are observed purchasing groceries from one retailer,[17] while consumers allocate, on average, a small share (4%) of their total supermarket expenditure to non-alcoholic drinks.

The four main retailers in the UK implement national pricing policies, following a Competition Commission (2000) investigation into supermarket behavior.[18] This means that if we observe a transaction price for a particular UPC in a store belonging to one of the retailers, Tesco say, we know the price that consumers shopping in other Tesco stores at the same time faced for that UPC. Using the large number of transactions in our data we can construct the price vector households 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.

[17] Thomassen et al. (2017) highlight the role of fixed shopping costs in leading consumers to undertake their grocery shopping in one or a small number of stores.

[18] Close to uniform pricing within retail chains has been documented in the US; see, for instance, DellaVigna and Gentzkow (2019) and Hitsch et al. (2017).

Table 3.3: *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.

On-the-go segment

The natural periodicity for on-the-go purchases is at the daily level; we define a choice occasion as any day on which the individual buys a cold beverage, including bottled water. We observe individuals for an average of 44 choice occasions each year, and in total, we have data on 286,576 on-the-go choice occasions. On 60% of choice occasions individuals choose to buy one of the products listed in Table B.2, and on 90% of these choice occasions they buy only one product.[19]

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 (see Table 3.3). This, coupled with the fact that the single portion cans and bottles are similarly priced across the large four supermarkets, motivates their aggregation into one composite retailer. The majority of transactions in the on-the-go segment are in local convenience stores. This means that for these choice occasions, unlike in the at-home segment, we do not observe the price of non-selected products in consumers' choice sets. Therefore, in the case of convenience stores, for all options in consumer choice sets we use a mean monthly price, where the price is constructed using all convenience store transactions in each of three regions (the south, central, and north regions of the UK).

[19]On the rare case when they buy multiple products (usually 2 or 3) we treat these as independent purchases.

Dependence across the at-home and on-the-go segments

We model consumer choice for at-home consumption and for on-the-go consumption separately.[20] A potential concern is the existence of non-separabilities across the at-home and on-the-go segments, e.g., recent at-home household purchases influence decisions that individuals make on-the-go. 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. This is consistent with the findings in Dubois et al. (2020), who, following Browning and Meghir (1991), test for non-separabilities between the two segments and find no evidence of demand dependence.

3.5 Price variation

The vector of prices that a consumer faces when making a purchase varies across time and retailers. Here we describe this variation and in Section 4.2 we discuss how it allows us to identify the key parameters driving consumer demand behavior.

The at-home segment is characterized by products that are sold in multi-portion sizes, and it is dominated by retailers that have national pricing policies. An important source of price variation is promotions, which differ in their timing, duration and depth, both across UPCs and retailers. In our data, 30% of transactions are multi-buy offers (e.g. a discount for purchasing two of the same UPC), and 20% are ticket price reductions (when a UPC has a temporarily low price).

We provide a graphical example of each promotion type in Figure 3.1, which shows the price for two UPCs over the most recent year of our data for two retailers. Panel (a) shows price series for a 2l bottle of Coke. In both retailers, with the exception of one week in Tesco, one unit of a 2l bottle is priced at £2. However, over most of the year each retailer runs a multi-buy offer, where 2 bottles can be purchased at a discounted per bottle price, though the depth of discount varies both over time and across retailers.[21] Panel (b) shows price series for a pack of 12×330ml cans of Coke. This UPC does not have a multi-buy offer, but is reasonably frequently subject to a ticket price reduction.

In Figure 3.1 average prices are similar across the two retailers, but the time path of price changes is different. This is true more generally. To illustrate this we compute measures of price stability suggested by DellaVigna and Gentzkow (2019).

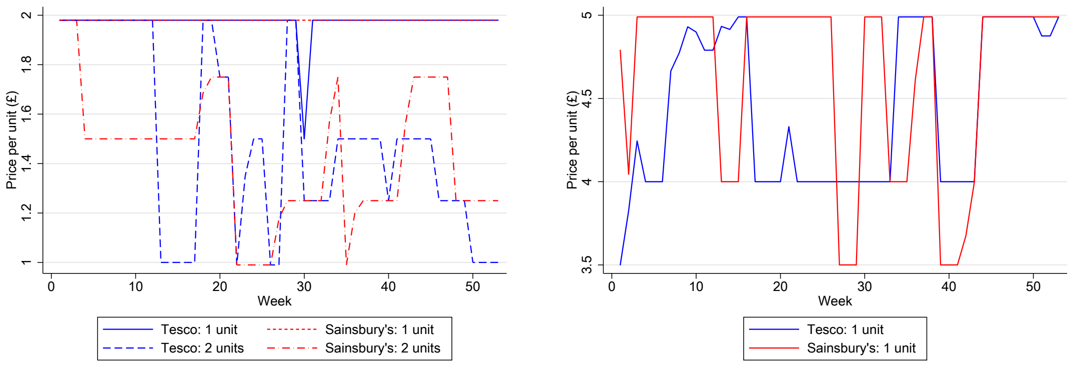
[20] When constructing market level demand we weight each segment so their expenditure shares match with those in the Britvic trade report (2014).

[21] In our demand model we treat one 2l bottle and two 2l bottles of Coke as different options, with distinct prices.

First, we compute the average log price for each product-retailer-year and then for each product-year compute the deviation in this for each retailer pair. The median deviation is 8 log points, indicating a relatively low level of cross-sectional differences in average prices across retailers. Second, we obtain the residuals from regressing log prices on product-retailer-year fixed effects and then for each product-year compute the correlation in residuals across each retailer pair. The median of these correlations is 0.13, indicating that the co-movement in prices over time across retailers is low. In addition, no retailer sets systematically low or high prices – among the big four retailers, Asda is the cheapest retailer the most (for 27% of product-weeks) and the most expensive the least (for 17%), and Sainsbury's is cheapest the least (for 22%) and most expensive the most (for 31%).

Figure 3.1: *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.

A concern with relying on price variation from promotions to estimate demand is that households respond to them by intertemporally switching their purchases (i.e stocking up during sales) and hence failing to model this behavior will result in an overestimate of own-price elasticities (Hendel and Nevo (2006a)). A number of papers have documented evidence of stockpiling in the US market for soft drinks (see Hendel and Nevo (2006b), Hendel and Nevo (2013), Wang (2015)).

Although we cannot rule out that there may be some stockpiling underlying transactions in our data, the evidence for it is much less clear than in the US. Specifically, UK households purchase soft drinks, on average, around twice as frequently (every 14 days) as those in the US (see Hendel and Nevo (2006b)), and when a household does purchase on sale there is no meaningful change in the timing of purchases.[22] Instead, we find that sales are associated with switching across pack types (e.g. cans to bottles), brands and sizes. One reason why stockpiling is less prevalent in the UK, compared with the US, is that the relatively long running nature of UK promotions create less incentives to stockpile. For instance, for each of the soft drinks products available in the at-home segment, the average time between a price change of 25% or more is 8 weeks, whereas in the US prices can fluctuate by large amounts from week to week (see an archetypal example in Figure 1 of Hendel and Nevo (2013)). A second reason is that transport and storage costs in the UK are likely to be much higher, with the average size of UK homes around half of those in US, and vehicle ownership rates 25% lower.[23]

In the on-the-go segment only 20% of spending is done in the large four supermarkets, with around 55% of expenditure occurring in conveniences stores. Promotions are less common in this segment, with price variation driven by regional differences in price in convenience stores, and variation in prices in convenience store relative to national retailers and vending machines.

3.6 Vertical relations

We do not observe the contracting relationship between the drinks firms (listed in Table 3.2) and the retailers. If drinks firms and retailers are restricted to linear contracts, the retail prices will likely reflect double marginalization.[24] In contrast, with non-linear contracts the contracting parties can avoid double marginalization, with retail margins defined relative to product costs, rather than wholesale prices. A 2013 report into the soft drinks market by the UK competition authority provides evidence that vertical relations are characterized by non-linear contracting (see Competition Commission (2013)). They cite evidence that at a regular bilateral "Joint Business Planning" meeting a soft drinks firm and retailer agree on wholesale prices, payments related to product visibility, recommended retail prices, and agreements on the number, type and timings of promotions. This evidence motivates our treatment of soft drinks firms setting final consumer prices, an outcome consistent with optimal non-linear contracts – see Villas-Boas (2007) and Bonnet and Dubois (2010).

[22] Hendel and Nevo (2006b) find, in the US, buying 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. We find changes of 0.23 and 0.14 respectively. See Appendix C.

[23] 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)). In 2014 the US had 816 vehicles per capita (U.S. Department of Energy (2019)), in 2017 the UK had 616 (ACEA (2019)).

[24] Specifically, manufacturers would maximize their profits by choosing their wholesale price, and retailers would maximize their profits taking wholesale prices are given.

4 Equilibrium model of the drinks market

To implement the tax design framework outlined in Section 2, we need to know how consumers switch across products in response to price changes, the level of price-cost margins on these products and how firms, in response to tax, adjust them. 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 approach enables us to model demand and substitution patterns over the many differentiated products in the market, while incorporating rich preference heterogeneity crucial to capturing realistic substitution patterns. We identify firms' unobserved marginal costs by coupling our demand estimates with the equilibrium conditions from an oligopoly pricing game (Berry (1994), Nevo (2001)).

4.1 Consumer demand

We model which, if any, drink product a consumer (indexed i) chooses on a choice occasion. 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 with each type of choice situation, but for notational parsimony we suppress a market segment index.

We index the drinks products by $j = \{1, \ldots, J\}$. The products vary by brand, which we index by $b = \{1, \ldots, B\}$, size, indexed by $s = \{1, \ldots, S\}$, and whether or not they contain sugar (for instance, the brand Coke is available in Regular, Diet and Zero variants). The consumer chooses between the available drinks products, and choosing not to buy a drink, which we denote by $j = 0$. The set of products available to the consumer, as well as the prices they face, depends on which retailer they visit – we index retailers by r and denote the set of available drink options in retailer r by Ω_r.

Consumer i in period t, with total period income or 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}.

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

Let $d = (1, \ldots, D)$ index the consumer groups shown in Table 3.1. We assume 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:

$$\nu(.) = -\alpha_i p_{jrt} + \boldsymbol{\beta}_i \widetilde{\mathbf{x}}_j^{(1)} + \gamma_{d(i)} \widetilde{x}_{jt}^{(2)} + \zeta_{d(i)b(j)s(j)rt},$$

where

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

We allow for consumer specific preferences for price (i.e. the marginal utility of income) and a subset of product characteristics denoted by $\widetilde{\mathbf{x}}_j^{(1)}$, which includes a constant to capture a preference for drinks versus not buying them, dummy variables indicating whether the product has strictly positive but less than 10g sugar per 100ml or weakly more than 10g per 100ml, dummy variables indicating if the product is a cola, lemonade, store brand soft drink or fruit juice, and an indicator for whether the size is large.[25] These individual level preferences play a key role in allowing the model to capture realistic substitution patterns across

[25]Defined as larger than 2l in the at-home segment or 500ml in the on-the-go segment.

products. $\widetilde{x}_{jt}^{(2)}$ is a measure of the stock of advertising for the product in the current period;[26] we allow the effect of advertising to vary across consumer groups.

$\zeta_{d(i)b(j)s(j)rt}$ denotes a set of consumer group specific shocks to utility. These include: brand-size effects, which control for unobserved consumer preferences that are time-invariant; brand- and size-retailer effects, which capture the possibility that, on average, consumer preferences over brand and size differ across retailers; and brand- and size-time effects, that control for shocks to demands through time.

