DISCUSSION PAPER SERIES

DP14582

CORRECTIVE TAX DESIGN AND MARKET POWER

Martin O'Connell and Kate Smith

INDUSTRIAL ORGANIZATION

PUBLIC ECONOMICS

ISSN 0265-8003

CORRECTIVE TAX DESIGN AND MARKET POWER

Martin O'Connell and Kate Smith

Discussion Paper DP14582
Published 08 April 2020
Submitted 04 April 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Industrial Organization
- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Martin O'Connell and Kate Smith

CORRECTIVE TAX DESIGN AND MARKET POWER

Abstract

We study the design of taxes aimed at limiting externalities in markets characterized by differentiated products and imperfect competition. In such settings policy must balance distortions from externalities with those associated with the exercise of market power; the optimal tax rate depends on the nature of external harms, how the degree of market power among externality generating products compares with non-taxed alternatives, and how consumers switch across these products. We apply the framework to taxation of sugar sweetened beverages. We use detailed data on the UK market for drinks to estimate consumer demand and oligopoly pricing for the differentiated products in the market. We show the welfare maximizing tax rate leads to welfare improvements over 2.5 times as large as that associated with policy that ignores distortions associated with the exercise of market power.

JEL Classification: D12, D43, D62, H21, H23, L13

Keywords: externality, corrective tax, market power, oligopoly

Martin O'Connell - martin_o@ifs.org.uk
IFS/UCL and CEPR

Kate Smith - kate_s@ifs.org.uk
IFS/UCL

Corrective Tax Design and Market Power

Martin O'Connell and Kate Smith*

April, 2020

Abstract

We study the design of taxes aimed at limiting externalities in markets characterized by differentiated products and imperfect competition. In such settings policy must balance distortions from externalities with those associated with the exercise of market power; the optimal tax rate depends on the nature of external harms, how the degree of market power among externality generating products compares with non-taxed alternatives, and how consumers switch across these products. We apply the framework to taxation of sugar sweetened beverages. We use detailed data on the UK market for drinks to estimate consumer demand and oligopoly pricing for the differentiated products in the market. We show the welfare maximizing tax rate leads to welfare improvements over 2.5 times as large as that associated with policy that ignores distortions associated with the exercise of market power.

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, UIUC, WashU 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. Correspondence: martin_o@ifs.org.uk and kate.smith@ifs.org.uk.

*Institute for Fiscal Studies and University College London.

1 Introduction

One-fifth of all consumer spending is undertaken in markets subject to taxes aimed, at least in part, at altering behavior to limit externalities.[1] Many of these markets are characterized by the presence of large multi-product firms that are likely to exercise substantial market power. For instance, soft drink markets, recently the subject of new taxes in several jurisdictions, are dominated by Coca Cola Enterprises and PepsiCo. Much of the long literature on the design of taxes to correct for externalities, dating back to Pigou (1920), assumes a perfectly competitive environment, and yet, as argued by Buchanan (1969), in imperfectly competitive environments price is already in excess of marginal cost and externality correcting policy that fails to take account of this can reduce welfare.

Our contribution is to study the design of corrective taxes in markets characterized by differentiated products and strategic firms, and to undertake a substantive empirical application to the taxation of sugar sweetened beverages. We set out a simple optimal tax model that shows how patterns of consumer substitution, positive price-cost margins and strategic price re-optimization affect the optimal corrective tax prescription. In our empirical application we estimate a detailed model of consumer demand and oligopoly price competition in the market for drinks, and compute the optimal sugar sweetened beverage tax. We show that despite substantial price-cost margins on these products, there is a case for levying a positive tax rate; in part, because consumers switch to other products with high margins. Nevertheless, the optimal rate lies below the rate a planner that ignores distortions associated with the exercise of market power would set.

Corrective taxes are common and marker power is pervasive, yet there is relatively little work that empirically investigates the design of these taxes in noncompetitive environments. An important exception is Fowlie et al. (2016), who study policy to reduce carbon emissions in the US cement market. In a setting in which firms produce a homogeneous good and derive market power from high entry costs, they show that policy that aims for full carbon abatement is welfare reducing and inferior to policy that recognizes pre-existing distortions from the exercise of market power. We focus on an environment in which firms derive market power from selling differentiated products, and highlight the importance of how the relative degree of market power and external harms across products, and multi-product firms' portfolio effects, impact optimal tax design.

[1] Spending on alcohol, tobacco, soft drinks, fuel, and motoring (all of which are subject to some kind of excise duty in the UK – Levell et al. (2016)) accounts for 24% of spending recorded in the UK's consumer expenditure survey (Living Costs and Food Survey (2017)).

In particular, we consider a setting in which there are many differentiated products available to consumers. The consumption of one set of products generates an externality (in proportion to some specific product attribute), while the remaining set generate no external costs. The products are supplied by a set of (potentially multi-product) firms that derive market power from the imperfectly substitutable nature of the products in the market. A social planner sets a linear tax on the externality generating product attribute, with the aim of improving welfare. To focus on the interaction between externality correction and imperfect competition we assume the planner sets the tax rate to maximize economic efficiency in the market; the planner does not have a redistributive motive,[2] nor a revenue raising constraint.[3] If the market was perfectly competitive, the optimal rate would be equal to the marginal external cost (if homogeneous across consumers, as in Pigou (1920)) or, when there is heterogeneity in marginal externalities, the optimal rate would be equal to a weighted average of marginal external costs (Diamond (1973)).

Under imperfect competition the optimal tax rate equals the traditional corrective component minus an adjustment for the distortion associated with the exercise of market power. In a market with just one product supplied by a monopolist, the optimal rate is equal to the marginal externality minus the equilibrium price-cost margin on the product. In a two product market (where one product is associated with externalities and one is not), the planner cares both about achieving an efficient level of total consumption, and achieving allocative efficiency across the two products. A higher equilibrium price-cost margin on the externality generating product acts to reduce the optimal rate, while a higher margin on the substitute (untaxed) product acts to increase it. The extent to which the margin on the non-externality generating product raises the optimal rate depends on how strongly the tax shifts consumption towards it from the taxed product; in the limit, if consumption switches one-for-one between the products, the optimal tax rate equals

[2]Allcott et al. (2019a) study the optimal design of tax on products that generate externalities and internalities under perfect competition, when the social planner has a preference for redistribution. They show pro-redistribution preferences impact the optimal rate through two channels; (i) if internalities are concentrated among the poor this raises the optimal rate, but (ii) if the poor have a strong preference for the product it lowers the optimal rate as this preference acts as a tag for ability.

[3]Sandmo (1975) shows that in the face of a revenue raising constraint, an efficiency maximizing planner that can set a linear tax on each product in the economy will set a tax rate on an externality generating good that entails a Pigovian component plus a distortionary Ramsey component. Kopczuk (2003) shows this additivity property holds under much more general conditions, including when there are redistributive motives. See Bovenberg and Goulder (2002) for a thorough review of work on how the interaction between corrective taxes and other distortionary taxes changes the Pigovian tax prescription and can limit the effectiveness of externality correcting taxation.

the marginal externality minus the difference in equilibrium price-cost margins between the two products. With many differentiated products, switching within the set of taxed products, as well as the alternative products that consumers switch most strongly towards, also influences the optimal rate.

We use the framework to study 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; as of December 2019, 43 countries and 8 US cities had some form of sugar sweetened beverage tax in place (GFRP (2019)). The market for these products is characterized by large multi-product firms that offer strongly branded products and are likely to enjoy significant market power.

It is common in modern public economics to use sufficient statistics to assess the welfare consequences of policy reforms (Chetty (2009)). This approach is used by Jacobsen et al. (2018) to quantify the welfare loss associated with the inability to levy product-specific Pigovian taxes, by Allcott et al. (2019a) who consider the optimal sin tax for sugary drinks when consumers misoptimize and the planner has preferences for redistribution, and by Ganapati et al. (2019) to measure incidence of input taxes in imperfectly competitive markets. In our setting, the welfare effects of changing the tax rate depend on consumer substitution between, and price-cost margins (which are unobservable – see Bresnahan (1989)) of, a large set of (asymmetrically) differentiated products. We specify a model of demand and supply in the market, which enables us to estimate elasticities and price-cost margins for disaggregate products and to simulate the effect of non-local tax changes, and therefore recover the optimal tax rate. To provide evidence that our empirical model successfully captures behavior in the market, we use quasi-experimental variation in price changes resulting from the recent introduction of the UK's sugar sweetened beverage tax to validate our estimated model.

We use longitudinal data from the UK on purchases of non-alcoholic drinks that households bring into the home and that individuals consume while on-the-go. Most empirical studies of sugar sweetened beverage taxes do not cover purchases made on-the-go, yet they are an important part of the market.[4] We obtain demand elasticities by estimating a model of consumer choice among the differentiated products in the drinks market (in the broad spirit of Berry et al. (1995)). We model pref-

[4]An exception is Dubois et al. (2020), who focus on modeling on-the-go demand for drinks in the UK market. Their interest is identifying which groups of individuals would lower their sugar intake most strongly in response to increases in the prices of on-the-go soft drinks.

erences over key product attributes as random coefficients, allowing the coefficient distributions to depend on consumer age, income and a measure of total dietary sugar. The overall preference distribution takes the flexible form of a mixture of normal distributions, relaxing functional form restrictions otherwise imposed on product demand curves.[5,6]

We use our demand estimates and the equilibrium conditions of an oligopoly pricing game to infer marginal costs (as, for instance, in Nevo (2001)). Our estimates suggest that, on average, prices are around double marginal costs, though there is considerable variation in price-cost margins across products. In particular, small pack sizes typically have larger price-cost margins (per liter) than bigger sizes. Our demand estimates suggest consumers switch more strongly away from large sizes in response to a tax, meaning a tax leads consumers to tilt their baskets of taxed products towards those with high margins.

We find that if the externality from the consumption of sugar from sweetened beverages is greater than £2.15 per kg of sugar (approximately 0.8¢/oz of product) then the optimal tax rate is positive. This is despite these products having substantial price-cost margins that are exacerbated by the tax – both by people switching to small high margin products, and firms optimally raising their margins by increasing prices by more than the tax. The positive optimal rate is driven by consumers switching towards alternative drinks products that are also supplied non-competitively. However, the optimal rate lies below the rate that a planner would set if they ignored distortions associated with the exercise of market power. If externalities are also generated by substitute goods that contain sugar, a tax on the sugar in sweetened beverages will be less effective at combating externalities. We show that the existence of these untaxed external costs leads to a reduction in the optimal rate of around 20%. We also show that the optimal rate increases modestly in the extent to which externalities are concentrated among those with high overall dietary sugar.

Our results highlight that tax policy that ignores distortions associated with the exercise of market power will lead to significant unrealized welfare gains; for a central estimate of the externality from sugar, optimal policy results in gains 2.5

[5]In particular, the flexible preference distribution helps relax curvature restrictions on demands. As highlighted by Weyl and Fabinger (2013), demand curvature is one important determinant of how equilibrium prices respond to tax changes.