We model the consumer specific preferences, $(\alpha_i, \boldsymbol{\beta}_i)$ as random coefficients. We specify the distribution for α_i as log-normal and $\boldsymbol{\beta}_i$ as normal, both conditional on consumer group d. The overall random coefficient distribution is a mixture of normal distributions.[27] The inclusion of rich unobserved heterogeneity adds flexibility to the curvature of market demand (see Griffith et al. (2018)), which is important for recovering realistic patterns of pass-through (Weyl and Fabinger (2013)).

4.2 Identification of demand parameters

Prices vary across brands, sizes, retailers and time. We include a rich set of controls to strip out the price variation likely to be correlated with demand; specifically:

Brand-size effects, $\xi_{d(i)b(j)s(j)}^{(1)}$. Consumers may value one brand over another (e.g. Coke over Pepsi) for reasons not fully captured by observed product characteristics; failure to control for this would likely result in correlation between ϵ_{ijt} and prices. We include a full set of brand-size interaction terms that allows for the possibility that the strength of unobserved brand effects varies across product sizes and pack types. Numerous brand-sizes are available in both sugar sweetened and diet variants. We control for the amount of sugar per 100ml in a product in the characteristic vector, $\widetilde{\mathbf{x}}_j^{(1)}$. We are therefore able to identify the mean (as well as standard deviation) of the consumer group specific preferences for sugar, based on the restriction that the impact of sugar on utility does not vary across brands.

Brand-retailer effects, $\xi_{d(i)b(j)r}^{(2)}$ and size-retailer effects, $\xi_{d(i)s(j)r}^{(4)}$. These capture the possibility that either the prominence of products belonging to different

[26] We measure monthly TV advertising expenditure in the AC Nielsen Advertising Digest. We compute product specific stocks based on a monthly depreciation rate of 0.8. This is similar to the rate used in Dubois et al. (2018) on similar data in the potato chips market.

[27] The means (conditional on d) of the constant, cola, lemonade, store brand, fruit juice and large random coefficients are collinear with $\xi_{d(i)b(j)s(j)}^{(1)}$. We normalize them to zero. We allow for correlation (conditional on d) between the preferences for drinks and sugar.

brands, or of large versus small sizes, may vary across retailers. They also capture average differences in consumer brand and size preferences across retailers.

Brand-time effects, $\xi^{(3)}_{d(i)b(j)t}$ and size-time effects, $\xi^{(5)}_{d(i)s(j)t}$. The inclusion of time (quarterly) varying brand effects controls for shocks to national level demands for each brand that may be correlated with prices. Additionally controlling for time varying size effects captures any tendency through time for demands for larger versus small sizes to fluctuate, e.g., due to seasonal trends. As discussed in Section 3.1, the large four retailers that dominate the market have national pricing policies; the time varying effects help control for national level shocks to demand that could be correlated with these prices. We also control (through $\widetilde{x}^{(2)}_{jt}$) for product level advertising, which will capture the effect on demand of the (overwhelmingly national) advertising in the UK drinks market.[28] For convenience stores we use mean regional prices. We include region-time varying drinks effects in demand to control for the possibility of regional shocks to demand for drinks.

The remaining price variation that we exploit is *product level time series variation that is differential across retailers*. Our key identifying assumption is that this remaining variation is exogenous, and, in particular, the shocks to consumers' payoff functions, ϵ_{ijt}, are i.i.d. Hence, a restriction we make is the absence of retailer-time specific shocks to product demands that contemporaneously correlate with prices.[29] As outlined in Section 3.5, in the UK the main retailers use national prices and the average price of a given product is similar across retailers. However, the degree of co-movement over time in prices for a given product across retailer is low and, among the main retailers, each is sometimes the most expensive and at other times the least. Much of this variation is driven by differences in promotions, that, as outlined in Section 3.6, are typically agreed in advance in bilateral meetings between drinks firms and retailers. We assume that this creates randomness in the prices faced by consumers that is not a consequence of retailers anticipating time varying demand shocks that differ for their consumers compared to those in other retailers. These institutional details, coupled with national nature of advertising and the absence of targeted price offers and coupons in the UK drinks market, gives us confidence that our identifying assumption is a reasonable one.

[28]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.

[29]The $(\xi^{(1)}_{d(i)b(j)s(j)}, \xi^{(2)}_{d(i)b(j)r}, \xi^{(3)}_{d(i)b(j)t}, \xi^{(4)}_{d(i)s(j)r}, \xi^{(5)}_{d(i)s(j)t})$ effects control for all pairwise interactions between (b, s, r, t) but not higher order interactions.

4.3 Supply model

We model price competition among the firms operating in the UK drinks market. Consistent with the evidence of non-linear contracting laid out in the Competition Commission (2013) report, and discussed in Section 3.6, we model drinks firms as setting retail prices. We assume that they simultaneously set prices to maximize profits in a Nash-Bertrand game.[30] In Section 5.6 we show how our optimal tax results are influenced by different supply-side models.

Let $\boldsymbol{p}_m = (p_{1m}, \dots, p_{Jm})$ denote the prices that drinks firms set in market m, where markets are temporal.[31] Market demand for product j is given by:

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

where M_m denotes the potential size of the market.[32] We denote the marginal cost of product j in market m as c_{jm}.[33]

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

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

We assume firms engage in Bertrand competition and that the prices we observe in the data are the Nash equilibrium outcome of this game, and thus they satisfy the set of first order conditions: $\forall f$ and $\forall j \in \mathcal{J}_f$,

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

---

[30] An exception to this is the set of store brands, which are no-frills, low priced alternatives to the branded products. We treat these as a competitive fringe, assuming they are priced at marginal cost.

[31] In the supply model and simulations we average over short-run price variation, as this likely reflects the predetermined price reduction strategies agreed on by drinks firms and retailers rather than fundamentals of demand or supply. Specifically, let $\mathcal{M}$ denote the set of (r, t) pairs in market m, the market price for product j is defined as $p_{jm} = (|\mathcal{M}|)^{-1} \sum_{(r,t) \in \mathcal{M}} p_{jrt}$. We present results for the most recent market covered by our data, 2012.

[32] 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.

[33] 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.

From this system of equations we can solve for the implied marginal cost, c_{jm}, for each product in each market.

Counterfactual market equilibrium. When solving for the optimal tax rate we need to solve for the associated counterfactual market equilibrium. Denote the set of sugar sweetened beverages by $\mathcal{S}$ and the total sugar content of option $j \in \mathcal{S}$ by z_j (noting that for $j \notin \mathcal{S}$ $z_j = 0$). Given some tax rate τ, levied on the sugar in sweetened beverages, the set of first order conditions are: $\forall f$ and $\forall j \in \mathcal{J}_f$,

$$q_{jm}(\boldsymbol{p}'_m) + \sum_{j' \in \mathcal{J}_f} (p'_{j'm} - \tau z_{j'} - c_{j'm}) \frac{\partial q_{j'm}(\boldsymbol{p}'_m)}{\partial p_{jm}} = 0. \tag{4.6}$$

For any τ, we can solve the system of equations to obtain the vector of counterfactual equilibrium prices, $\boldsymbol{p}'_m = (p'_{1m}, \ldots, p'_{Jm})$.

Solving for the optimal tax rate also requires us to compute the derivative of the equilibrium price vector with respect to the tax rate, $\frac{d\boldsymbol{p}'_m}{d\tau}$. To obtain this we differentiate the first order conditions with respect to the tax rate and solve the resulting system of equations. See Appendix F for details.

5 Results

We estimate our empirical model of supply and demand in the drinks market and embed this within the tax design framework set out in Section 2 to quantify the importance of accounting for market power when designing externality correcting policy. Estimating a structural model of behavior in the market necessarily entails making assumptions about the nature of demand and firm competition. We use data on prices over the introduction of the UK's sugar sweetened beverage tax in 2018 to perform an out-of-sample validation of our estimated model.

5.1 Consumer substitution patterns

We estimate the demand model outlined in Section 4.1 using simulated maximum likelihood[34] and report the coefficient estimates in Appendix D. The estimated coefficients exhibit some intuitive patterns: those with more added sugar in their

[34]We allow all parameters to vary by consumer group and estimate the choice model separately by groups. 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. We conduct all post demand estimation analysis on the full sample.

diets (based on their purchases in the preceding year) have stronger preferences for sugary drinks products, and those with below median income are more sensitive to price, have stronger preferences for soft drinks and weaker preferences for 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 jointly 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 -2.4 (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 2.8, a further 25% of products have own-price elasticities with magnitude less than 1.8. This is consistent with the findings of Bonnet and Réquillart (2013) who estimate own-price elasticities for soft drink products in France between -2.13 and -3.95. The distribution of the cross-price elasticities exhibits a high degree of skewness, with the mean close to the 75th percentile. This reflects consumers' willingness to switch between products close together in product characteristic space.

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	-2.394	0.021	0.018	0.012	0.018	0.012	0.031	0.021	0.017	0.012	0.010	0.006	0.007
10x330ml can	0.037	-3.169	0.020	0.030	0.020	0.029	0.034	0.052	0.019	0.029	0.016	0.009	0.011
Diet													
2l bottle	0.013	0.008	-2.434	0.016	0.022	0.016	0.012	0.008	0.022	0.016	0.004	0.009	0.004
10x330ml can	0.014	0.021	0.027	-3.284	0.027	0.040	0.014	0.020	0.026	0.039	0.007	0.013	0.006

Notes: Numbers show the mean 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.

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 Cokes and Pepsi, and the mean elasticities for other (non-cola) sugar sweetened and diet beverages, and for fruit juice. The table highlights a number of intuitive patterns: (i) consumers are more willing to switch across cola products of the same variety (sugar vs. non-sugar) than they are 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.

Switching between sets of products

In Table 5.2 we summarize the effects of increasing the price of all sugar sweetened beverages by 1%.[35] This leads to a 1.55% fall in liters demanded of sugar sweetened beverages, i.e. our estimates yield an own-price elasticity for sugar sweetened beverages of 1.55. This is in line with Allcott et al. (2019a), who use US scanner data and an instrumental variable approach to estimate an own-price elasticity for sugar sweetened drinks ranging between -1.37 and -1.48, depending on the specification.

Around 33% of the reduction in demand for sugar sweetened beverages is diverted to alternative drinks, indicating that in response to a rise in price for sugar sweetened drinks, there is substantial switching to alternative drinks products. As we discuss in Section 2, if alternative products are supplied non-competitively, the degree of switching in response to a marginal tax rise between sugar sweetened beverages and alternatives is an important determinant of the optimal tax rate; the stronger this is, all else equal, the higher is the optimal rate. 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 5.4, the relatively high degree of substitution between the two product sets indicated by the diversion ratio plays an important role in determining the optimal rate.[36]

The 1% increase in the price of sugar sweetened beverages (at observed prices) leads to essentially no change in overall drinks expenditure. In Section 5.5 we explore under what conditions changes in equilibrium spending on the numeraire good in response to a tax impacts optimal policy.

[35] 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).

[36] 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 5.2: *Switching due to an increase in the price of sugar sweetened beverages*

Own price elasticity for sugar sweetened beverages	-1.55
	[-1.60, -1.50]
% lost demand diverted to substitute drinks	32.6
	[32.0, 33.7]
% change in overall drinks expenditure	-0.001
	[-0.017, 0.015]

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 the marginal cost of each product, and hence the price-cost margins and Lerner indexes (margin over price) at observed prices. In Table 5.3 we show the average price, cost, margin (all expressed per liter) and Lerner index for sugar sweetened beverages and alternative products; in Appendix D we show these by brand.

The average Lerner index is 0.46 for sugar sweetened beverages and 0.40 for substitute products.[37] This indicates that firms exercise a significant degree of market power when setting the prices both of sugar sweetened beverages and alternative drinks. The substantial price-cost margins of sugar sweetened beverages will act to lower the optimal tax rate on these products, relative to if they were competitively supplied. The converse is true for alternative drinks: given that consumers switch towards them from sugar sweetened beverages, positive margins for these products will act to raise the optimal sugar sweetened beverage tax. However, the quantitative impact of positive margins on the optimal tax rate depends on their value at the optimal rate, which will differ from their value at observed prices if firm respond to the tax by adjusting their margins, as well as the correlation in margins with how strongly consumers switch away from (or towards) the product in equilibrium in response to a marginal tax rise.

In Figure 5.1 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

[37] These averages include the store brands, which we assume are competitively priced. Among branded drinks the average Lerner index is 0.53; this broadly accords with evidence from accounting data, with gross margins in this market reported to be between 35-70% (see Competition Commission (2013)).

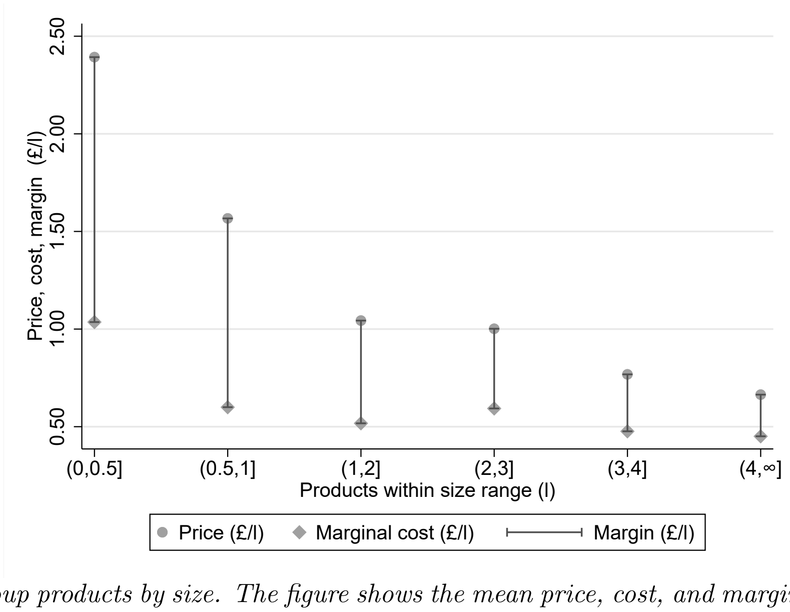
products are, on average, more expensive. Average marginal costs are broadly constant across the size distribution, with the exception of small single portion sizes, which, on average, have higher costs. Price-cost margins are declining in size – the average margin (per liter) is more than twice as large 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 5.3: *Summary of costs and margins*

	Sugar sweetened beverages	Alternative products
Price (£/l)	1.07	1.08
Marginal cost (£/l)	0.52	0.60
Price-cost margin (£/l)	0.56	0.48
Lerner index (margin/price)	0.46	0.40

Notes: We recover marginal costs for each product in each market. The table shows the price, marginal cost, price-cost margin 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).

Figure 5.1: *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).

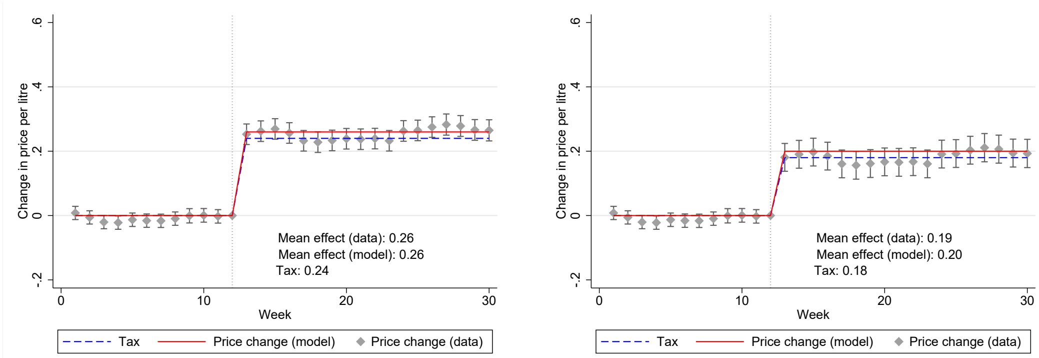
5.3 Tax pass-through and model validation

The margins presented in the preceding section are estimates at the observed price equilibrium. In response to the introduction of a tax, firms may choose to adjust their margins, for instance, increasing margins by passing the tax through to prices by more than one-for-one. 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.[38] We also provide direct evidence that our model succeeds in generating realistic tax pass-through predictions.

To do this we use the introduction of the UK's Soft Drinks Industry Levy (SDIL) in 2018 to compare our model predictions with what happened when this policy was introduced. This tax was introduced more recently than the period covered by our main dataset. We therefore supplement it with a weekly database of UPC level prices and sugar contents for drinks products, 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.[39]

Figure 5.2: *Comparison of model predictions with observed price changes*

(a) Higher rate treatment group (b) Lower rate treatment group



Notes: Grey markers show the estimated price changes (relative to the week preceding the introduction of the tax) for the set of products subject to the higher and lower rates. Full details are given in Appendix E. 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 model of the UK drinks market.

The UK's 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 prod-

[38] 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 of demand curvature. One way of relaxing restrictions on the curvature of market demands is through incorporating rich individual level preferences heterogeneity (see Griffith et al. (2018)).

[39] 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.

ucts with sugar content > 8g/100ml. We estimate the price changes for the sets of products subject to each rate and for the set of drinks products not subject to the tax – full details are provided in Appendix E. We find evidence that the tax was slightly overshifted, with price increases of 26p/liter for products subject to the higher rate and 19p/liter for products subject to the lower rate, implying average pass-through rates of 105-108%, with no change in the price of untaxed products. We simulate the effect of the tax using our estimated model of supply and demand. Figure 5.2 shows the estimated price changes in the data (grey markers) for the high and low tax groups (the figure for untaxed products is shown in the Appendix E), and the predicted price changes using our model, which 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. (2019), 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)).

5.4 Corrective tax results

In the preceding sections we show that (i) in response to price rises for sugar sweetened beverages, there is considerable consumer switching to alternative products, (ii) firms exercise a substantial degree of market power when setting the price of sugar sweetened beverages and alternative products, (iii) distortions from market power are largest among relatively small products, which, as a percentage of price will have the lowest tax burden under a sugar sweetened beverage tax, and (iv) in response to the UK's Soft Drinks Industry Levy firms slightly overshifted the tax.

These facts suggest that market power among both sugar sweetened beverages and substitute products will play an important role in determining optimal tax policy. In this section we use the demand and supply estimates to solve for the optimal tax rate, as well as the rates that would be set by alternative planners that ignore the various dimensions of market power.

Externalities

Externalities from sugar sweetened beverage consumption arise due to increased health care costs that are associated with their intake; see Scientific Advisory Committee on Nutrition (2015) for a review of 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 is also means that they are less likely than solid sources of calories to sate appetites, and therefore are associated with weight gain.[40] The bulk of associated health care costs are not borne by the individual, but instead by others, for example, through higher taxpayer costs of publicly funded systems and from increased premiums in insurance based systems. For instance, 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.

To place a value on the size of externalities from the sugar in sweetened beverages we draw on two pieces of epidemiological evidence. First, Wang et al. (2012) estimate the impact of a population reduction in consumption of sugar sweetened beverages on the net present value of long-run health care costs. We use this to compute the average implied health care saving per unit reduction in sugar from sweetened drinks. Health care costs in the UK are almost entirely borne by taxpayers, so we assume this represents a fiscal externality. Second, based on the World Health Organization recommendation that added sugar consumption should be below 10% of dietary calories, we assume that consumers with dietary sugar below this level do not generate externalities through their sugar sweetened beverage consumption. Together this implies the externality per 10g of sugar consumption is around £0.05 for the 80% of consumers with dietary sugar above 10% of their calorie intake (see Appendix F for details). A typical sugar sweetened beverage contains 3g of sugar per oz, so this translates into an externality of approximately 1.7 pence per oz. In Section 5.5 we show how our results vary with the nature of externalities, including their overall size, whether they are associated with alternative, non-taxed products, and their convexity in total consumption.

Tax policy

To illustrate the importance of various dimensions of market power on the optimal tax rate and associated welfare implications, we compare an 'Optimizing' social planner that sets the optimal tax rate with two alternative planners – a 'Pigouvian' planner and a 'Buchanan' planner. For ease of exposition, and as in Section 2, we refer to the externality generating products (i.e. the sugar sweetened beverages) as

[40] Typically fruit juice and milk based drinks are exempt from beverage taxes, because, while they are also rich in sugar, they have off-setting positive health benefits. However, this is controversial (e.g., see Public Health England (2018)). In Section 5.5 we consider the implications for tax design if externalities arise from these (untaxed) products.

"sin products" and the other, non-externality generating, drinks in the market as "alternative products". The "tax rate" to which we refer in this section is levied on the sugar in sweetened beverages.[41]

The Pigouvian planner sets a tax rate, τ^P, which accounts only for distortions that arise from externalities, ignoring any distortions associated with the exercise of market power. The Buchanan planner sets a tax rate, τ^B, which accounts for the distortions from externalities and market power of sin products, but ignores the distortions from market power of alternative products. In contrast, the Optimizing planner chooses the efficiency maximizing tax rate, which we denote τ^*, accounting for distortions that arise from the existence of externalities and the exercise of market power, both for sin and alternative products. In terms of the expressions in equation (2.3), the three planners set tax rates implicitly defined as follows:

$$\tau^* = \widetilde{\Phi}(\tau^*) - \widetilde{\mu}^S(\tau^*) + \widetilde{\mu}^N(\tau^*) \times \widetilde{\Theta}(\tau^*) \tag{5.1}$$
$$\tau^P = \widetilde{\Phi}(\tau^P) \tag{5.2}$$
$$\tau^B = \widetilde{\Phi}(\tau^B) - \widetilde{\mu}^S(\tau^B). \tag{5.3}$$

In Table 5.4 we report the three tax rates, along with the terms that enter into the tax formula. The optimizing planner sets a tax rate of £1.89 per kg of sugar, which corresponds to tax rates of around 0.8¢/oz for a typical sugar sweetened beverage.[42] At the time of writing US sugar sweetened drink tax rates vary between 1 and 2¢/oz, and the "higher rate" of the UK tax corresponds to around 1¢/oz. The optimal rate results in an average increase in the price of sugar sweetened drinks of 18%, (median pass-through is 109%).[43]

This optimal rate reflects three forces. At that rate the externality correcting component of the tax formula, $\widetilde{\Phi}(.)$, takes the value 4.22. This is offset by the correction for market power of the sin products, $\widetilde{\mu}^S(.)$, which takes the value -3.92. However, the market power among alternatives acts to the raise the optimal tax; its effect is given by the product of $\widetilde{\mu}^S(\tau)$ and $\widetilde{\Theta}(\tau)$ and is 1.61 (=0.37×4.34). The Pigouvian planner, who ignores distortions from the exercise of market power, overshoots the optimal tax rate, setting a rate of £4.24 per kg of sugar, or 1.8¢/oz for

[41]Fruit juices and flavored milks are exempt from the tax, as is the case under most taxes implemented in practice. In Section 5.5 we consider the implication for tax policy when externalities are also associated with the sugar in these products.

[42]Coca Cola contains 106g of sugar per liter. Therefore the tax rate implies tax per liter of Coca Cola is £0.20, which converted to cents per oz based on a $\$ - £$ exchange rate of 0.75, is 0.8¢ per oz.

[43]This feeds through into an 18% reduction in the probability, on average, a consumer purchases a sugar sweetened beverage and, conditional on buying, a reduction in volume of around 16%. See Table F.1 in Appendix F.

a typical sugary beverage. Conversely, the Buchanan planner, who ignores market power among alternative products, undershoots, setting a tax rate of £0.78 per kg, or 0.3¢/oz for a typical sugary beverage.