[6]A potential threat to the validity of our demand estimates is the presence of neglected dynamics. In particular there is evidence in the US that consumers stockpile soft drinks (Hendel and Nevo (2006b), Wang (2015)). We provide evidence that stockpiling is much less relevant in the UK context; when consumers purchase on sale they tend to switch brands or pack type, with no evidence of significant changes in the timing of purchase.

times as large as those that ignore market power distortions. Our results add to the small, but growing, literature that uses empirically rich treatments of markets to evaluate how imperfect competition affects fundamental tax design questions. For example, Miravete et al. (2018a, 2018b) show that the peak and shape of the Laffer curve associated with an ad valorem tax rate depends on the strategic pricing behavior of distillers, and quantify welfare gains that would be realized if government instead set product specific taxes/prices. A number of papers consider optimal subsidy design in health insurance markets in which providers exercise market power (see Tebaldi (2017), Polyakova and Ryan (2019) and Einav et al. (2019)) and show targeted subsidies engender equilibrium pricing responses and spillovers to non-targeted groups.

To show how the degree of market power exercised by firms influences the potential welfare gains from levying a tax on externality generating products, we simulate the optimally set tax under counterfactual firm ownership structures. A more competitive market structure leads to welfare gains (in the absence of tax), as increases in consumer surplus swamp reductions in firm profitability and increased externalities. In addition, a tax on externality generating goods leads to larger welfare gains under more competitive market structures, pointing towards a complementarity between competition and corrective tax policy.

Almost all jurisdictions that have introduced taxes on sugar sweetened beverages do so on a volumetric basis, rather than in proportion to sugar content. There is a rapidly growing literature that uses the implementation of these taxes to study the effects on prices and quantities.[7] A complementary set of papers use estimates of consumer demand based on periods and locations with no tax in place to simulate the introduction of taxes similar to those used in practice.[8] We contribute to this literature by comparing the performance of the optimal tax on sugar to a number of more common tax structures. We find that an optimally set volumetric tax achieves only 60% of the welfare gains achieved by the optimally set sugar tax rate. Some localities, notably Philadelphia, apply a volumetric tax to both artificially and sugar sweetened beverages as a revenue raising measure; we show that it is much more costly in welfare terms to raise revenue with this instrument compared to a tax levied only on sugar sweetened beverages.

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

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

The rest of this paper is structured as follows. In Section 2 we consider the design of corrective taxes in markets, such as that for drinks products, in which firms set prices above marginal costs. Section 3 describes the UK market for drinks and the micro panel data we use on purchase decisions made for consumption outside as well as in the home. In Section 4 we present our empirical model of consumer demand and firm pricing competition. Section 5 presents our empirical tax results. A final section draws together the implications of our results and concludes.

2 Corrective tax design in imperfect competition

Our aim is to highlight how distortions associated with the exercise of market power influence the efficiency maximizing rate of tax on externality generating products. We consider a market that comprises a set of differentiated products, a subset of which have externalities associated with their usage. The products are provided by firms who set their prices under conditions of imperfect competition. We begin by considering a stylized market in which there are just two products, before generalizing the analysis to a market with many products.

We consider a social planner whose task it is to set a tax rate for the externality generating goods. The planner's objective is to maximize efficiency. We abstract from possible redistributive motives, focusing instead on how imperfect competition alters the optimal externality correcting tax prescription.[9]

2.1 A two product market

Set-up. Consider a market that comprises two products, $j = \{1, 2\}$. Consumer i, facing prices, $\mathbf{p} = (p_1, p_2)$, chooses how to allocate her income, y_i, between the two 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 either product does not induce any income effects. This assumption is reasonable when focusing on a market that accounts

[9]Under perfect competition and when 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 and consumers are utility maximizing (Saez (2002)). Jaravel and Olivi (2019) show that this extends to an economy characterized by imperfect competition. Kaplow (2012) shows that accompanying externality correcting taxes with a distribution-neutral adjustment to the income tax system can offset the redistributive effects of the corrective taxes across the income distribution.

for a small share of total consumer spending,[10] and it assists with focusing on a planner that seeks to maximize economic efficiency. We denote market level demand for product j by $q_j(\mathbf{p}) = \sum_i q_{ij}(\mathbf{p})$ and total consumer surplus from participation in the market by $v(\mathbf{p}) = \sum_i v_i(\mathbf{p})$

Each unit of product 1 consumed creates an externality. We initially assume the externality is homogeneous across individuals and denote it by ϕ. Product 2 is a substitute for product 1; its consumption does not create any externalities. A social planner chooses the rate of tax, τ, to set on product 1. Both products are supplied imperfectly competitively at constant marginal cost; equilibrium prices are such that:

$$\begin{aligned} p_1 - \tau - c_1 &= \mu_1 \\ p_2 - c_2 &= \mu_2, \end{aligned}$$

where c_j denotes the marginal cost and μ_j denotes the equilibrium price-cost margin for product j (per unit, for instance liter, of consumption). The equilibrium prices and margins depend on the rate of tax levied on product 1 (as well as the marginal costs of both products). They also depend on whether the products are supplied by duopolists or a monopolist.[11] For notational simplicity we suppress the dependence of prices and margins on the tax rate, while in this section we remain agnostic about the product ownership structure.

We assume that the numeraire is competitively supplied, and its consumption does not generate any externalities. We relax each of these assumptions when we empirically implement our results in Section 5.

Optimal policy. We consider a social planner that chooses the rate of tax to maximize total welfare, which equals the total consumer surplus from participation in the market, $v(\mathbf{p})$, minus total externalities plus tax inclusive profits. Tax inclusive profits on product 1 are given by $(p_1 - c_1)q_1$ and are equal to the sum of net profits $(p - \tau - c_1)q_1$ and tax revenue, τq_1. The planner's problem is:

$$\max_{\tau} v(\mathbf{p}) - \phi q_1 + (p_1 - c_1)q_1 + (p_2 - c_2)q_2. \tag{2.1}$$

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

[11] For instance, if the two products are supplied by separate firms that compete in a Bertrand game $\mu_j = -q_j(\mathbf{p})/\frac{\partial q_j(\mathbf{p})}{\partial p_j}$. Solving the two optimal pricing equations yields equilibrium prices $(p_1(\tau), p_2(\tau))$ (where we suppress the dependence of prices on marginal cost), and associated margins $(\mu_1(\tau), \mu_2(\tau))$.

The optimal tax rate, τ^*, is implicitly defined by:

$$\tau^* = \phi - \left(\mu_1 - \mu_2 \times \frac{dq_2}{d\tau} \Big/ \left(-\frac{dq_1}{d\tau}\right)\right), \tag{2.2}$$

where $\frac{dq_j}{d\tau} = \frac{\partial q_j}{\partial p_1}\frac{dp_1}{d\tau} + \frac{\partial q_j}{\partial p_2}\frac{dp_2}{d\tau}$ is the derivative of equilibrium consumption of product j with respect to the tax. We expect $\frac{dq_1}{d\tau} < 0$ and, as the goods are substitutes, $\frac{dq_2}{d\tau} > 0$. We refer to the expression $\frac{dq_2}{d\tau}/\left(-\frac{dq_1}{d\tau}\right)$ as the switching ratio; it captures the extent to which any reduction in equilibrium consumption of the externality generating product induced by a marginal increase in the tax rate is redirected towards the substitute good.

When the two products are supplied competitively (so $\mu_j = 0$ for $j = \{1,2\}$ regardless of the level of τ) the optimal policy is a Pigovian tax ($\tau^* = \phi$) and the first best is achieved. Whenever the products are supplied under imperfect competition, the optimal tax rate is equal to the Pigovian rate plus an adjustment for non-competitive pricing.

Under imperfect competition it is instructive to consider two special cases. First, suppose demands for the two products are independent (i.e. $q_j(p_1,p_2) = q_j(p_j)$ for $j = \{1,2\}$), so $\frac{dq_2}{d\tau}/\left(-\frac{dq_1}{d\tau}\right)$=0 and the market for product 1 is a single-product monopoly market. In this case the optimal tax rate is (implicitly defined by) $\tau^* = \phi - \mu_1$, product 1 is priced at the efficient level, $p_1 = c_1 + \phi$, and the equilibrium price of product 2 is left unaffected by the tax. Second, suppose instead total consumption across the two good is fixed, so in response to price changes consumers only reallocate their demand between the two products, which implies $\frac{dq_2}{d\tau}/\left(-\frac{dq_1}{d\tau}\right) = 1$. In this case the optimal tax rate is $\tau^* = \phi - (\mu_1 - \mu_2)$ and the difference in equilibrium prices of the two products is $p_1 - p_2 = (c_1 - c_2) + \phi$. The tax achieves an efficient allocation (of the fixed consumption level) across the two products.

In practice, $\frac{dq_2}{d\tau}/\left(-\frac{dq_1}{d\tau}\right)$ is likely to lie somewhere between 0 and 1; the imperfect competition adjustment to the Pigovian tax rate partly reflects how policy changes total consumption in the market and partly how it influences the allocation of consumption across the two products. To see this, note that we can re-write equation (2.2) as $\tau^* = \phi - \left[\left(1 - \frac{dq_2}{d\tau}/\left(-\frac{dq_1}{d\tau}\right)\right)\mu_1 + \frac{dq_2}{d\tau}/\left(-\frac{dq_1}{d\tau}\right)(\mu_1 - \mu_2)\right]$. The more strongly the reduction in equilibrium quantity of product 1 from a marginal change in the tax rate is directed to product 2 (i.e. the closer $\frac{dq_2}{d\tau}/\left(-\frac{dq_1}{d\tau}\right)$ is to 1), the more weight is placed on the difference in equilibrium margins of the two goods.

2.2 Many differentiated products

In practice, corrective taxes are typically used in markets in which there are many differentiated products. To the extent that there is variation across the equilibrium price-cost margins of these products and in whether their consumption generates externalities,[12] this will influence the optimal tax prescription. In addition, it matters whether the tax is levied directly on the product characteristic that is associated with externalities, or whether the tax is levied on a per unit basis. For instance, a tax on sugar sweetened beverages can either be levied directly on sugar, or on a volumetric (i.e. per liter) basis.

Suppose there are many products $j = \{1, \dots, J\}$. 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 $j \in \mathcal{N}$), $z_j = 0$. Consider a tax levied on z. The products are supplied in an imperfectly competitive environment with equilibrium prices satisfying:

$$\begin{aligned} p_j - \tau z_j - c_j &= \mu_j \quad \forall j \in \mathcal{S} \\ p_j - c_j &= \mu_j \quad \forall j \in \mathcal{N}. \end{aligned}$$

In Appendix A we show that in this case the optimal tax rate can be expressed as follows:

Proposition 1. *Define: (i) the derivative of the total equilibrium quantity of the set of externality generating products with respect to the tax as* $\frac{dQ^{\mathcal{S}}}{d\tau} = \sum_{j \in \mathcal{S}} \frac{dq_j}{d\tau}$*, (ii) the share that product* $j \in \mathcal{S}$ *contributes to this derivative as* $w_j^{\mathcal{S}} = \frac{dq_j}{d\tau} / \frac{dQ^{\mathcal{S}}}{d\tau}$*, (iii) the analogous expressions for the set of products that do not generate externalities (i.e.* $\frac{dQ^{\mathcal{N}}}{d\tau} = \sum_{j \in \mathcal{N}} \frac{dq_j}{d\tau}$ *and* $w_j^{\mathcal{N}} = \frac{dq_j}{d\tau} / \frac{dQ^{\mathcal{N}}}{d\tau}$*), and (iv) the derivative of the total equilibrium quantity of the externality generating attribute with respect to the tax rate as* $\frac{dZ}{d\tau} = \sum_{j \in \mathcal{S}} z_j \frac{dq_j}{d\tau}$*. The optimal tax rate is then implicitly defined by:*

$$\tau^* = \phi - \frac{1}{\frac{dZ}{d\tau} / \frac{dQ^{\mathcal{S}}}{d\tau}} \left(\sum_{j \in \mathcal{S}} w_j^{\mathcal{S}} \mu_j - \sum_{j \in \mathcal{N}} w_j^{\mathcal{N}} \mu_j \times \frac{dQ^{\mathcal{N}}}{d\tau} \Big/ \left(-\frac{dQ^{\mathcal{S}}}{d\tau} \right) \right). \tag{2.3}$$

This expression generalizes the optimal tax formula in the two good case (equation (2.2)). Now the rate depends on the *weighted average* price-cost margin among the *sets of* externality and non-externality generating products. As the tax rate

[12] For instance, in the case of sugar sweetened beverages, a given amount of consumption of a product with 10g of sugar per 100ml, all else equal, is likely to be associated with more externalities than one with 5g sugar per 100ml.

varies, the average margin term 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/towards products with different equilibrium margins. Now, how much margins on non-taxed products raise the optimal rate depends on the extent to which any reduction in consumption of the *set of* externality generating product induced by a marginal increase in the tax rate is redirected towards the *set of* substitute goods. The many product optimal tax expression also depends on the ratio of the marginal change in equilibrium quantity of the externality generating attribute and equilibrium quantity of the externality generating goods with respect to the tax rate (i.e. $\frac{dZ}{d\tau}/\frac{dQ^{\mathcal{S}}}{d\tau}$).[13] This term results from the tax being levied on the externality generating product attribute rather than volumetrically on the externality generating products (see Appendix A for the expression for a volumetric tax).

2.3 Extensions

Heterogeneity in externalities. Marginal externalities may be heterogeneous, either because externalities depend non-linearly on an individual's total intake of the externality generating attribute or because, conditional on consumption, some individuals' intake is more problematic than others. In this case the externality component in equation (2.3), ϕ is replaced by the weighted average of consumer specific marginal externalities, $\sum_i \omega_i \phi_i$, where the weight, ω_i, is the contribution of individual i to the marginal change in the equilibrium quantity of the externality generating characteristic with respect to a marginal change in the tax rate (see Appendix A for the full expression). 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 expression $\sum_i \omega_i \phi_i$ takes a similar form as 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 ω_i incorporate the equilibrium pricing response of firms in the market.

Externality leakage. In some circumstances a policymaker may be restricted to set a tax on a subset of externality generating products, perhaps due to some

[13] In the case of a tax on the sugar in sweetened beverages, this captures the ratio of the marginal change in sugar consumption with respect to a small change in the tax over the marginal change in liters of sugar sweetened beverage consumption with respect to the tax.

political constraint.[14] In this case, the corrective component in equation (2.3), ϕ, is scaled by the ratio $\frac{dZ^{\mathcal{A}}}{d\tau}\Big/\frac{dZ^{\mathcal{S}}}{d\tau}$, where $\frac{dZ^{\mathcal{A}}}{d\tau}$ denotes the marginal reduction in the externality generating characteristic from taxed and untaxed products associated with an increase in the tax rate, and $\frac{dZ^{\mathcal{S}}}{d\tau}$ denotes the marginal reduction in the externality generating characteristic from taxed products only (the full expression is provided in Appendix A). If, in equilibrium, a marginal increase in the tax rate induces switching from taxed to untaxed products that create an externality, then $\frac{dZ^{\mathcal{A}}}{d\tau}\Big/\frac{dZ^{\mathcal{S}}}{d\tau} < 1$, and, all else equal, the optimal tax rate is lower.

Full externality internalization. A policymaker may choose to ignore the distortions associated with the exercise of market power, aiming instead at full externality internalization, relative to the market equilibrium with zero tax in place. One option is to set a Pigovian tax, $\tau = \phi$. However, the pricing response of firms can undermine this. For instance, in a single product market with equilibrium pass through that is 150%, a Pigovian tax will lead to prices rising by 50% than the marginal externality. In this case the planner can achieve full internalization of externalities by setting a tax rate of $\tau = \frac{\phi}{\rho}$, where ρ is the pass-through rate (defined as the change in the equilibrium consumer price divided by the tax). In Appendix A we formalize the problem a planner solves when aiming for full externality internalization, relative to the zero tax equilibrium quantities, and show that tax policy will depend on the weighted average pass-through rate across all taxed products, as well as the equilibrium margin adjustment on non-taxed alternatives.

Internalities. Corrective taxes are sometimes justified on the basis of the presence of internalities – costs consumers impose on themselves by making choices that fail to maximize their underlying utility (e.g. Gruber and Koszegi (2004) and O'Donoghue and Rabin (2006)). Internalities may arise for many reasons including consumer self-control problems, incorrect beliefs and inattention. Our framework accommodates internalities that lead to consumer welfare taking the form $v_i(\mathbf{p}) - \varphi_i \sum_j z_j q_{ij}$, where φ_i can be interpreted as the marginal internality. We show in Appendix A that if demand is generated from a discrete choice random utility model and internalities arise from consumers over-estimating their underlying preference for a particular attribute (z) when making consumption decisions, the expression for consumer welfare will take this form.

[14] A leading example is when a good can be imported tax-free (see Fowlie et al. (2016) who study greenhouse gas emissions leakage due to imported concrete). In the case of sugar sweetened beverage taxes, some legislators have argued for a broadening of the base to cover other sources of dietary sugars (for instance, see House of Commons Health Committee (2018)).

2.4 Empirical implementation

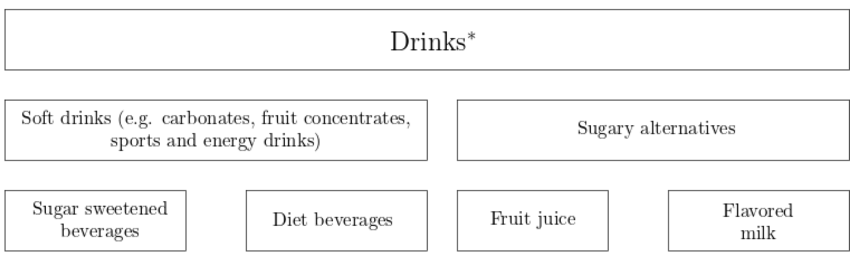
We apply our framework to the topical issue of the taxation of sugar sweetened beverages. We estimate consumer demand and firm competition in the UK market for non-alcoholic drinks; the model allows us to simulate equilibrium quantities (allowing for the endogenous response of prices) and price-cost margins for any given tax policy. We calibrate two key parameters over which there is considerable uncertainty: the magnitude of externalities from sugar sweetened beverages, and the degree of market power outside the drinks market.

Our analysis assumes that firms compete in their price setting, but hold fixed the portfolio of products they offer and non-price features of these products. A tax that is levied directly on the sugar content of products potentially incentivises firms to reduce the sugar content of some of their products to reduce tax liability (though this will depend on how this changes production costs and the strength of consumer preference for sugar). We return to this point when discussing our results in Section 6.

3 The drinks market

We model behavior in the UK market for drinks. Our market definition includes all chilled or ambient non-alcoholic beverages with the exception of water and unsweetened milk. Figure 3.1 shows a classification of drinks that we use to refer to different sets of products throughout the rest of the paper. We refer to one subset of the drinks as soft drinks. These include carbonates, fruit concentrates and sports and energy drinks. Soft drinks can be further divided into sugar sweetened beverages and diet (or artificially sweetened) beverages. We refer to those drinks that are not soft drinks as sugary alternatives. These include fruit juice and flavored milk; they are generally exempt from sugar sweetened beverage or soft drinks taxes.

Figure 3.1: *Drinks classification*



**drinks refers to all non-alcoholic drinks with the exception of water and unsweetened milk.*

3.1 Externalities from sugar sweetened beverages

There is considerable evidence that consumption of sugar sweetened beverages increases the risk of developing a number of diseases.[15] Sugar sweetened beverages are high in sugar and the sugar is in liquid form; this means it is digested quickly, which leads to spikes in insulin and a higher propensity to develop type II diabetes. Calories consumed in liquid form are also less likely to sate appetites, which means people are less likely to compensate for their intake with reduced calories from other sources and thus consumption of these drinks leads to weight gain. Sugar sweetened beverage intake is also associated with increased blood pressure and a higher risk of cardiovascular disease, as well as causing tooth decay.

The higher disease burden associated with sugar sweetened beverages leads to costs borne by people other than the person consuming the products (i.e. externalities). A central source of externalities are increased public costs of funding heath care systems. These can result from 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)). Wang et al. (2012) estimate that a 15% reduction in sugar sweetened beverage consumption in US would lead to a $17.1 billion saving in heath care costs over 10 years. Although a portion of this saving would be realized by the consumer themselves, this is likely to be small (for instance, Cawley and Meyerhoefer (2012) estimate 88% of the US medical costs of treating obesity are borne by third parties). These externalities have led many governments, including the UK (Scientific Advisory Committee on Nutrition (2015)), to specifically target reductions in the intake of sugar sweetened beverages.

There is also concern about high levels of added sugar (including from foods) in diet more broadly. The World Health Organization recommends average intake of added sugars should not exceed 10% of total dietary energy (World Health Organization (2015)), while the UK has adopted the more stringent target of 5%.[16] In our analysis of a sugar sweetened beverage tax, we allow for the possibility that the nature of externalities from sugar sweetened beverages interact with broader dietary

[15] Allcott et al. (2019b) provide a useful summary of the evidence. The Scientific Advisory Committee on Nutrition (2015) provide a thorough review of the medical literature.

[16] These targets are stated in terms of "free sugars", which are similar to added sugar but also include naturally occurring sugars in fruit juices and honey.

sugar, and we consider the implications for the optimal tax on these products if there are externalities created by switching to other markets.[17]

3.2 Purchase data

We use micro data on the grocery purchases of a sample of consumers living in Great Britain (i.e. the UK excluding Northern Ireland). The data contain information on household level purchases for home consumption ("at-home"), as well as purchases made by individuals for consumption outside of the home (i.e. "on-the-go"). On-the-go consumption is an important part of soft drink intake – accounting for 30% of total soft drink consumption and 40% of total sugar consumption from soft drinks.[18] Our data are collected by the market research firm Kantar and comprise two parts: the Kantar Worldpanel covers the at-home segment of the market and the Kantar On-The-Go Survey covers the on-the-go segment.

The Kantar Worldpanel contains details of all the grocery purchases (including food, drink, alcohol, toiletries, cleaning produce and pet foods) that are made and brought into the home by a representative sample of just over 30,000 British households from January 2008 to December 2012. Participating households use a hand held scanner to record all grocery purchases at the UPC level (i.e at the disaggregate level at which items are barcoded). 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 where the transaction took place, and transaction level prices.

The Kantar On-The-Go Survey is based on a random sample of just under 3000 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 do not cover bars and restaurants). The data contain details of the item 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.

[17]Allcott et al. (2019a) argue that sugar from soft drinks also gives rise to internalities. While our main focus is on externality correcting taxation, as outlined in Section 2.3, our optimal tax set-up can accommodate some forms of consumer misoptimization. In Section 5.3 we show how our results vary with the magnitude of marginal externalities; these number can be reinterpreted as how results vary with size of marginal internalities.

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

3.3 Consumers

We use the term consumer to refer to households in the at-home segment, and individuals in the on-the-go segment. 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 depending on consumer age and with a measure of the total sugar in the consumer's diet in the preceding year. This allows us to capture any differences in demand behavior along dimensions over which marginal externalities from sugar sweetened beverage intake might vary.