Table 5.4: *Tax rates set by alternative planners*

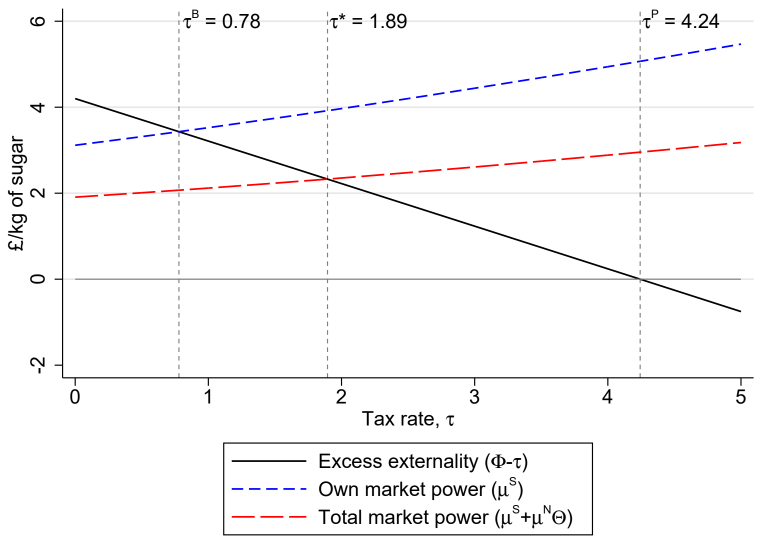
	Tax rate (τ)		Externality correction ($\widetilde{\Phi}(\tau)$)		Market power of: Sin products ($\widetilde{\mu}^S(\tau)$)		Market power of: Alternatives ($\widetilde{\mu}^N(\tau)$)		Equilibrium switching ($\widetilde{\Theta}(\tau)$)
Planner	(1)		(2)		(3)		(4)		(5)
Optimizing	1.89	=	4.22	−	-3.92	+	4.34	×	0.37
(included?)			✓		✓		✓		✓
Pigouvian	4.24	=	4.24	−	*-5.07*	+	*4.94*	×	*0.43*
(included?)			✓		✗		✗		✗
Buchanan	0.78	=	4.21	−	-3.43	+	*4.11*	×	*0.33*
(included?)			✓		✓		✗		✗

Notes: The tax rate set by each planner (listed in the first column) is shown in column (1). This is equal to the relevant components (listed in columns (2)–(5)) that are used to implicitly define the rate, as described in equations (5.1) – (5.3). The terms in italics are the values of the component at that tax rate, but that are not included in the implicit formula for the tax rate associated with that planner.

Figure 5.3 represents this graphically by showing how the various components of the tax formula vary with the tax rate. The blue line shows that the correction for market power on the sin goods, $\widetilde{\mu}^S(.)$, is increasing approximately linearly in the tax rate. In other words, a higher rate of tax exacerbates the distortions associated with the exercise of market power for this set of products. This happens for two reasons. First, the tax leads firms to increase product level margins by overshifting the tax – as prices rise, relatively elastic consumers switch away, leaving firms optimally pricing for the remaining set of slightly less elastic consumers, acting to raise profit maximizing margins. Second, consumers adjust their basket of taxed goods towards products with relatively high margins – the tax leads to more switching away from large (low margin) sugar sweetened beverages, increasing the weight in the $\widetilde{\mu}^S(.)$ term of the optimal tax formula on high margin small products.

The red line shows the total correction for market power, $\widetilde{\mu}^S(.) + \widetilde{\mu}^N(.) \times \widetilde{\Theta}(.)$. It lies below the blue line, reflecting the role that imperfect competition among alternative products plays in lowering the overall market power correction. The red line is less steeply sloped than the blue line: at higher tax rates, a larger fraction of the distortions from the exercise of market power on externality generating goods is offset by distortions from market power of substitutes. This happens for two reasons. Firstly, at a higher tax rate, and hence higher prices among the taxed products, a marginal increase in the tax rate causes more switching to alternative products i.e. $\widetilde{\Theta}(.)$ is increasing in τ. Secondly, consumers switch most strongly towards high margin alternatives like fruit juice, which drives up the weighted average margin on alternative products, $\widetilde{\mu}^N(.)$.

Figure 5.3: *Components of tax rate*



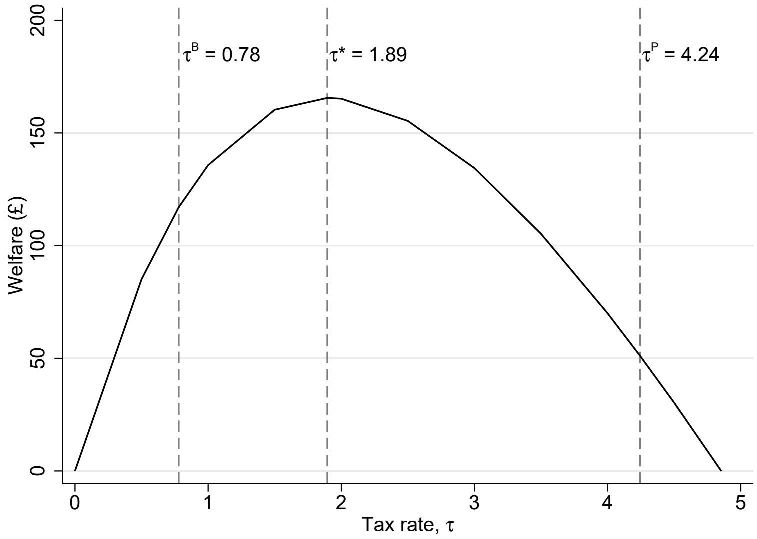
Notes: The black line plots the value of the externality correction minus the tax rate, $\widetilde{\Phi}(\tau) - \tau$, at each value of τ. The dashed blue line plots the correction for market power among taxed goods, $\widetilde{\mu}^S(\tau)$, and the dashed red line plots the total market power correction, $\widetilde{\mu}^S(\tau)+\widetilde{\mu}^N(\tau)\widetilde{\Theta}(\tau)$, for each value of τ. The vertical dashed lines plot the values of the rates set by the Optimizing, Pigouvian and Buchanan planners.

The black line shows the excess of the tax rate over the externality correcting term i.e. $\tau - \widetilde{\Phi}(\tau)$. The tax rates chosen by the three alternative planners are given by the intersection of this line and the blue line (Buchanan), red line (Optimizing) and the horizontal axis (Pigouvian).

Welfare

In Figure 5.4 we show how the choice of tax rate impacts welfare. The optimal rate increases welfare by £166 million. The sub-optimally low rate set by the Buchanan planner achieves only 70% of these gains, while the sub-optimally high tax rate set by the Pigouvian planner leads to only 30% of possible gains. This highlights how naive policymaking that either ignores distortions from the exercise of market power completely, or that accounts only for such distortions from taxed products, ignoring positive margins on other products, can lead to substantial unrealized welfare gains.

Figure 5.4: *Welfare function*



Notes: The black line shows how the change in total welfare depends on τ. The vertical dashed lines plot the values of the rates set by the Optimizing, Pigouvian and Buchanan planners.

Table 5.5: *Welfare components*

		Change (relative to zero tax) in:					
		Welfare components					
Planner	Tax rate	Consumer surplus	External costs	Tax rev.	SSBs profits	Alts profits	Total welfare
Optimizing	1.89	-626	-550	487	-405	159	166
	[1.83, 1.97]	[-651, -612]	[-582, -527]	[478, 503]	[-422, -396]	[155, 171]	[148, 188]
Pigouvian	4.24	-1149	-919	721	-743	303	51
	[4.19, 4.28]	[-1194, -1123]	[-969, -885]	[706, 743]	[-773, -725]	[295, 324]	[25, 79]
Buchanan	0.78	-287	-266	253	-186	71	117
	[0.67, 0.85]	[-299, -281]	[-282, -255]	[248, 261]	[-194, -182]	[69, 76]	[108, 130]

Notes: Impact of tax on total welfare equals change in consumer surplus minus changes in externalities plus change in gross profits. Gross profits are equal to the sum of tax revenue, net profits from sugar sweetened beverages (SSBs) and net profits from alternative (alts) drinks. Numbers are in £million per year. 95% confidence intervals are given in square brackets.

Table 5.5 shows the components of total welfare under each of the three alternative tax rates. The increase in welfare under the optimal tax rate is comprised of a fall in consumer surplus of £626 million and sugar sweetened beverage profits of £405 million, which is more than offset by a reduction in the external costs of sugar sweetened beverage consumption of £550 million, tax revenue of £487 million, and an increase in profits for alternative drinks products of £159 million. Pigouvian policy leads to much larger falls in both consumer surplus and sugar sweetened beverage profits than optimal policy, which are not fully compensated for by higher tax revenue and a larger reduction in externality costs. Conversely, the sub-optimally low tax rate set by the Buchanan planner leads to smaller reductions in external costs and lower tax revenue than under the optimal policy, which are not offset by lower falls in consumer surplus and profits.

5.5 Sensitivity

Our main results are based on a calibration of the externalities from sugar sweetened beverages that, while motivated by epidemiological evidence, there is considerable uncertainty around. In this section, we consider how variation in the nature of externalities impacts our results. We also consider the impact of the existence of market power associated with the numeraire good. We summarize the sensitivity exercise in Table 5.6. Column (1) repeats the optimal tax rate and associated welfare gain in Table 5.5, and shows the *difference* in tax rates and welfare changes under the Pigouvian and Buchanan planners. The remaining columns show this under alternative assumptions about the nature of externalities, and market power for the numeraire good.

Externality magnitude. Columns (2) and (3) of Table 5.6 show the effect of the scale of externalities being 25% lower and 25% higher than those in our baseline calibration and implied by Wang et al. (2012).[44] The optimal tax rate and associated welfare gain are unsurprisingly sensitive to the scale of externalities: when externalities are smaller, the optimal tax rate and associated welfare gains are smaller, and, conversely, when externalities are larger, the optimal tax rate and welfare gains are larger. However, the effects of ignoring the different dimensions of market power are robust to the size of the externalities: at all three levels the Pigouvian planner sets a sub-optimally high rate that results in lower welfare by over £100m and the Buchanan planner sets a sub-optimally low tax rate that results in lower welfare by around £50m.

Externality convexity. Our baseline results assume that externalities from sugar sweetened beverage consumption arise only for those consumers that exceed World Health Organization recommendations to consume less than 10% of their calories from added sugar. Columns (4) and (5) of Table 5.6 show the impact of assuming that externalities arise from all consumers regardless of their total dietary sugar,

[44] One reason the social costs of sugar sweetened beverage consumption may be larger than the estimates by Wang et al. (2012) is because of internalities. In Appendix A we show conditions under which internalities can be accommodated by our optimal tax framework.

and only for those consumers with more than 15% of calories from added sugar.[45] The more concentrated externalities are among high sugar consumers the higher is the optimal rate and associated welfare gain. This effect, however, is relatively weak, reflecting only a modest positive relationship between changes in sugar from the tax and total dietary added sugar. At each degree of externality convexity, the sub-optimal rates set by the Pigouvian and Buchanan planner lead to substantial unrealized welfare gains that are similar in magnitude to the baseline case.

Externality spillovers. The baseline optimal tax rate leads to an average reduction in the sugar consumers get from sweetened beverages of 34%. However, because, in part, they switch to fruit juice and sweetened milk, sugar from these sources rises by around 8%.[46] In columns (6) and (7) of Table 5.6 we show how tax policy is impacted if externalities are associated with the sugar in these alternative drinks, and, additionally, with the numeraire good.[47,48] In this case a tax levied on the sugar in sweetened beverages is a less effective instrument for tackling externalities since some consumers switch to alternative sources of external costs. This is reflected in a lower optimal tax rate and associated welfare gains relative to the baseline. However, again the gap between optimal and Pigouvian or Buchanan policy in both tax rates and welfare changes in similar to the baseline.

Numeraire good market power. In the final two columns we consider the impact that market power for the numeraire good has on optimal policy. In particular, in column (8) we assume the margin on the numeraire good is 50% of the weighted average margin for alternative drinks, and in column (9) we assume it is equal to the weighted average margin on alternative drinks. The more market power is associated with the numeraire good the higher is the optimal tax rate. This is because the tax causes a shift in consumption from sugar sweetened beverages towards the numeraire good.[49] Therefore, a higher numeraire good margin raises the market power of alternative goods, which acts to raise the optimal tax rate. Higher market power on the numeraire good also impacts the relative welfare gains of optimal

[45]To isolate the effect of convexity in the externality function from its scale we hold fixed the average marginal externality across all consumers.

[46]For details see Table F.1 in Appendix F.

[47]For concreteness, here we assume the numeraire good has 54.5g of sugar per £1 of expenditure, which is based on a commonly purchased chocolate bar.

[48]See Appendix A for optimal tax formula when there are externality spillovers.

[49]Note, this does not contradict Table 5.2. There we show a marginal increase in sugar sweetened beverage prices from their zero tax level leads to little switching to the numeraire good. However, at the price equilibrium that prevails at the optimal tax rate, a marginal increase in the tax rate does lead to some switching towards the numeraire.

policy over Pigouvian and Buchanan policy; the unrealized gains from Pigouvian policy shrink, but those associated with Buchanan policy rise. In other words, the mistake a social planner makes by correcting only for market power on the taxed set of products grows as total market power of alternatives rises.

Table 5.6: *Tax policy under alternative assumptions about externalities and the numeraire good*

		Externality						Numeraire margin	
	Baseline	Size		Convexity		Spillovers			
		-25%	+25%	Linear	>15%	All drinks	+Num.	$0.5\times\tilde{\mu}^{\mathcal{N}}$	$\tilde{\mu}^{\mathcal{N}}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Optimal rate	1.89	1.03	2.75	1.72	2.21	1.69	1.51	2.17	2.43
Difference from optimal:									
Pigouvian	+2.35	+2.14	+2.56	+2.28	+2.48	+2.23	+1.95	+2.08	+1.82
Buchanan	-1.12	-1.01	-1.22	-1.08	-1.17	-1.07	-0.96	-1.39	-1.65
Δ W under optimal rate	166	53	324	139	216	138	121	189	210
Difference from optimal:									
Pigouvian	-114	-123	-108	-115	-114	-113	-107	-78	-53
Buchanan	-49	-50	-47	-48	-48	-48	-43	-66	-82

Notes: Table reports the tax rate, and associated welfare change, for an Optimizing planner, and the differences in tax rate and welfare changes relative to optimal policy under Pigouvian and Buchanan planners. Column (1) reports our baseline results (summarizing information in Table 5.5). Columns (2) and (3) show results if the scale of externalities is 25% below or above their baseline value. Columns (4) and (5) show results if all consumers, and only those with dietary added sugar in excess of 15% of calories, generate externalities. Columns (6) and (7) shows results if there are externalities associated with the sugar in alternative drinks, and if additionally sugar from the numeraire good generates externalities. Columns (8) and (9) show results if there is market power associated with the numeraire good, and, in particular, if its price-cost margin is half of, and equal to, the weighted average margin of alternative drinks.

5.6 Tax policy and market competition

The potential for an optimally set tax on sin products to improve welfare depends on the degree of competition in the market. In the simple case in which one sin product is supplied non-competitively and all alternatives are supplied competitively, as long as the marginal externality is at least as large as the product's price-cost margin, moving towards a more competitive market structure *reduces* welfare, but this is completely unwound by optimal tax policy, which achieves the first best – competition and optimal tax policy are *perfect substitutes.*[50]

However, when there are many products, differing in their margins and the extent to which they create externalities, and when non-externality generating alternative goods are supplied non-competitively, the relationship between competition

[50] In particular, suppose a monopolist sets a fixed margin on its product given by $\mu \leq \phi$, where ϕ is the marginal 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.

and optimal tax policy is more nuanced. To illustrate this we compare the impact of tax policy under the observed product ownership structure with a set of counterfactual ownership structures designed to highlight the roles played by distortions from the exercise of market power for both the sin products and for alternatives. In particular, we consider tax policy when (i) all products are produced by single product firms, $\{SPF\}$, (ii) the sin products (i.e. sugar sweetened beverages) are produced by a joint profit maximizer and the alternative products are produced by single product firms, $\{JPM^{\mathcal{S}}, SPF^{\mathcal{N}}\}$ (iii) the sin products are produced by single product firms and the alternatives by a joint profit maximizer, $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$ and (iv) both sets of products are produced by one joint profit maximizer, $\{JPM\}$.[51]

Table 5.7: *Impact of tax policy under alternative market structures*

	No tax	Tax							
		Optimizing		Pigouvian			Buchanan		
	Welfare	τ^*	ΔW	τ^P	ΔW	% loss	τ^B	ΔW	% loss
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Observed ownership		1.89	166	4.24	51	*69*	0.78	117	*29*
Counterfactual ownership:									
$\{SPF\}$	423	2.24	242	4.24	158	*35*	1.32	211	*13*
$\{JPM^{\mathcal{S}}, SPF^{\mathcal{N}}\}$	290	1.61	104	4.24	-26	*125*	0.83	83	*20*
$\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$	-197	2.22	248	4.24	156	*37*	1.37	220	*11*
$\{JPM\}$	-1545	1.31	69	4.24	-93	*236*	-0.17	-22	*132*

Notes: We solve for the rates set by the Optimizing, Pigouvian and Buchanan planners (as expressed in equations (5.1)–(5.3)) under counterfactual ownership structures. The first row shows effects under the observed ownership structure (summarizing information in Table 5.5). The remaining rows show effects under the counterfactual ownership structures described in the text. Column (1) show the impact of different market structures under zero tax on welfare relative to the observed ownership structure. Numbers in columns (3), (5) and (7) show the effect on welfare under the corresponding market structure of tax policy. All welfare numbers are expressed in £millions. Column (6) and (9) show the % loss in welfare under Pigouvian and Buchanan policy relative to optimal policy. Confidence intervals are reported in Table F.2 in Appendix F.

Column (1) of Table 5.7 reports how welfare changes under these counterfactual market structures at zero tax, relative to the observed structure.[52] Moving from the observed ownership structure towards the more competitive $\{SPF\}$ market structure leads to a welfare increase; the gains from lower distortions from the exercise of market power more than offset the loss due to larger distortions from

[51] Throughout we assume different sizes of the same brand, regular vs. diet variant, and pack type (bottles vs. cans) belong to the same firm (so the single product firms own multiple sizes). For each ownership structure we assume the store brands are priced competitively.

[52] Confidence intervals are reported in Table F.2 in Appendix F.

externalities. The converse is true when moving to the less competitive market structure $\{JPM\}$; welfare falls because the much larger loss from the distortions from market power exceed the gains from lower externalities. Ex ante, it is unclear whether moving from the observed ownership structure to either $\{JPM^{\mathcal{S}}, SPF^{\mathcal{N}}\}$ or $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$ makes the market more or less competitive. In each case the impact is to somewhat lower the costs of distortions from market power: in both cases there is an increase in welfare net of externalities. This is reinforced by lower external costs when moving to market structure $\{JPM^{\mathcal{S}}, SPF^{\mathcal{N}}\}$ (due to relatively high sin good prices) and hence welfare rises. Conversely, it is more than offset by higher external costs when moving to $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$ (due to relatively low sin good prices) and hence welfare falls.

Columns (2) and (3) of Table 5.7 report the optimal tax rates and welfare under each market structure, relative to when no tax is in place under the same market structure. Under single product firms ($\{SPF\}$) the optimal tax rate and welfare gain are both higher than under the observed ownership structure. Conversely, under one joint profit maximizer ($\{JPM\}$) the optimal tax rate and welfare gains are lower. With market structure $\{JPM^{\mathcal{S}}, SPF^{\mathcal{N}}\}$ the optimal tax rate is lower than under $\{SPF\}$. This is driven mainly by higher price-cost margins on sin products leading to a lower optimal tax rate. The optimal rate and welfare gain under $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$ are similar to under $\{SPF\}$. This is because, although the weighted average margin on the alternative products is higher under $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$, this is offset by a reduction in the equilibrium switching term, which is driven in part by the fact that the firm producing all the alternative drinks responds to tax by marginally *raising* the prices of these products.

The remaining columns of Table 5.7 show the tax rate set by the Pigouvian and Buchanan planners, the associated welfare change and % welfare loss from failing to pursue optimal policy. Under each market structure the Pigouvian planner sets a sub-optimally high rate, and the Buchanan planner a sub-optimally low rate. The size of unrealized welfare gains from failing to set the optimal rate varies with market structure. When the sin goods are produced by single product firms ($\{SPF\}$ and $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$) the lost welfare from setting the Pigouvian rather than optimal rate is less, compared with the other market structures, reflecting the relatively low equilibrium price-cost margins for sin products. When sugar sweetened beverages are priced by a joint profit maximizer, price-cost margins for sin products are high and Pigouvian policy is actually welfare reducing, relative to no tax.

The costs of setting the Buchanan tax rate instead of the optimal tax rate are lower under three of the four counterfactual market structures than they are under the observed market structure. Under $\{SPF\}$ and $\{JPM^{\mathcal{S}}, SPF^{\mathcal{N}}\}$ the alternative products are supplied relatively competitively and therefore the cost from policy ignoring distortions from market power for these goods under these structures is relatively low. Under market structure $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$ it is the relatively low degree of switching between the product sets that dampens the losses from pursuing Buchanan over optimal policy. In contrast, under the market structure $\{JPM^{\mathcal{S}}\}$ the costs of pursuing Buchanan (which entails subsidizing the sin products) over the optimal policy is much higher than under the observed market structure, because the margins on alternative products are very high and consumer willingness to switch between the products sets is higher, relative to under $\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$.

6 Summary and discussion

In this paper we show that accounting for distortions from the exercise of market power, both for externality generating goods and untaxed alternatives, is crucial when designing corrective taxes. All else equal, higher equilibrium price-cost margins for sin products lower the welfare-maximizing tax rate. However, the exercise of market power among alternative goods creates an off-setting effect: higher equilibrium price-cost margins of alternatives, and the more strongly these goods are substitutable for the externality generating products, leads to a higher optimal tax rate. The effectiveness of tax policy depends on relative price-cost margins; substantial unrealized welfare gains may result from ignoring positive margins altogether, or from failing to account for the margins of alternative products. We illustrate this in an application to sugar sweetened beverage taxation, quantifying the size of these unrealized welfare gains and exploring the relationship between tax policy and competition. We think our insights apply more widely, and conclude by discussing extensions that would be useful in implementing our approach in other settings.

Firms may respond to tax by re-designing their products, or altering their product offering. The incentives that firms have to undertake product re-design depend on production costs, the nature of demand, and also the structure of the tax. In most instances sugar sweetened beverage taxes are volumetric (levied per fluid ounce), creating no incentive for firms to reduce the sugar content of their products.[53] However, other tax structures contain non-linearities that strongly encourage products re-design.[54] A promising avenue for future research is to use data covering the in-

[53]The majority of products are already available in a sugary and non-sugary variety.

[54]For instance, the UK's soft drinks tax, implemented in 2018, has a multi-rate tax schedule designed to encourage firms to reduce product sugar content.

troduction of such systems to study firms' choices over product re-design, and hence the implications for tax policy.