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.

Table 3.1 shows the groups into which we place consumers. In the at-home segment we split households based on whether there are any children (people aged under 18) in the household or not. In the on-the-go segment we separate individuals aged 30 and under from those aged above 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 purchase less than

10% of their calories from added sugar (corresponding to meeting the World Health Organization's guideline) 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".

3.4 Firms, brands and products

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.[19] The firms Coca Cola Enterprises and Pepsico/Britvic dominate the market, having a combined market share exceeding 65% in the at-home segment and close to 80% in the on-the-go segment. 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 (also known as 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.

The majority of soft drinks brands are available in sugar sweetened ("regular") and artificially sweetened ("diet" and/or "zero") variants. In Table B.1 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. This variation in sugar content means a tax levied directly on sugar will have different implications to one levied volumetrically (i.e. per liter of product sold).

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 units. 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 units of the same brand-variant most commonly involve 2, or sometimes 3, units of the same pack (or UPC) and are typically a consequence of multi-buy offers. Multi-buy offers in the UK market are long running, so the set of UPCs for which multiple units are popular is broadly stable over time.

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

We incorporate the choice consumers make over size into our model of demand. Specifically, we define products as brand-variant-size combinations, and we model the consumer's choice of product from a discrete set of alternatives. For each brand-variant, the set of possible sizes includes both the available pack sizes (i.e. UPCs) and the most common multiple unit purchases of UPCs.[20] In Table B.1 we show, for each brand-variant, the number of sizes available to consumers in the at-home

[20]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.

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.[21] On-the-go sizes are always designed as a single serving, while at-home sizes are typically multi-portion.

3.5 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.

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 a household purchases on a choice occasion.[22] 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.1) on 60% of choice occasions on which drinks are purchased. On choice occasions in which a household chooses multiple (typically 2 or 3), we assume that (conditional on household specific preferences) these purchases are independent (for instance, 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 model the choice over which retailer to shop with as given, and assume it is driven by factors such as proximity of nearby stores and overall preferences for grocery outlet (for which we control in demand), and not consumers 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)

[21] 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.

[22] We focus on households that record making regular purchases; this excludes transactions (accounting for less than 2% of the total number) made by households who record making fewer than 10 shopping trips a year. We also focus on households who record making at least one drink purchase.

consumers are observed purchasing groceries from one retailer,[23] while consumers allocate, on average, a small share (4%) of their total supermarket expenditure to non-alcoholic drinks.

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.

The four main retailers in the UK implement national pricing policies.[24,25] 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.

[23] 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. They model consumer supermarket choice and within store allocation of expenditure across aggregated grocery goods (e.g. meat, dairy etc.). Integrating into their set-up choice over differentiated products in a specific market is a promising avenue for future research.

[24] The supermarkets agreed to implement national pricing policies following a Competition Commission investigation into supermarket behavior (Competition Commission (2000)).

[25] 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).

On-the-go segment

The natural periodicity for on-the-go purchases is at the daily level;[26] 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.1, and on 90% of these choice occasions they buy only one product.[27]

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).

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.[28] A concern with this is that recent at-home household purchases influence decisions that individuals make on-the-go (for instance, a recent at-home purchase may make an individual less likely to buy while on-the-go). We check for evidence of such non-separabilities across the at-home and on-the-go segments. Specifically, for individuals in the on-the-go sample we test whether recent purchases of drinks by their household in the at-home segment influences either their propensity to purchase drinks or the quantity they buy, finding no evidence of such dependence (see also Dubois et al. (2020)); details are provided in Appendix C.

[26] As in the at-home segment, we focus on individuals who record regularly, dropping less than 3% of total transactions that are made by those who record fewer than 5 purchases each year.

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

[28] When constructing market level demand we weight each segment such that their share of total sugar from sugar sweetened beverage matches that in the National Diet and Nutrient Survey (an individual level dietary intake survey, representative of the UK population).

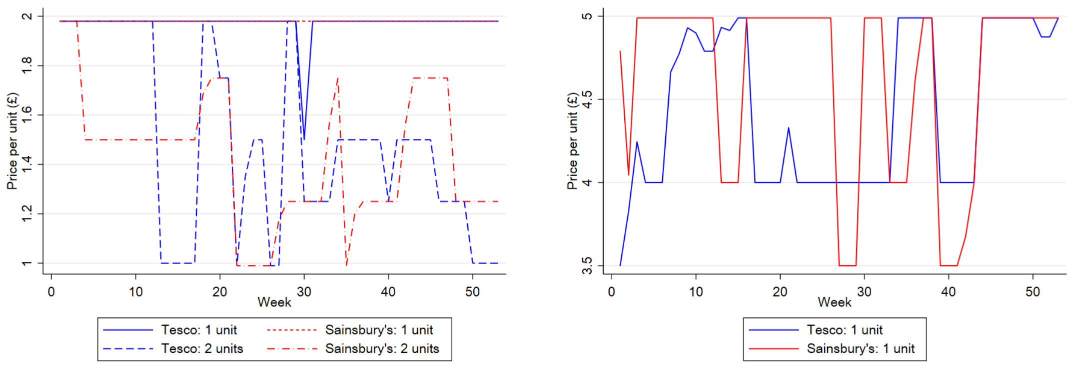
3.6 Price variation

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 (i.e. price reductions), which differ in their timing, duration and depth, both across UPCs and retailers. In our data, 30% of transactions are multi-buy offers (for instance, a discount for purchasing two of the same UPC), and 20% are ticket price reductions (when a UPC has a temporarily low price).

Figure 3.2: *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.

We provide a graphical example of each promotion type in Figure 3.2, which shows the price for two UPCs over the most recent year of our data for two retailers (Tesco and Sainsbury's). Panel (a) shows price series for a 2l bottle of Coke. In both retailers, (with the exception of one week in Tesco) 1 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.[29] 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 3.2 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

[29]In our demand model we treat one 2l bottle and two 2l bottles of Coke as different options.

compute measures of price stability suggested by DellaVigna and Gentzkow (2019). 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 (for products that are branded and available in multiple retailers), Asda is the cheapest retailer the most (for 27% of product-weeks) and the most expensive the least (for 17%) amount of time, and Sainsbury's is cheapest the least (for 22%) and most expensive the most (for 31%) amount of time.

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.[30] 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 price promotions create less incentives to stockpile. For instance, for each of the soft drinks products summarized in Table B.1, 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

[30] 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.

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.[31]

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. Price 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.

4 Estimating demand and supply

To implement our optimal corrective tax framework, we need to know how consumers switch across disaggregate 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 the possibility that preferences vary on each type of choice situation, but for notational parsimony we suppress a market segment index.

We index the drink products by $j = \{1, \dots, J\}$. The products vary by brand, which we index by $b = \{1, \dots, B\}$, size, indexed by $s = \{1, \dots, S\}$, and whether or not they contain sugar (for instance, the brand Coke comes 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

[31]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)).

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 (non-price) characteristics of product j and $\mathbf{x}_t = (\mathbf{x}_{1t}, \ldots, \mathbf{x}_{Jt})$ (note $p_0 = 0$ and $x_0 = 0$); $\boldsymbol{\theta}_i$ 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 $\widetilde{\mathbf{x}}_j^{(1)}$; $\widetilde{\mathbf{x}}_j^{(1)}$ includes a constant, which captures 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.[32] These individual level preferences play a key role in allowing the model to capture realistic substitution patterns across products. $\widetilde{x}_{jt}^{(2)}$ is a measure of the stock of advertising for the product in the current period;[33] 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.[34] 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

Our key identification assumption is that, conditional on our demand controls (including those for unobserved product attributes), the residual price variation is exogenous (and, in particular, the shocks to consumer's payoff functions, ϵ_{ijt}, are i.i.d.). The main form of price variation we exploit is *differential time series variation across retailers*, which we assume is driven by cost differences, store specific decisions related to unanticipated excess stock, random price reduction strategies and idiosyncratic variation in vertical contracts.

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

[33]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.

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

Our demand controls include brand-size effects, $\xi^{(1)}_{d(i)b(j)s(j)}$; these control for time-invariant unobserved characteristics that vary across brands and sizes. For instance, consumers may value one brand over another for reasons not fully controlled for by observed product characteristics; failure to control for this would likely result in correlation between ϵ_{ijt} and prices. By interacting brand with size effects we allow for the possibility that strength of unobserved brand effects vary 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, $\widetilde{\mathbf{x}}^{(1)}_j$. 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).

The time (quarterly) varying brand effects, $\xi^{(3)}_{d(i)b(j)t}$, control for shocks to national level demands for each brand. We additionally control for time varying size effects $\xi^{(5)}_{d(i)s(j)t}$, which capture any tendency through time for demands for larger versus small sizes to fluctuate. As discussed in Section 3.2, 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. In addition, we 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.[35] For convenience stores we use mean regional prices. We include region-quarter varying drinks effects in demand to control for the possibility of regional shocks to demand for drinks.

We also control for brand-retailer effects, $\xi^{(2)}_{d(i)b(j)r}$, and size-retailer effects, $\xi^{(4)}_{d(i)s(j)r}$. These capture the possibility that either the prominence of products belonging to different brands, or of large versus small sizes, may vary across retailers. They also capture average differences in consumer brand and size preferences across retailers.

An important restriction we make is the absence of retailer-time shocks to product demands that correlate with price setting.[36] As outlined in Section 3.6, average prices across retailers are similar, but co-movement in prices is low, with, for instance, the use of price promotions not synced across 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. Given the national nature of much retailing,

[35]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.

[36]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.

pricing and advertising in the UK drinks market, and the absence of targeted price offers and coupons, we believe that this assumption is reasonable.

4.3 Supply model

We model price competition among the firms operating in the UK drinks market. We assume that they simultaneously set prices to maximize profits in a Nash-Bertrand game, abstracting from modeling retailer-manufacturer interactions. This outcome can be achieved by use of non-linear vertical contracts (see Villas-Boas (2007), Bonnet and Dubois (2010)).[37] In Section 5.4 we show how our optimal tax results are influenced by different supply-side models.

Let $\boldsymbol{p}_m = (p_{1m}, \ldots, p_{Jm})$ denote the prices that drinks firms set in market m, where markets are temporal.[38] 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.[39] We denote the marginal cost of product j in market m as c_{jm}.[40]

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

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

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

[37]Non-linear contracts with side transfers between manufacturers and retailers allow them to reallocate profits and avoid the double marginalization problem. Bonnet and Dubois (2010) show evidence of price equilibria in the French bottled water market consistent with use of non-linear contracts.

[38]In the supply model we average over short-run price variation, as this likely reflects random price promotion strategies 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.

[39]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.

[40]Note, in Section 2 we express quantity in terms of units (i.e. 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 analogues in Section 2 without changing the nature of the firms' problem.

set of first order conditions: $\forall f$ and $\forall j \in \mathcal{J}_f$,

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

From this system of equations we can solve for the implied marginal cost, c_{jm}, 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 $\forall j \in \mathcal{J}_f$,

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

For any τ, we can solve the system of equations to obtain the vector of counterfactual equilibrium prices, $\boldsymbol{p}'_m = (p'_{1m}, \ldots, p'_{Jm})$.[41]

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. For details see Appendix F.