Many "sin" taxes aimed at combating unhealthy behaviors apply within reasonably narrowly defined markets. This makes it feasible to estimate product level substitution patterns and price-cost margins for the sin goods and the most relevant alternative products. However, sources of environmental externalities are often used in many sectors characterized by market power.[55] The existing literature on the interaction of market power and environmental policy primarily focuses on imperfect competition among producers of a particular source of externality. A key message of our paper is that distortions from the existence of market power among alternative goods are also important in determining effective tax policy; for instance, the appropriate rate of tax for carbon depends on the magnitude of distortions from non-competitive behavior by suppliers of carbon-based fuels relative to those of alternative energy sources. In this context, an empirical challenge is to estimate price-cost margins and switching across such a wide set of goods. A potentially fruitful avenue is to harness alternative empirical methods designed to estimate mark-ups across broad sectors, for instance, by drawing on the burgeoning literature that uses production function techniques (De Loecker et al. (2020)).

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.

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.

Berry, S. T. and P. A. Haile (2010). Nonparametric Identification of Multinomal Choice Demand Models with Heterogeneous Consumers. *Cowles Foundation Discussion Paper* (1718), 1–41.

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

[55]See, e.g., Bushnell et al. (2008) on electricity markets, and Hastings (2004) on gasoline markets.

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.

Britvic (2014). Britvic Soft Drinks Review 2014. Britvic Soft Drinks Ltd.

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

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, March). 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.

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.

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

DellaVigna, S. and M. Gentzkow (2019). Uniform Pricing in US Retail Chains. *Quarterly Journal of Economics*, 74.

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*.

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

Global Food Research Program (GFRP) (2019). 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, April). 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.

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.

Hitsch, G. J., A. Hortacsu, and X. Lin (2017). Prices and Promotions in U.S. Retail Markets: Evidence from Big Data. *SSRN Electronic Journal*.

Jaravel, X. and A. Olivi (2019). Optimal Taxation and Demand-Led Productivity. *SSRN Electronic Journal*.

Kaplow, L. (2012). Optimal Control of Externalities in the Presence of Income Taxation. *International Economic Review 53*(2), 487–509.

Lancaster, K. (1971). *Consumer Demand: A New Approach*. Columbia University Press New York.

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.

Oates, W. E. and D. L. Strassmann (1984, June). Effluent fees and market structure. *Journal of Public Economics 24*(1), 29–46.

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

Pigou, A. C. (1920). *The Economics of Welfare*. McMillan&Co., London.

Polyakova, M. and S. Ryan (2019). Subsidy Targeting with Market Power. *NBER Working Paper 26367*.

Preonas, L. (2019). Market Power in Coal Shipping and Implications for U.S. Climate Policy.

Public Health England (2018). Sugar reduction: What about the drinks not included in the Soft Drinks Industry Levy? Blog: Public health matters.

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). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics 83*(2), 217–230.

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 (2019). The Impact of Soda Taxes: Pass-through, Tax Avoidance, and Nutritional Effects. *SSRN Electronic Journal*.

Thomassen, Ø., H. Smith, S. Seiler, and P. Schiraldi (2017). Multi-Category Competition and Market Power: A Model of Supermarket Pricing. *American Economic Review 107*(8), 2308–2351.

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. *The 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.

Appendix

Corrective tax design and market power

Martin O'Connell and Kate Smith
September, 2020

A Optimal tax formulae

A.1 Derivation of equation (2.3)

The first order condition of the planner's problem (equation (2.2)), after application of Roy's Identity, is:

$$-\sum_i \Phi_i'(Z_i(\tau^*))Z_i'(\tau^*) + \sum_j (p_j(\tau^*) - c_j)q_j'(\tau^*) = 0.$$

Note that $q_j'(\tau) \equiv \sum_k \frac{\partial q_j}{\partial p_k}\frac{dp_k(\tau)}{d\tau}$ and $Z_i'(\tau) = \sum_j z_j q_{ij}'(\tau)$.

Defining the marginal externality, $\phi_i(\tau) \equiv \Phi_i'(Z_i(\tau))$, and noting $p_j(\tau) - c_j = \mu_j(\tau) + \tau z_j$, we can re-write this expesssion as:

$$-\sum_i \phi_i(\tau^*)Z_i'(\tau^*) + \sum_j (\mu_j(\tau^*) + \tau^* z_j)q_j'(\tau^*) = 0,$$

or

$$\tau^* = \frac{\sum_i \phi_i(\tau^*)Z_i'(\tau^*)}{Z'(\tau^*)} - \frac{1}{Z'(\tau^*)}\sum_j \mu_j(\tau^*)q_j'(\tau^*) \tag{A.1}$$

$$= \frac{\sum_i \phi_i(\tau^*)Z_i'(\tau^*)}{Z'(\tau^*)} - \frac{1}{Z'(\tau^*)}\left(\sum_{j\in\mathcal{S}} \mu_j(\tau^*)q_j'(\tau^*) + \sum_{j\in\mathcal{N}} \mu_j(\tau^*)q_j'(\tau^*)\right), \tag{A.2}$$

where $Z'(\tau) \equiv \sum_j z_j q_j'(\tau) = \sum_i Z_i'(\tau)$.

Define the weights: $\omega_i(\tau) \equiv \frac{Z_i'(\tau)}{Z'(\tau)}$, $w_j^{\mathcal{S}}(\tau) \equiv \frac{q_j'(\tau)}{Q'^{\mathcal{S}}(\tau)}$ and $w_j^{\mathcal{N}}(\tau) \equiv \frac{q_j'(\tau)}{Q'^{\mathcal{N}}(\tau)}$, where $Q'^{\mathcal{S}}(\tau) \equiv \sum_{j\in\mathcal{S}} q_j'(\tau)$ and $Q'^{\mathcal{N}}(\tau) \equiv \sum_{j\in\mathcal{N}} q_j'(\tau)$, we can write (as in equation (2.3)):

$$\tau^* = \sum_i \phi_i(\tau^*)\omega_i(\tau^*) - \frac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)}\sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}(\tau^*)\mu_j(\tau^*) +$$
$$\frac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)}\sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}(\tau^*)\mu_j(\tau^*) \times \frac{Q'^{\mathcal{N}}(\tau^*)}{-Q'^{\mathcal{S}}(\tau^*)}.$$

A.2 Extensions

Externality spillovers. Suppose there are three sets of products: (i) set $\mathcal{S}$ contain attribute z and are subject to tax; (ii) set $\mathcal{L}$ contain attribute z but are outside the scope of the tax; (iii) the remaining set of products contain none of attribute z and are untaxed. It is useful to denote the products in set (ii) and (iii) by $\mathcal{N}$.

Define $Z'^{\mathcal{S}}(\tau) = \sum_{j\in\mathcal{S}} z_j q'_j(\tau)$ and $Z'^{\mathcal{A}}(\tau) = \sum_{j\in\{\mathcal{S}\bigcup\mathcal{L}\}} z_j q'_j(\tau)$ to be the impact of a marginal change in the tax rate on equilibrium intake of the externality generating attribute from the set of taxed products and the set of all products containing the attribute respectively. In this case the optimal rate of tax can be expressed:

$$\tau^* = \frac{Z'^{\mathcal{A}}(\tau^*)}{Z'^{\mathcal{S}}(\tau^*)} \times \sum_i \phi_i(\tau^*)\omega_i(\tau^*) - \frac{Q'^{\mathcal{S}}(\tau^*)}{Z'^{\mathcal{S}}(\tau^*)} \sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}(\tau^*)\mu_j(\tau^*) + \frac{Q'^{\mathcal{S}}(\tau^*)}{Z'^{\mathcal{S}}(\tau^*)} \sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}(\tau^*)\mu_j(\tau^*) \times \frac{Q'^{\mathcal{N}}(\tau^*)}{-Q'^{\mathcal{S}}(\tau^*)},$$

where $\omega_i(\tau) = \frac{Z_i'^{\mathcal{A}}(\tau)}{Z'^{\mathcal{A}}(\tau)}$. In this case, the externality correcting component in multiplied by the ratio $\frac{Z'^{\mathcal{A}}(\tau^*)}{Z'^{\mathcal{S}}(\tau^*)}$. If, in equilibrium, a marginal increase in the tax rate induces switching from taxed to untaxed products that create externalities, then $\frac{Z'^{\mathcal{A}}(\tau^*)}{Z'^{\mathcal{S}}(\tau^*)} < 1$, and, all else equal, the optimal tax rate is lower.

Numeraire good market power. Denote consumer i's demand for the numeraire good by $X_i(\mathbf{p})$, where $X_i(\mathbf{p}) = y_i - \sum_j p_j q_{ij}(\mathbf{p})$, and total demand for the numeraire good by $X(\mathbf{p}) = \sum_i X_i(\mathbf{p})$. In our baseline analysis we assume the numeraire good is competitively supplied. Suppose it is non-competitively supplied with a price-cost margin (which, because its price is normalized to 1, is the same as the Lerner index) of $\tilde{\mu}$. In this case the planner's problem becomes:

$$\max_\tau v(\mathbf{p}(\tau)) - \sum_i \Phi_i(Z_i(\tau)) + \sum_j (p_j(\tau) - c_j)q_j(\tau) + \tilde{\mu}X(\tau)$$

and the optimal tax characterization (equation (2.3)) becomes:

$$\tau^* = \sum_i \phi_i(\tau^*)\omega_i(\tau^*) - \frac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)} \sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}(\tau^*)\mu_j(\tau^*) + \frac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)} \sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}(\tau^*)\mu_j(\tau^*) \times \frac{Q'^{\mathcal{N}}(\tau^*)}{-Q'^{\mathcal{S}}(\tau^*)} + \frac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)}\tilde{\mu} \times \frac{X'(\tau^*)}{-Q'^{\mathcal{S}}(\tau^*)}.$$

The original characterization is augmented by a new term: $\frac{Q'^{\mathcal{S}}(\tau^*)}{Z'(\tau^*)}\tilde{\mu} \times \frac{X'(\tau^*)}{-Q'^{\mathcal{S}}(\tau^*)}$. Interpretation of the impact of this term is similar to the impact of the set of alternative products, $\mathcal{N}$, on the optimal tax. When the numaraire good is a substitute/complement to the product set $\mathcal{S}$, a higher numeraire good margin, all else equal, acts to raise/reduce the optimal tax rate.

Internalities. Here we illustrate how our framework can accommodate certain forms of internalities. Suppose consumers suffer from internalities and their underlying utility takes the form $v_i(\mathbf{p})-\varphi_i \sum_j z_j q_{ij}$, and there are heterogeneous marginal externalities. The optimal tax can then be written as in equation (2.3), but with $\overline{\phi}_i(\tau) = \phi_i(\tau) + \varphi_i$ i.e. the sum of the marginal externality and internality, in place of ϕ_i.

The following model of consumer choice leads to utility taking the form $v_i(\mathbf{p}) - \varphi_i \sum_j z_j q_{ij}$. Suppose consumer i chooses one product from the set of available products Ω according to the choice model:

$$\max_{j\in\Omega}\{\tilde{U}_{ij} = \alpha_i(y_i - p_j) + \tilde{\beta}_i z_j + \epsilon_{ij}\}$$

where α_i is the marginal utility of income, $\tilde{\beta}_i$ is the weight the consumer places on the attribute z_j and ϵ_{ij} is a random shock to utility. Suppose $\tilde{\beta}_i$ is an over-estimate of the consumer's underlying preferences for attribute z_j; the true weight is $\beta_i < \tilde{\beta}_i$ and the "true" utility from product j is $U_{ij} = \alpha_i(y_i - p_j) + \beta_i z_j + \epsilon_{ij}$.