4.4 Demand estimates

We estimate the demand model outlined in Section 4.1 using simulated maximum likelihood[42] and report the coefficient estimates in Appendix D. The estimated coefficients exhibit some intuitive patterns; those with more added sugar in their 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.

[41] Note that for the set of store brand products, we do not model price re-optimization – for store brand sugar sweetened beverages we assume pass-through of any tax is 100%, and for store brand diet beverages we assume consumer prices remain unchanged.

[42] 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.

The variance parameters of the random coefficients are significant both statistically and in size, indicating an important role for unobserved preference heterogeneity.

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 market level own- and cross-price demand elasticities; in Table 4.1 we summarize the market level product own- and cross-price elasticities. The mean own-price elasticity is around -2.4, 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. The distribution of the cross-price elasticities exhibits a high degree of skewness, with the mean close to the 75^{th} percentile. This reflects consumers willingness to switch between products close together in product characteristic space.

Table 4.1: *Summary of own- and cross-price elasticities*

	No. elasticities per market	Mean	Percentile 25$^{\text{th}}$	Percentile 50$^{\text{th}}$	Percentile 75$^{\text{th}}$
Own-price	175	-2.431	-2.795	-2.434	-1.765
Cross-price	18757	0.013	0.003	0.007	0.014

Notes: In each market there are 175 own-price elasticities (one for each product) and 18757 cross-price elasticities (between product pairs available either in the at-home or on-the-go segment). Numbers summarize the distribution of market elasticities based on the most recent year covered by our data (2012).

Table 4.2 illustrates consumers' tendency to switch between similar products by showing 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 10x330ml packs of Cokes and Pepsis, and the mean elasticities for other (non-cola) sugar sweetened and diet beverages, and for fruit juice. The table illustrates 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 alternative 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 product level own and cross-price elasticities for popular products in the at-home and on-the-go segments.

Table 4.2: *Selected elasticities for cola products*

	Coke						Pepsi						Non-colas		
	Regular		Diet		Zero		Regular		Diet		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.	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	0.017	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	0.028	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	0.022	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	0.026	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.

In Table 4.3 we summarize the effects of increasing the price of all sugar sweetened beverages by 1%. The resulting reduction in demand (in liters) for sugar sweetened beverages is 1.48% (i.e. our estimates correspond to an own-price elasticity for sugar sweetened beverages of 1.48). The diversion ratio (defined as the percentage of the reduced sugar sweetened beverage demand that is diverted to each group of substitute products) is 27.3% for diet drinks and 6.7% for alternative sugary drinks. The percent change in expenditure on non-alcoholic drinks is 0.05% – the price increase leads to a modest increase in drinks expenditure. 95% confidence intervals are given in brackets.[43]

Table 4.3: *Switching from sugar sweetened beverages*

Own-price elasticity of sugar sweetened beverages	Diversion ratio: Diet beverages	Diversion ratio: Sugary alternatives	Elasticity of drinks expenditure
-1.48	27.3%	6.7%	0.05
[-1.52, -1.43]	[26.8%, 28.1%]	[6.4%, 7.0%]	[0.04, 0.07]

Notes: We simulate the effect of a 1% price increase for all sugar sweetened beverage products. Column 1 shows the % reduction in volume demanded of sugar sweetened beverages, columns 2 and 3 shows how much of the volume reduction is diverted to diet beverages and sugary alternatives and column 4 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.

The optimal tax formula, given by equation (2.3), partly depends on how much any reduction in the equilibrium quantity of taxed drinks induced by a marginal

[43] To calculate the confidence intervals, we obtain the variance-covariance matrix for the parameter vector estimates using standard asymptotic results. We then take 100 draws of the parameter vector from the joint normal asymptotic distribution of the parameters and, for each draw, compute the statistic of interest, using the resulting distribution across draws to compute Monte Carlo confidence intervals (which need not be symmetric).

tax increase is shifted to non-taxed substitutes. The diversion ratios suggest a significant amount of demand for sugar sweetened beverages will be diverted to non-taxed drinks, while the elasticity of total drinks expenditure indicates only a modest degree of switching from the numeraire. However, these diversion ratios and elasticities summarize the demand effects *at observed prices*. The optimal tax formula depends on changes in equilibrium quantities (which depend on supply responses) and are evaluated at the optimal tax rate. We fully incorporate this when we solve for the optimal tax rate in Section 5.

4.5 Supply estimates

We use the first order conditions of the firms' profit maximization problem (equation (4.5)) to solve for the marginal cost of each product. This enables us to compute the equilibrium price-cost margins (which we express per liter) and price-cost mark-ups (margin over price) at the observed market equilibrium (where no sugar sweetened beverage tax is in place). In Table 4.4 we summarize the distribution of prices, costs, margins and mark-ups across products; in Appendix D we show these by brand. The average mark-up is 0.55 (price is around double marginal cost), though there is considerable variation around this. This broadly accords with evidence from accounting data, with gross margins in this market being reported to be around 50-70% (see Competition Commission (2013)).

Table 4.4: *Summary of costs and margins*

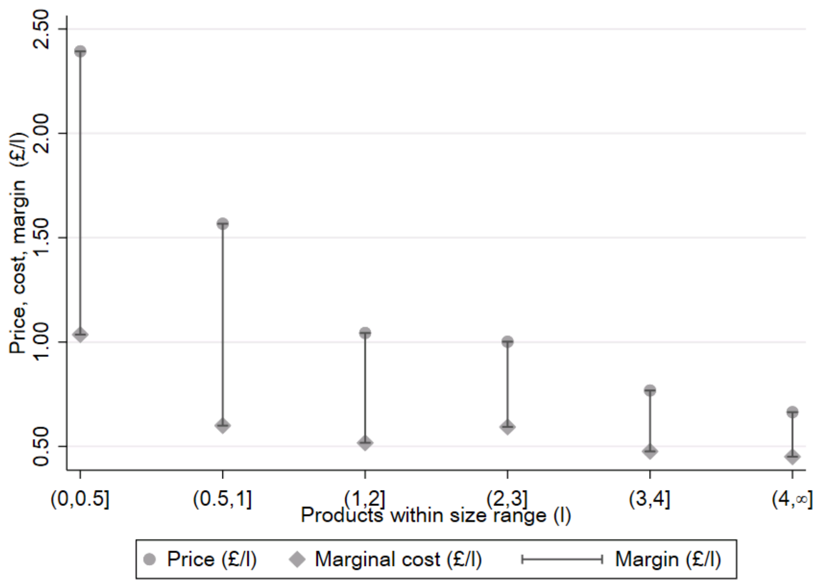
		Percentile		
	Mean	25^{th}	50^{th}	75^{th}
Price (£/l)	1.44	0.83	1.16	1.96
Marginal cost (£/l)	0.67	0.31	0.61	0.83
Price-cost margin (£/l)	0.77	0.43	0.56	0.98
Price-cost mark-up (margin/price)	0.55	0.41	0.50	0.67

Notes: We recover marginal costs for each product in each market. We report summary statistics for the most recent year covered by our data (2012). Margins are defined as price minus cost and expressed in £ per liter; mark-ups are margins over price.

Equilibrium price-cost margins play an important role in determining the optimal tax policy. All else equal, the higher (lower) are margins on externality (non-externality) generating options, the lower (higher) will be the optimal tax rate on the externality product attribute. At observed prices, the (unweighted) average margin across sugar sweetened beverages is 0.78, while it is 0.76 across alternative drinks. How these margins adjust in equilibrium with the tax, and how consumers switch within the two sets of options and between them is important in determining the optimal tax rate.

In Figure 4.1 we show how prices, marginal costs, and equilibrium price-cost margins vary with product size. There is strong non-linear pricing; in per liter terms smaller products are, on average, more expensive. Average marginal costs are broadly constant across the size distribution, with the exception of small single portion sizes (i.e. with sizes no larger than 500ml), 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 is important in driving the optimal tax rate, as one way consumers respond to the tax is by shifting their basket of taxed products towards small, high margin sizes – this acts to exacerbate the distortions associated with the exercise of market power.

Figure 4.1: *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).

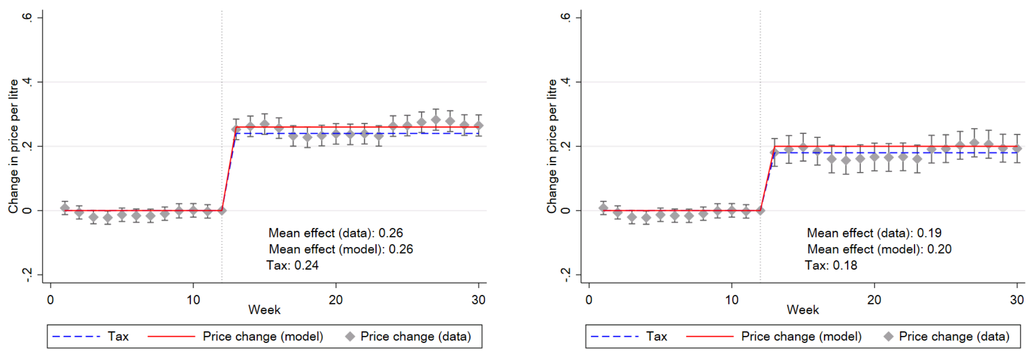
4.6 Model validation

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 of the market. We use 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.[44]

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 products with sugar content > 8g/100ml. We use an event study approach to 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 4.2 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. The predicted price changes from the model are very close to the observed price changes.

Figure 4.2: *Comparison of model predictions with event study*

(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.

These patterns are broadly consistent with the literature that uses ex post evaluation methods to estimate 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

[44]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.

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. Bollinger and Sexton (2018)). A likely reason for low tax pass-through in Berkeley is, given the small geographical area in which the tax is operation, consumers can readily avoid the tax through cross-border shopping.

5 Corrective tax results

In this section, we embed our estimated empirical model of supply and demand in the drinks market into the tax design framework set out in Section 2 to solve for the optimal tax rate on sugar sweetened beverages and its effect on prices, purchases, and welfare. We also consider how the tax's performance is affected by the structure of the market and how it compares to alternative tax instruments.

We repeat, for reference, the implicit formula for the optimal rate of tax levied on the externality generating product attribute (the sugar in sweetened beverages), equation (2.3), with the exception that we split out the effect of switching to the set of untaxed drinks products from the effect of switching to the numeraire good:

$$\tau^* = \underbrace{\tilde{\phi}}_{\text{Externalities}} - \frac{1}{\frac{dZ}{d\tau} \Big/ \frac{dQ^{\mathcal{S}}}{d\tau}} \left(\underbrace{\sum_{j \in \mathcal{S}} w_j^{\mathcal{S}} \mu_j - \sum_{j \in \mathcal{N}} w_j^{\mathcal{N}} \mu_j \times \frac{dQ^{\mathcal{N}}}{d\tau} \Big/ \left(-\frac{dQ^{\mathcal{S}}}{d\tau} \right)}_{\text{Market power in drinks market}} - \underbrace{\tilde{\mu} \times \frac{dX}{d\tau} \Big/ \left(-\frac{dQ^{\mathcal{S}}}{d\tau} \right)}_{\text{Numeraire good market power}} \right). \tag{5.1}$$

Here we denote the externality term by $\tilde{\phi}$; the precise form this takes will depend on whether there is heterogeneity in marginal externalities and whether externalities arise from consumption of untaxed goods. We use $\tilde{\mu}$ to denote the mark-up on the numeraire good, and $\frac{dX}{d\tau}$ to denote the marginal effect of the tax on equilibrium consumption of the numeraire good.[45]

Our demand and supply model allows us to simulate, for any tax rate, the degree of switching between drinks products and from drinks to the numeraire, and the equilibrium price-cost margins on products in the drinks market. However, it does not provide us with information on the marginal externalities nor on the mark-up

[45] Consumption of the numeraire good is given by $X = \sum_i \left(y_i - \sum_j p_{ij} q_{ij} \right)$. See Appendix F for how the planner's problem is modified to accommodate a non-competitively supplied numeraire good.

on the numeraire good. We use existing evidence to calibrate these parameters, and first describe how the patterns of consumer switching and firms' endogenous margin adjustment affect the optimal tax rate. We then show how the optimal rate and associated components of welfare vary with the calibrated parameters.