Define the expected value of the consumer's "decision" utility (i.e. the function the consumer optimizes when making consumption decisions), $\tilde{v}_i(\mathbf{p}) = \mathbb{E}_{\boldsymbol{\epsilon}}[\tilde{U}_{ij^*}]$ where $j^* = \arg\max\{\tilde{U}_{ij}\}$ denotes the product she selects. The consumer level expected demand (probability) for product j is $q_{ij}(\mathbf{p}) = \mathbb{P}(\tilde{U}_{ij} > \tilde{U}_{ik} \quad \forall \quad k \neq j)$. The consumer's expected utility takes the form:

$$\begin{aligned} V_i(\mathbf{p}) = \mathbb{E}_{\boldsymbol{\epsilon}}[U_{ij^*}] =& \mathbb{E}_{\boldsymbol{\epsilon}}[\tilde{U}_{ij^*}] - \mathbb{E}_{\boldsymbol{\epsilon}}(\tilde{\beta}_i - \beta_i) z_{j^*} \\ =& \tilde{v}_i(\mathbf{p}) - \sum_{j\in\Omega}(\tilde{\beta}_i - \beta_i) z_j q_{ij}(\mathbf{p}) \end{aligned}$$

Relabelling $\tilde{v}_i(.) = v_i(.)$ and $\varphi_i = (\tilde{\beta}_i - \beta_i)$, we have $V_i(\mathbf{p}) = v_i(\mathbf{p}) - \sum_{j\in\Omega} \varphi_i z_j q_{ij}$. The marginal internality is given by the size of the consumer's overestimate of their preference for attribute z_j.

B Additional data tables

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

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.

In Table B.2 we list the firms and brands in the market, as well as 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: *Brands, sugar contents and sizes*

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

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

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:

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

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 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:

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

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 and $\text{packtypeswitch}_{i\tau}$ is a dummy variable equal to 1 if the household purchased a pack type that is different from the pack type they bought last.

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% 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.

Table C.3: *Sales and product switching*

	(1) Brand switch	(2) Size switch	(3) Pack type switch
Purchase on sale?	0.0181***	0.0241***	0.0160***
	(0.0012)	(0.0012)	(0.0007)
Mean of dependent variable	0.5432	0.5221	0.1272
N	1823157	1823157	1823157
Time effects?	Yes	Yes	Yes
Retailer effects?	Yes	Yes	Yes
Decision maker fixed effects?	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. Robust standard errors are shown in parentheses.

To summarize, we find limited evidence of stockpiling behavior in our data; although we cannot conclusively rule it out, the any effects appear to be very 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. In the at-home segment we allow preferences over price, branded soft drinks, store brand soft drinks and fruit juice to vary systematically with whether the household has above or below median equivalized household income. The positive coefficients on the interaction of price with low income implies those on low incomes are systematically more sensitive to price.

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.

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	0.227	0.257	0.169	0.261	0.247	0.284
		(0.045)	(0.040)	(0.044)	(0.036)	(0.035)	(0.033)
	Sugar medium	0.683	0.884	0.727	0.545	0.822	0.853
		(0.092)	(0.088)	(0.085)	(0.072)	(0.068)	(0.065)
	Sugar high	-0.045	0.516	0.692	-0.212	0.131	0.589
		(0.064)	(0.063)	(0.062)	(0.050)	(0.047)	(0.046)
	Advertising	0.346	0.265	0.336	0.313	0.355	0.335
		(0.057)	(0.055)	(0.052)	(0.045)	(0.041)	(0.039)
Interaction with low income	× Price	0.129	0.110	0.125	0.181	0.180	0.120
		(0.051)	(0.045)	(0.044)	(0.041)	(0.039)	(0.038)
	× Branded soft drinks	0.233	0.136	0.256	0.308	0.363	0.182
		(0.112)	(0.103)	(0.099)	(0.090)	(0.086)	(0.087)
	× Store soft drinks	0.167	0.509	0.320	0.593	0.516	0.403
		(0.123)	(0.129)	(0.121)	(0.118)	(0.103)	(0.113)
	× Fruit juice	-0.152	-0.437	-0.463	-0.339	-0.141	-0.323
		(0.167)	(0.149)	(0.162)	(0.132)	(0.127)	(0.140)
Variance	Price	0.116	0.075	0.150	0.074	0.123	0.109
		(0.019)	(0.013)	(0.023)	(0.012)	(0.012)	(0.012)
	Sugary	2.248	2.255	1.993	1.524	1.572	1.464
		(0.209)	(0.197)	(0.169)	(0.128)	(0.122)	(0.113)
	Drinks	2.211	2.659	1.706	1.572	1.412	1.390
		(0.191)	(0.212)	(0.200)	(0.141)	(0.126)	(0.130)
	Large	0.888	0.989	0.425	0.670	0.708	0.487
		(0.200)	(0.163)	(0.139)	(0.125)	(0.130)	(0.117)
	Cola	2.063	1.499	2.674	1.504	1.743	1.476
		(0.274)	(0.211)	(0.288)	(0.190)	(0.174)	(0.146)
	Lemonade	4.544	2.560	1.595	2.166	1.833	1.623
		(0.713)	(0.428)	(0.381)	(0.423)	(0.375)	(0.278)
	Store soft drinks	2.577	2.995	1.873	2.481	1.688	2.388
		(0.229)	(0.248)	(0.194)	(0.208)	(0.146)	(0.195)
	Fruit juice	3.318	2.925	3.826	2.324	2.242	2.907
		(0.340)	(0.279)	(0.350)	(0.209)	(0.203)	(0.261)
Covariance	Sugary-Drinks	-1.585	-1.801	-1.112	-1.136	-1.051	-0.878
		(0.171)	(0.173)	(0.145)	(0.116)	(0.109)	(0.098)
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	1.482	1.171	0.344	0.939	1.221	1.044
		(0.080)	(0.067)	(0.258)	(0.089)	(0.044)	(0.057)
	Sugar medium	2.435	2.316	2.622	0.811	1.230	1.812
		(0.194)	(0.115)	(0.116)	(0.085)	(0.062)	(0.093)
	Sugar high	1.385	0.826	1.720	-0.249	0.079	0.711
		(0.092)	(0.055)	(0.065)	(0.053)	(0.034)	(0.043)
	Advertising	0.848	0.538	0.269	0.484	0.467	0.643
		(0.063)	(0.037)	(0.048)	(0.038)	(0.026)	(0.038)
Variance	Price	0.258	0.151	0.027	0.264	0.183	0.384
		(0.044)	(0.021)	(0.016)	(0.035)	(0.013)	(0.031)
	Sugary	7.241	4.318	7.396	9.429	8.280	6.432
		(0.491)	(0.205)	(0.389)	(0.418)	(0.266)	(0.256)
	Drinks	4.265	2.310	6.170	4.260	2.628	2.495
		(0.278)	(0.152)	(0.341)	(0.197)	(0.099)	(0.114)
	Large	3.359	3.995	4.533	5.864	3.299	3.754
		(0.235)	(0.194)	(0.220)	(0.223)	(0.102)	(0.141)
	Cola	6.152	3.234	3.110	7.073	6.426	6.207
		(0.444)	(0.146)	(0.176)	(0.314)	(0.215)	(0.215)
	Lemonade	4.814	1.527	4.611	1.184	1.139	5.618
		(0.457)	(0.182)	(0.359)	(0.203)	(0.100)	(0.417)
	Fruit juice	6.522	2.221	3.160	5.670	3.980	1.402
		(0.937)	(0.296)	(0.432)	(0.486)	(0.253)	(0.225)
	Flavored milk	6.132	3.123	3.748	7.212	2.241	0.208
		(0.975)	(0.322)	(0.419)	(0.955)	(0.302)	(0.091)
Covariance	Sugary-Drinks	-2.252	-2.684	-5.766	-5.066	-3.443	-2.231
		(0.298)	(0.174)	(0.338)	(0.233)	(0.176)	(0.141)
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	-2.394	0.022	0.022	0.007	0.023	0.005	0.035	0.024	0.022	0.021
	Diet 2x2l	0.012	-3.163	0.012	0.007	0.012	0.010	0.013	0.065	0.015	0.024
Capri Sun	10x200ml	0.011	0.011	-2.795	0.011	0.030	0.007	0.012	0.014	0.021	0.032
Schweppes Lemonade	Regular 2x2l	0.007	0.013	0.023	-2.423	0.023	0.006	0.008	0.015	0.023	0.057
Robinsons	Squash 1l	0.011	0.010	0.028	0.011	-1.387	0.008	0.014	0.012	0.025	0.029
	Fruit diet 1l	0.006	0.020	0.017	0.006	0.019	-1.440	0.007	0.024	0.019	0.018
Pepsi	Regular 2l	0.031	0.021	0.022	0.007	0.026	0.006	-1.444	0.027	0.022	0.020
	Max 2x2l	0.012	0.058	0.013	0.007	0.013	0.011	0.015	-2.592	0.014	0.023
Topicana	1l	0.006	0.008	0.013	0.007	0.016	0.005	0.007	0.009	-2.333	0.019
Lucozade	Regular 6x380ml	0.008	0.016	0.024	0.021	0.023	0.006	0.008	0.017	0.024	-2.961
Outside option		0.006	0.008	0.011	0.006	0.016	0.007	0.008	0.010	0.018	0.014

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	-2.236	0.160	0.063	0.044	0.044	0.088	0.277	0.080	0.029	0.027
	Diet 500ml	0.251	-2.631	0.025	0.017	0.018	0.037	0.077	0.261	0.012	0.010
Fanta	Regular 500ml	0.262	0.066	-2.886	0.120	0.115	0.229	0.083	0.037	0.071	0.058
Dr Pepper	Max 500ml	0.239	0.058	0.157	-2.728	0.117	0.207	0.095	0.040	0.086	0.054
Cherry Coke	Regular 500ml	0.250	0.064	0.156	0.121	-2.821	0.230	0.076	0.033	0.067	0.066
Oasis	Regular 500ml	0.243	0.064	0.151	0.105	0.112	-2.625	0.077	0.034	0.065	0.058
Pepsi	Regular 500ml	0.862	0.151	0.062	0.054	0.042	0.087	-2.525	0.103	0.033	0.027
	Regular 500ml	0.256	0.529	0.028	0.023	0.019	0.039	0.106	-2.641	0.015	0.011
Ribena	Regular 500ml	0.230	0.061	0.134	0.125	0.093	0.186	0.083	0.036	-2.550	0.050
Lucozade	Regular 330ml	0.131	0.031	0.068	0.049	0.057	0.102	0.043	0.016	0.031	-2.084
Outside option		0.072	0.060	0.029	0.021	0.021	0.041	0.027	0.034	0.012	0.045

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	1.14	0.47	0.67	0.51
			[0.45, 0.49]	[0.65, 0.69]	[0.50, 0.52]
	Capri Sun	1.17	0.61	0.57	0.47
			[0.59, 0.62]	[0.55, 0.59]	[0.46, 0.49]
	Innocent fruit juice	3.34	1.89	1.44	0.48
			[1.85, 1.95]	[1.39, 1.49]	[0.47, 0.49]
	Schweppes Lemonade	0.52	0.16	0.35	0.68
			[0.15, 0.18]	[0.34, 0.37]	[0.66, 0.70]
	Fanta	1.44	0.56	0.88	0.61
			[0.53, 0.60]	[0.85, 0.91]	[0.59, 0.62]
	Dr Pepper	1.33	0.51	0.82	0.61
			[0.48, 0.55]	[0.78, 0.85]	[0.59, 0.62]
	Schweppes Tonic	1.65	0.90	0.75	0.54
			[0.87, 0.93]	[0.72, 0.78]	[0.52, 0.56]
	Sprite	1.26	0.45	0.81	0.64
			[0.43, 0.48]	[0.78, 0.83]	[0.63, 0.66]
	Cherry Coke	1.53	0.64	0.89	0.54
			[0.61, 0.68]	[0.85, 0.92]	[0.52, 0.56]
	Oasis	2.31	0.85	1.47	0.63
			[0.75, 0.95]	[1.36, 1.56]	[0.59, 0.68]
Pepsico/Britvic	Robinsons	1.20	0.40	0.80	0.67
			[0.38, 0.43]	[0.77, 0.82]	[0.66, 0.69]
	Pepsi	1.02	0.46	0.56	0.56
			[0.44, 0.47]	[0.54, 0.58]	[0.55, 0.58]
	Tropicana fruit juice	2.20	1.18	1.02	0.46
			[1.15, 1.21]	[0.99, 1.05]	[0.45, 0.48]
	Robinsons Fruit Shoot	1.80	0.76	1.05	0.58
			[0.72, 0.79]	[1.02, 1.08]	[0.56, 0.59]
	Britvic fruit juice	2.05	1.08	0.97	0.47
			[1.05, 1.12]	[0.94, 1.00]	[0.46, 0.49]
	7 Up	1.22	0.53	0.69	0.61
			[0.51, 0.56]	[0.66, 0.71]	[0.60, 0.63]
	Copella fruit juice	1.40	0.36	1.03	0.73
			[0.33, 0.40]	[1.00, 1.07]	[0.70, 0.75]
	Tango	1.13	0.42	0.72	0.67
			[0.39, 0.45]	[0.69, 0.74]	[0.65, 0.69]
GSK	Ribena	1.77	0.96	0.82	0.46
			[0.93, 0.99]	[0.78, 0.85]	[0.44, 0.47]
	Lucozade	1.62	0.86	0.76	0.48
			[0.84, 0.90]	[0.72, 0.78]	[0.46, 0.49]
	Lucozade Sport	1.49	0.90	0.59	0.39
			[0.88, 0.93]	[0.56, 0.61]	[0.38, 0.40]
	Vimto	1.09	0.55	0.54	0.52
			[0.53, 0.56]	[0.53, 0.56]	[0.51, 0.53]
	Irn Bru	1.56	0.77	0.78	0.53
			[0.73, 0.82]	[0.73, 0.82]	[0.50, 0.55]
Merrydown	Shloer	1.59	0.80	0.80	0.50
			[0.77, 0.83]	[0.77, 0.82]	[0.48, 0.52]
Red Bull	Red Bull	4.74	3.08	1.66	0.34
			[2.99, 3.19]	[1.55, 1.75]	[0.32, 0.36]
Total		1.44	0.67	0.77	0.55
			[0.65, 0.69]	[0.74, 0.79]	[0.53, 0.56]

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.

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).[56] 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 fruit juices are exempt from the tax.

We define three sets of products. First, the "higher rate treatment group" are those products with at least 8g of sugar per 100ml,at the time the tax was introduced and therefore are subject to the higher tax rate. Second, the "lower rate treatment group" are those products that have 5-8g of sugar per 100ml, and therefore are subject to the lower tax rate. The remaining set of products are exempt, either because their sugar content is less than 5g per 100ml, or because they are milk-based or fruit juice. There was some reformulation in anticipation of the introduction of the SDIL. We categorize products based on the post reformulation sugar contents.[57]

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:

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

[56] 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.

[57] 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.

where p_{jrt} denotes the price per liter of product j in retailer r in week t,[58] τ_t are week effects, ξ_j are product fixed effects, and ρ_r are retailer fixed effects.

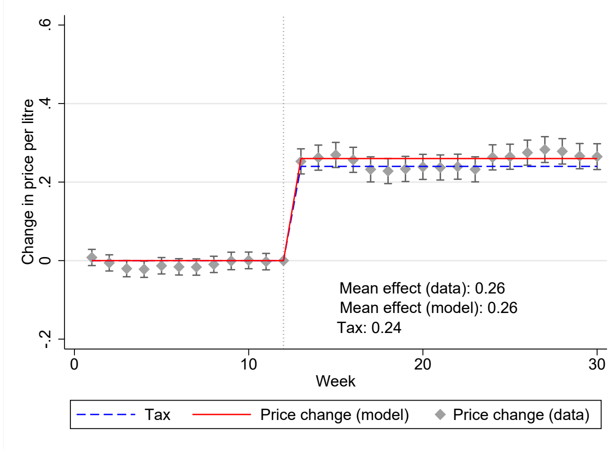
Figure E.1(a) plots the estimated price changes, relative to the week preceding the introduction of the tax, for the higher rate treatment group ($= \hat{\beta}^{hi} \times \text{Post}_t + \sum_{t \neq 12} \hat{\tau}_t$). Figure E.1(b) plots the analogous estimates for the lower rate treatment group ($= \hat{\beta}^{lo} \times \text{Post}_t + \sum_{t \neq 12} \hat{\tau}_t$). 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.

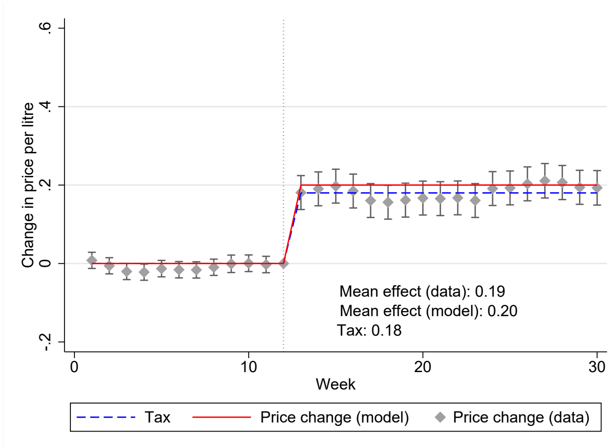
[58] This is the VAT-exclusive price per liter.

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

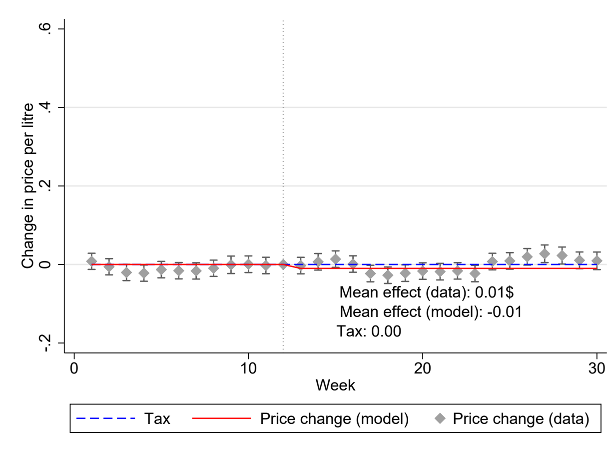
(a) High treatment group



(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 15% fall 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. This means the 15% fall in sugar sweetened beverage consumption translates into a fall in calories from these products of (0.56-0.47)*170=15kcal per adult per day.[59] 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 £0.04 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 £0.05 per 10g of sugar for this group.

F.2 Solving for the tax rate

The optimal, Pigouvian and Buchanan tax rates all take the form $\tau^* = G(\tau^*)$ (see equations (5.1)-(5.3)). In each case we solve for the tax rate τ^*, by the following procedure:

1. Guess a tax rate, τ^r
2. Find the solution to the system of firm first order conditions, given by equation (4.6), $\mathbf{p}(\tau^r)$

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

3. Find the solution to the systems of tax derivatives of firms' first order conditions, given by equation (F.1), $\frac{d\mathbf{p}(\tau^r)}{d\tau}$

4. Compute $G(\tau^r)$

5. Set $\tau^{r+1} = G(\tau^r)$ and repeat until convergence.

The derivative of firms' first order conditions with respect to the tax rate is given by: $\forall\ j$

$$\sum_{j'} \frac{\partial q_{jm}}{\partial p_{j'm}} \frac{dp_{j'm}}{d\tau} + \sum_{j' \in \mathcal{J}_f} \left(\frac{dp_{j'm}}{d\tau} - z_{j'} \right) \frac{\partial q_{j'm}}{\partial p_{jm}} + \sum_{j' \in \mathcal{J}_f} (p_{j'm} - \tau z_{j'} - c_{j'm}) \sum_{j''} \frac{\partial^2 q_{j'm}}{\partial p_{jm} \partial p_{j''m}} \frac{dp_{j''m}}{d\tau} = 0. \tag{F.1}$$

Note an alternative way of solving for the optimal tax rate is, instead of solving $\tau^* = G(\tau^*)$, solve the planner's first order condition, equation (A.1). For each trial rate it remains necessary to solve for the associated equilibrium prices and the derivative of equilibrium prices with respect to the tax. This alternative procedure results in a numerical value for the optimal tax rate that is the same (up to the tolerance threshold of 1e-4) as the above procedure.

F.3 Additional tables

Table F.1 summarizes the impact of the optimal tax rate on purchases and prices, separately for sugar sweetened beverages, zero sugar soft drinks and alternative sugary drinks (i.e. fruit juice and flavored milks).

Table F.2 replicates information in Table 5.7 and additionally reports 95% confidence intervals.

Table F.1: *Impact of optimal tax rate on purchases*

	Sugar sweetened beverages	Zero sugar soft drinks	Alternative (sugary) drinks
%Δ price change	17.9	-0.5	-0.1
	[17.6, 18.3]	[-0.6, -0.5]	[-0.2, -0.1]
%Δ purchase probability	-17.3	7.9	7.8
	[-17.8, -17.0]	[7.7, 8.3]	[7.5, 8.2]
conditional on purchase:			
%Δ volume	-15.6	1.6	-0.1
	[-16.3, -15.1]	[1.4, 1.8]	[-0.2, -0.0]
%Δ sugar intensity	-3.1		0.0
	[-3.3, -3.0]		[0.0, 0.0]
%Δ sugar	-33.7		7.7
	[-34.7, -33.0]		[7.4, 8.1]

Notes: Price changes refers to average change across products, weighted using pre-tax market share. Numbers in the second panel are averages across consumers. Numbers are reported for optimal rate, τ^*. 95% confidence intervals are given in square brackets.

Table F.2: *Impact of tax policy under alternative market structures*

	No tax	Tax					
		Optimizing		Pigouvian		Buchanan	
	Welfare	τ^*	ΔW	τ^P	ΔW	τ^B	ΔW
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Observed ownership	–	1.89	166	4.24	51	0.78	117
		[1.83, 1.97]	[148, 188]	[4.19, 4.28]	[25, 79]	[0.67, 0.85]	[108, 130]
Counterfactual ownership:							
$\{SPF\}$	423	2.24	242	4.24	158	1.32	211
	[385, 442]	[2.19, 2.32]	[221, 273]	[4.19, 4.29]	[131, 192]	[1.23, 1.39]	[195, 234]
$\{JPM^{\mathcal{S}}, SPF^{\mathcal{N}}\}$	290	1.61	104	4.24	-26	0.83	83
	[238, 318]	[1.53, 1.69]	[90, 120]	[4.18, 4.28]	[-50, -6]	[0.72, 0.89]	[74, 93]
$\{SPF^{\mathcal{S}}, JPM^{\mathcal{N}}\}$	-197	2.22	248	4.24	156	1.37	220
	[-228, -179]	[2.16, 2.32]	[227, 280]	[4.19, 4.29]	[129, 192]	[1.29, 1.45]	[204, 245]
$\{JPM\}$	-1545	1.31	69	4.24	-93	-0.17	-22
	[-1665, -1428]	[1.16, 1.42]	[54, 81]	[4.19, 4.28]	[-127, -67]	[-0.35, -0.09]	[-24, -19]

Notes: Table replicates information in Table 5.7 and additionally displays 95% confidence intervals.