Baseline calibration

In Section 3.1 we summarize the well-established evidence that links consumption of sugar sweetened beverages to externalities. However, placing a numerical value on the marginal externality associated with an extra gram of sugar from these products is challenging. We begin by considering a marginal externality of £4.00 per kg of sugar from sweetened beverages (which translates to approximately 1.5c/oz of sugar sweetened beverage[46]). This value is similar to that implied by epidemiological estimates of the impact on health care costs (e.g. Wang et al. (2012)).[47] In this case $\tilde{\phi} = \phi = 4$. Below we show how the optimal rate varies with the magnitude, degree of heterogeneity, and source of the externalities from sugar consumption.

The optimal rate also depends on the degree to which there is market power associated with the numeraire good (which represents what consumers switch towards when lowering their drinks expenditure),[48] and the direction and strength of consumer switching towards the numeraire good. We calibrate the mark-up on the numeraire good using an estimate for the UK economy-wide mark-up from De Loecker and Eeckhout (2018), which implies $\tilde{\mu} = 0.4$. The average mark-up on drinks products, based on our estimates, is around 30% higher than this.[49] Below we show how the optimal rate depends on the value of the numeraire mark-up.

5.1 Optimal tax rate

Under our baseline calibration of the marginal externality function and the price-cost mark-up on the numeraire good, the optimal rate of tax on the sugar in sweetened beverages is £1.74 per kg of sugar (95% confidence interval [1.62, 1.85]). How-

[46]A typical sugar sweetened beverage contains 10g of sugar per 100ml, or around 3g per oz, which implies a marginal externality per oz of 1.5¢.

[47]Wang et al. (2012) estimate that a 15% reduction in consumption of sugar sweetened beverages among US adults aged 24-65 would result in health care costs savings of $17.1 billion over 10 years. Converting this to savings per person, per kg of sugar and adjusting for differences in the cost of providing health care in the UK implies an externality of roughly £4 per kg of sugar.

[48]Note that as we are free to normalize the price of the numeraire to 1, $\tilde{\mu}$ can equivalently be interpreted as the numeraire price-cost margin or mark-up.

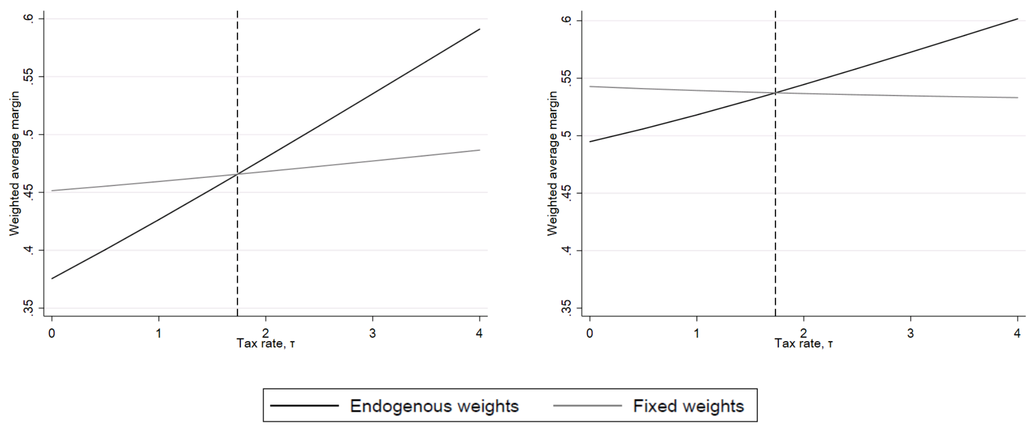
[49]De Loecker and Eeckhout (2018) adopt the convention of measuring mark-ups as price over marginal cost, and estimate that this is 1.68 in the UK economy. This corresponds to a mark-up defined as margin over price on the numeraire of around 0.4. The average of our estimated mark-ups on drinks is 0.55.

ever, the optimal tax rate lies well below the rate that would be optimal under perfect competition (a Pigovian tax of £4 per kg of sugar). It also lies below the rate that a planner that ignores distortions associated with the exercise of market power would set: a planner that takes the allocation in the absence of tax as a benchmark and aims for full externality internalization relative to this baseline, would set a tax rate of £3.64 per kg of sugar (95% confidence interval [3.54, 3.68]).[50]

The reason why the optimal rate lies below the rate aiming at full internalization of externalities is the existence of positive price-cost margins for sugar sweetened beverage products. This is reflected in the optimal tax formula by the weighted average margin term, $\sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}\mu_j$. This expression reflects both the equilibrium product level price-cost margins set by drinks firms on the taxed products (i.e. μ_j) and, through the weights, switching within the set of taxed products. In particular, the weights capture how much each product contributes to the derivative of total equilibrium quantity of the set of taxed products with respect to the tax rate. Both the product level margins and weights may vary with the tax rate.

Figure 5.1: *Weighted average margins*

(a) Taxed products (b) Untaxed products



Notes: The black lines show the weighted average margin on taxed products, $\sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}\mu_j$ (left hand panel), and on untaxed products, $\sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}\mu_j$ (right hand panel), at the tax rate shown on the horizontal axes. The grey lines show weighted average margins constructed using weights held fixed at their values at τ^*. The dashed vertical line indicates the optimal tax rate, τ^*, when $\phi = 4$ and $\tilde{\mu} = 0.4$.

Figure 5.1(a) shows that the weighted average margin on the taxed products increases with the tax. In other words, for sugar sweetened beverages, the tax acts to exacerbate distortions associated with the exercise of market power. This happens for two reasons. First, firms choose to raise the price-cost margins they set

[50] A planner aiming to achieve full internalization of externalities sets a tax rate below the value of the marginal externality because firms' pricing responses act to amplify the effect of the tax on prices. For a formal statement of the problem this planner solves see Appendix A.

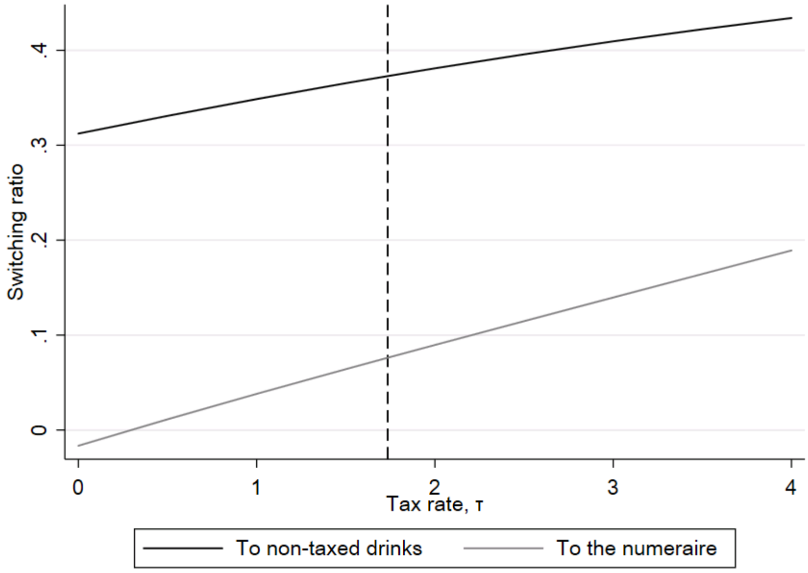
(meaning tax is slightly overshifted, with average pass-through of 109%) – shown by the grey line, which holds fixed the product weights. This is driven by relatively elastic consumers switching away as prices increase, leaving firms optimally pricing for the remaining set of slightly less elastic consumers. Second, consumers adjust their basket of taxed goods towards products with relatively high margins. This effect, shown by the fact that the black line is more steeply sloped than the grey line, is primarily driven by consumers switching more strongly away from large (relatively low margin) products than smaller products.

However, the effect that distortions associated with the exercise of market power among the taxed products plays in suppressing the optimal rate is, to some extent, offset by consumers switching to other products that are also supplied non-competitively. The effect of market power among non-taxed drinks on the optimal tax rate is captured in the tax formula through the term $\sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}\mu_j$. Figure 5.1(b) shows that the weighted average price-cost margin on non-taxed substitute drinks increases with tax, and this is driven by the equilibrium impact of consumer switching (the grey line is roughly flat, while the black is upwards sloping). This is because consumers switch most strongly towards natural fruit juices, and these products have relatively high price-cost margins.

The influence of market power among the non-taxed drinks on the optimal tax rate is determined by the term $\frac{dQ^{\mathcal{N}}}{d\tau}\Big/\left(-\frac{dQ^{\mathcal{S}}}{d\tau}\right)$: the fraction of the reduction in the equilibrium quantity of the taxed products, induced by a marginal tax change, that is diverted to untaxed alternative drinks. In Figure 5.2 we show how this switching ratio varies with the tax rate. When the tax rate is close to zero a marginal increase in the rate leads to just over 30% of the resulting reduction in the equilibrium quantity of taxed beverages being offset by an increase in equilibrium quantity of untaxed drinks (diet products and alternative sugary drinks). This switching ratio rises with the tax rate to just under 40% at the optimal rate. The figure also shows that the impact of a marginal change in the tax rate on numeraire good consumption is rising in the rate; at the optimal rate, just under 10% of the reduction in the equilibrium quantity of taxed products is offset by higher numeraire good consumption. At the optimal rate the numeraire good is therefore a substitute for sugar sweetened beverages, which means that the positive numeraire good margin acts to raise the optimal tax rate.[51]

[51]The optimal tax rate also depends on the ratio $\frac{dZ}{d\tau}\Big/\frac{dQ^{\mathcal{S}}}{d\tau}$ i.e. the responsiveness of total sugar from taxed products to a marginal change in the tax rate over the responsiveness of total quantity of the taxed products. The ratio is around 0.1 for all values of the tax, reflecting that average sugar per liter of sugar sweetened beverage is around 100g. This term has the role of "converting" the margin components of the optimal tax formula from per liter to per kg of sugar.

Figure 5.2: *Switching ratios from taxed products*



Notes: The black line shows the switching ratio between taxed and untaxed drink products at the tax rate shown on the horizontal axis (i.e. $\frac{dQ^{\mathcal{N}}}{d\tau} / \left(-\frac{dQ^{\mathcal{S}}}{d\tau}\right)$*). The grey line shows the switching ratio for the numeraire good (i.e.* $\frac{dX}{d\tau} / \left(-\frac{dQ^{\mathcal{S}}}{d\tau}\right)$*). The dashed vertical line indicates the optimal tax rate,* τ^**, when* $\phi = 4$ *and* $\tilde{\mu} = 0.4$*.*

5.2 Impact on drinks purchases, sugar and welfare

In Table 5.1 we summarize the impact on purchases and prices, of the tax rate that is optimal under our baseline calibration. We describe the effects separately for sugar sweetened beverages, zero sugar soft drinks and alternative sugary drinks (i.e. fruit juice and flavored milks, which are not subject to the tax).

The tax results in a 14.2% increase in the average price of sugar sweetened beverages. The median pass-through rate across products is 109% (the 25^{th} and 75^{th} percentiles are 105% and 113%). On average, the price of zero sugar soft drinks falls by 0.6%. These price changes lead to a 13.7% fall in the average probability of consumers selecting a sugar sweetened beverage on a shopping trip. Conditional on buying a sugar sweetened beverage, there is an average reduction in volume (liters) purchased of 13.5% and a 2.4% reduction in the sugar intensity (gram/liters) of purchases (reflecting consumer substitution towards brands that are less sugar intense). Consumers increase the probability that they buy zero sugar soft drinks by 7.3%, and also buy slightly higher volumes of these types of drinks, conditional on choosing to buy. Together, this implies a 28.4% decline in the sugar from soft drinks. However, consumers also switch towards alternative, untaxed sugary drinks, increasing the sugar from these products by 7.3%. This means that the overall fall in sugar from drinks is 22.1%.

Table 5.1: *Impact on purchases*

	Sugar sweetened beverages	Zero sugar soft drinks	Alternative (sugary) drinks
%Δ price change	14.2	-0.6	-0.1
	[13.9, 14.6]	[-0.6, -0.5]	[-0.2, -0.1]
%Δ purchase probability	-13.7	6.9	7.3
	[-14.0, -13.4]	[6.8, 7.2]	[7.1, 7.7]
conditional on purchase:			
%Δ volume	-13.5	1.6	-0.1
	[-14.1, -13.0]	[1.4, 1.8]	[-0.2, 0.0]
%Δ sugar intensity	-2.4		0.0
	[-2.5, -2.3]		[0.0, 0.0]
%Δ sugar	-28.4		7.3
	[-29.2, -27.7]		[7.0, 7.6]

Notes: Price changes refer 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^**, at* $\phi = 4$ *and* $\tilde{\mu} = 0.4$*. 95% confidence intervals are given in square brackets.*

In Table 5.2 we report the impact of the optimal tax rate on welfare. It leads to an increase in total welfare of £129 million per year. This is comprised of a fall in consumer surplus of £691 million and in soft drinks firms' profits of £259 million, which is more than offset by tax revenue of £558 million, and a reduction in the external costs of sugar sweetened beverage consumption of £509 million (there is also a small increase in the profits associated with the numeraire good).

We also report the impact on welfare of setting the Pigovian tax rate, and the rate that aims at full internalization of externalities. The Pigovian rate leads to a welfare gain of £21 million per year, and the tax rate set by a planner that aims at full internalization of externalities improves welfare by £49 million. The optimal sugar tax leads to welfare gains that are over 6 and 2.5 times as large, respectively; ignoring the presence of distortions associated with the exercise of market power when setting the tax rate leads to smaller welfare gains. In Appendix D we show that these differences in welfare gains are highly statistically significant.

Table 5.2: *Welfare changes under alternative objectives*

		Change in:					
		Welfare components					
	Tax rate	Cons. surplus	Tax rev.	Drinks profits	Num. profits	Ext. costs	Total welfare
Optimal sugar tax	1.74	-691	558	-259	13	-509	129
Pigovian tax	4.00	-1345	904	-489	61	-890	21
Rate aiming at full internalization	3.64	-1254	865	-457	52	-842	49

Notes: Description of the different tax rates is provided in the text. Num. profits refers to profits outside the drinks market (i.e. from the numeraire good), Ext. costs refers to externality costs. Numbers are in £million per year and are reported for $\phi = 4$ and $\tilde{\mu} = 0.4$. In Table G.2 in Appendix D we show differences in welfare numbers across the different policies are statistically significant.

5.3 Impact of externalities and numeraire good margin

Size of externalities

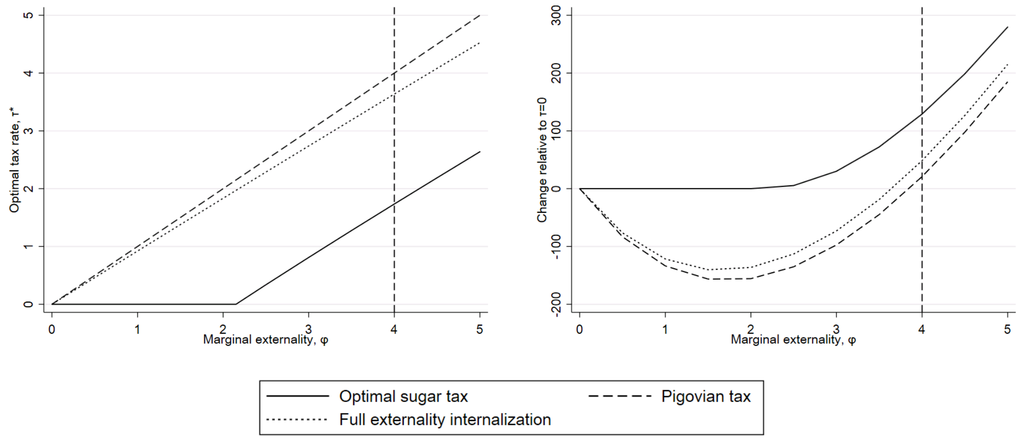
In Figure 5.3(a) we plot how the optimal tax rate, as well as the Pigovian rate, and the tax rate set by a planner aiming at full internalization of externalities, varies with the size of marginal externalities, ϕ. If the marginal externality from the sugar in sweetened beverages is greater than £2.15 per kg of sugar, a positive tax rate improves welfare. Assuming that the planner faces a non-negativity constraint on tax revenue, the optimal tax for marginal externalities below this level is zero. For a marginal externality above £2.15, the optimal tax rate is increasing approximately linearly in the marginal externality. At all positive values of ϕ the Pigovian tax rate and the rate that aims at full internalization of externalities lie above the optimal rate. Figure 5.3(b) plots how the welfare changes achieved by the three tax policies vary with the size of marginal externalities. For all values of ϕ, ignoring the distortions associated with the exercise of market power leads to smaller welfare gains (and reduces welfare if ϕ <3.70).

Heterogeneity in externalities

Although there is evidence that sugar obtained through consumption of sweetened beverages is particularly harmful, there is also concern about high levels of dietary sugar more broadly. The World Health Organization has a target that no more than 10% of dietary calories should be obtained from added sugar, which acknowledges that sugar intake for those with high levels of dietary sugar is likely to be more harmful than for those with modest levels.

Figure 5.3: *Externality size*

(a) Optimal tax rate (b) Change in welfare



Notes: The left panel shows how the optimal tax rate levied on the sugar in sweetened beverages, the Pigovian tax rate (optimal under perfect competition) and the tax rate aimed at full internalization of externalities varies with the calibrated value of the marginal externality, ϕ. The right panel shows the resulting change in welfare. The dashed vertical line shows the baseline calibration, $\phi = 4$. Numbers are based on $\tilde{\mu} = 0.4$.

We consider how heterogeneity in marginal externalities impacts the optimal tax prescription. Denoting by $\mathfrak{a}_i$ our measure of total dietary sugar (given by the share of total calories purchased as added sugar by the consumer's household in the previous calendar year), we specify the marginal externality from the sugar in sweetened beverages for consumer i as $\phi_i = A\exp(b \times \mathfrak{a}_i) - 1$. b controls for the curvature of the function and, conditional on b, A determines the level of externalities. We vary b, and calibrate A so that the (unweighted) marginal externality across consumers is fixed and equal to its value in the baseline calibration; increasing b translates into increasing marginal externalities for those with high levels of dietary sugar relative to those with low levels.[52]

In our baseline calibration, marginal externalities are constant across the total dietary sugar distribution, and the optimal tax rate is $\tau^* = 1.74$, with an associated increase in welfare of £129 million per year. When those at the 95$^{\text{th}}$ percentile of the added sugar distribution have marginal externalities 2.5 times the size of those at the 5$^{\text{th}}$ percentile, the optimal rate is 1.80 and the welfare gain is £140 million; when the multiple is 6 times, the optimal rate is 1.89 and the welfare gain £150 million. The optimal tax rate rises with the degree of heterogeneity in marginal externalities because there is a positive correlation across consumers in total dietary sugar and

[52] We vary b from 0 to 10. When $b = 0$, $A = 5$ and $\phi_i = 4 \, \forall \, i$. When $b = 10$, $A = 1.2$, and at the 5^{th} percentile of the distribution of a_i (a_i=6%), $\phi_i = 1.2$, while at the 95^{th} percentile (a_i=19%), $\phi_i = 7.2$.

the equilibrium level reduction in sugar from sweetened beverages associated with a marginal tax rise. However, as the strength of this correlation is not large, the optimal tax rate increases only modestly in the degree of heterogeneity.

Externality leakage

Under our baseline externality calibration, we assume that externalities arise from the sugar in sweetened beverages. However, as shown in Table 5.1, consumers switch to other sources of sugar. If the marginal externality from sugar intake is given by $\phi = 4$, and this is associated both with sugar sweetened beverages and (untaxed) fruit juices and flavored milks, the optimal tax rate is 1.55 and the associated welfare increase is £103 million. Both the optimal rate and welfare gain lie below the case when externalities arise only through intake of sugar in sweetened beverages. This is because the tax on sugar sweetened beverages is now relatively less effective at reducing externalities as some consumers switch from taxed to untaxed sources. The optimal rate falls further, to 1.41 (and the associated welfare gain to £98 million), if consumption of the numeraire good also creates externalities.[53]

Numeraire good margin

Figure 5.2, shows that (when the tax rate is above 0.2) a marginal increase in the rate, in equilibrium, causes consumers to switch towards the numeraire good. Therefore, in equilibrium, the numeraire good is a substitute for sugar sweetened beverages, and its mark-up (in the same way as those on alternative drinks products) acts the raise the optimal rate. However, the effect of the numeraire good mark-up on the optimal rate is modest; when it is 0, the optimal rate is 1.51, when it 0.55, equal to the average mark-up in the drinks market, the optimal tax rate is 1.84. The reason for this is that a marginal increase in the tax rate induces relatively modest switching to the numeraire good: the fraction of the reduction in equilibrium sugar sweetened beverage consumption caused by a marginal rate increase that switches to the numeraire is 0 when $\tau = 0.2$, rising to just 0.08 when $\tau = 1.74$.

5.4 Tax policy and competition

The potential for an optimally set tax on externality generating products to improve welfare depends on the structure of, and degree of competition in, the market. In

[53] We simulate the case when the numeraire good has a sugar content of 54.5g per £1 of expenditure, which is the sugar intensity of a commonly purchased chocolate bar, and this sugar creates externalities that are the same as for sugar sweetened beverages in the baseline case, i.e. $\phi = 4$.

the simplest case of a monopolist producing one externality generating product, for sufficiently large externalities, a reduction in market power reduces welfare, but this is completely unwound by the optimal tax that achieves the first best – competition and optimal tax policy are perfect substitutes.[54] However, when there are many products, differing in margins and the extent to which they create externalities, the relationship between competition and optimal tax policy is more nuanced.

To highlight how the degree of competition in the market and tax policy interact, we simulate the effects of changing the market structure through varying the degree to which firms internalize portfolio effects when setting their prices. A multi-brand firm derives market power from the fact that if it raises the prices of products belonging to one of its brands, some of the consumers that switch away from those brands will switch to alternative products that it owns. This results in the firm setting higher prices than would be optimal in the absence of these portfolio effects.

It is useful to re-express the firms' first order conditions (equation 4.5) in vector notation:

$$\mathbf{p}_m = \mathbf{c}_m - \left[\Omega \otimes \left(\frac{\partial \mathbf{q}_m(\mathbf{p}_m)}{\partial p_m}\right)^T\right]^{-1} \mathbf{q}_m(\mathbf{p}_m).$$

Ω is a $J \times J$ matrix encoding the ownership structure of products in the market. Under the true (i.e. observed) market structure element $(j,k) = 1$ if products (j,k) are owned by the same firm, otherwise $(j,k) = 0$. We consider the following counterfactual market structure given by the ownership matrix $\widetilde{\Omega}$, where element:

$$(j,k) = \begin{cases} 1 & \text{if products } (j,k) \text{ belong to the same brand} \\ \theta_w \in [0,1] & \text{if products } (j,k) \text{ belong to the same firm but not the same brand} \\ \theta_b \in [0,1] & \text{if products } (j,k) \text{ belong to different firms} \end{cases}.$$

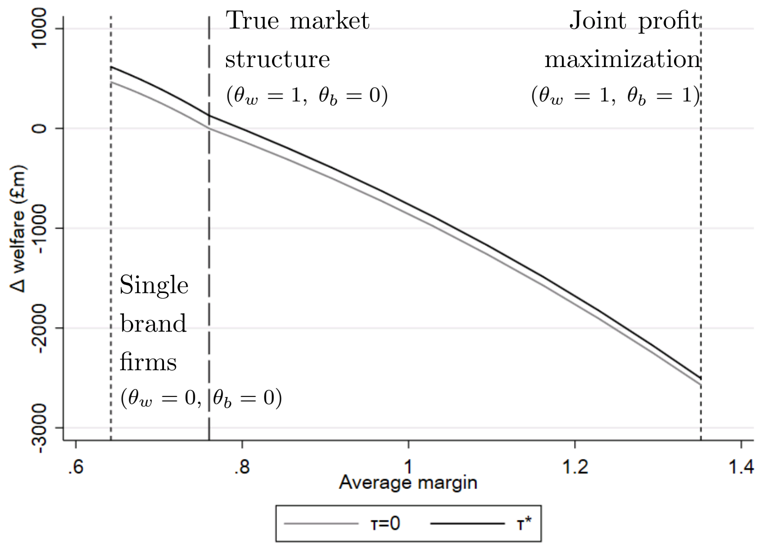
$(\theta_w = 1, \theta_b = 0)$ yields $\widetilde{\Omega} = \Omega$ – the true market structure. When $(\theta_w = 0, \theta_b = 0)$ the market structure is given by a set of single brand firms (i.e. divesting the multi-brand firms into separate single brand firms), and when $(\theta_w = 1, \theta_b = 1)$ firms owning the branded drink products behave as joint profit maximizers. More generally, when $\theta_b = 0$, reducing θ_w increases the competitiveness in the market compared with the true market structure, and when $\theta_w = 1$ raising θ_b from 0 decreases the degree of competition in the market relative to the true market structure.[55]

[54]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 (given by a lower μ) leads to lower welfare. However, the optimal tax, $\tau = \phi - \mu$ will exactly off-set this, bringing the market back to the first best.

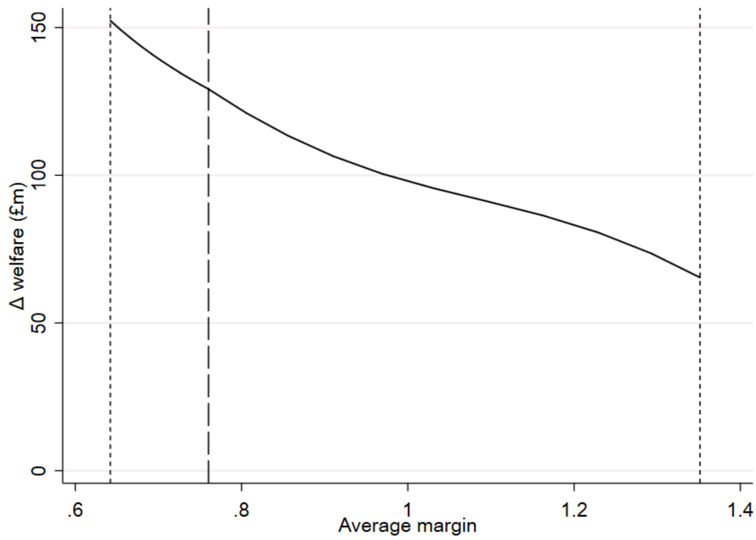
[55]Weyl and Fabinger (2013) and Mahoney and Weyl (2017) take a related approach by using a "conduct" parameter to capture the degree of competition in a model of symmetric product differentiation, while in their study of pricing in the beer market Miller and Weinberg (2017)

Figure 5.4: *Impact of market competitiveness on welfare*

(a) Welfare relative to observed market structure and $\tau = 0$



(b) Welfare gain from optimal tax policy



Notes: We simulate the market equilibrium under counterfactual ownership structures, $\widetilde{\Omega}$. The figure, to the left of the dashed line, shows the effect of varying θ_w, holding fixed $\theta_b = 0$, between 0 (single brand firms) and 1 (the observed market structure), and, to the right of the dashed line, varying θ_b, holding fixed $\theta_b = 1$, between 0 (the observed market structure) and 1 (joint profit maximization). The horizontal axis shows the average equilibrium price-cost margin (when there is no tax in place). In panel (a) the grey line shows the change in welfare as θ_w and θ_b are varied when the tax is zero, relative to the true market structure (and no tax), and the black line shows the change in welfare under the optimal tax, relative to the true market structure (and no tax). Panel (b) plots the welfare gain from optimal tax – which equals the difference between the black and grey lines in panel (a).

estimate the degree to which two major firms internalize between firm portfolio effects (i.e. a parameter similar to θ_w).

In Figure 5.4 we plot the welfare consequences of varying the market structure. The vertical dashed line corresponds to the true market structure ($\theta_w = 1, \theta_b = 0$). Moving to the left of this line corresponds to lowering the within-firm internalization of portfolio effects (θ_w), holding fixed the between effects, θ_b, at 0. This results in a lower average equilibrium margin (shown on the horizontal axis). Moving to the right of the dashed line corresponds to raising θ_b, holding $\theta_w = 1$, which translates into a higher equilibrium margin. The grey line in panel (a) shows the impact on total welfare associated with varying the market structure. The more competitive is the market the higher is total welfare; the increase in consumer surplus from lower prices more than offsets a reduction in profits and higher externalities from sugar consumption. The black line shows the welfare consequences of levying the optimal tax rate; for all market structures the optimal tax improves welfare.

Panel (b) shows that the gains from the optimal tax (i.e. the difference between the two lines in panel (a)) are larger the more competitive is the market. Under the true market structure, $\tau^* = 1.74$ and the gain in welfare is £129m, whereas, under single brand firms $\tau^* = 1.86$ and the associated gain in welfare is £152m and under joint profit maximization $\tau^* = 1.35$ and welfare rises by £65m (see Table 5.3). This points towards a complementarity in competition and tax policy – a more competitive market raises welfare, and also raises the welfare gains associated with optimal tax policy – and contrasts with the single product monopolist case.

5.5 Alternative tax bases

The majority of jurisdictions that have implemented taxes on soft drinks have set the tax base as the volume of sugar sweetened beverages sold (as opposed to products' sugar content), which has the implication that soft drinks with higher sugar contents do not attract proportionately more tax. We solve for the optimal rate if the planner chooses to levy the tax volumetrically on sugar sweetened beverages. We also solve for the optimal rate for a tax that is levied on an ad valorem basis on sugar sweetened beverages (a structure that is used in the Chilean sugar sweetened beverage tax). In some jurisdictions (for instance, Philadelphia) a volumetric tax is levied on all soft drinks (including zero sugar variants). This form of tax has typically been motivated as a revenue raising measure. We therefore solve for the tax rate that generates equivalent tax revenue to the optimally set volumetric tax applied only to sugar sweetened beverages. All rates, and their impact on welfare, are detailed in Table 5.4.

Table 5.3: *Welfare effect of optimal tax under different market structures*

			Change in:						
			Welfare components						
		Optimal tax rate	Cons. surplus	Tax rev.	Drinks profits	Num. profits	Ext. costs	Total welfare
True ownership structure	No tax		–	–	–	–	–	–
($\theta_w = 1, \theta_b = 0$)	Tax	1.74	-691	558	-259	13	-509	129
			(-691)	(558)	(-259)	(13)	(-509)	(129)
Single brand firms	No tax		1024	0	-570	141	130	465
($\theta_w = 0, \theta_b = 0$)	Tax	1.86	222	634	-796	124	-434	617
			(-803)	(634)	(-226)	(-17)	(-564)	(152)
Joint profit maximization	No tax		-4560	0	1572	-31	-449	-2570
($\theta_w = 1, \theta_b = 1$)	Tax	1.35	-4927	345	1282	24	-772	-2504
			(-367)	(345)	(-290)	(56)	(-322)	(65)

Notes: Column (2) shows the optimal tax rate on the sugar in sweetened beverages under different firm ownership structures, columns (3)-(8) show changes in welfare and its components relative to the true ownership structure when no tax is in place. For single brand firms and joint profit maximization, numbers in parentheses show difference between tax and no tax equilibrium values. Num. profits refers to profits outside the drinks market (i.e. from the numeraire good), Ext. costs refers to externality costs. Numbers are in £million per year and are reported for $\phi = 4$ and $\tilde{\mu} = 0.4$. In Table G.3 in Appendix D we show differences the welfare gains from optimal tax across the different ownership structures are statistically significant.

Table 5.4: *Welfare changes under alternative tax bases*

		Change in:					
		Welfare components					
	Optimal tax rate	Cons. surplus	Tax rev.	Drinks profits	Num. profits	Ext. costs	Total welfare
Optimal sugar tax	1.74 per kg	-691	558	-259	13	-509	129
Optimal volumetric tax on SSBs	0.13 per l	-585	497	-229	-1	-394	76
Optimal ad valorem tax on SSBs	19 %	-539	879	-595	7	-308	60
Rev. eq. volumetric tax on soft drinks	0.06 per l	-536	497	-213	-3	-152	-102

Notes: Description of the different tax rates is provided in the text. Num. profits refers to profits outside the drinks market (i.e. from the numeraire good), Ext. costs refers to externality costs. Numbers are in £million per year and are reported for $\phi = 4$ and $\tilde{\mu} = 0.4$. In Table G.2 in Appendix D we show differences in welfare numbers across the different policies are statistically significant.

Setting a tax either volumetrically or on an ad valorem basis leads to welfare gains that are smaller than a tax levied directly on the sugar in sweetened beverages. These taxes are not directly levied on the source of externalities and this leads them to be less efficient at lowering the most socially costly consumption. For instance, unlike a tax on sugar they do not incentivize substitution between high to moderately sugary brands.[56] A volumetric tax on all soft drinks (revenue equivalent to one only on sugar sweetened soft drinks) leads to a fall in welfare; compared with the volumetric sugar sweetened beverage tax, it leads to a moderately smaller reduction in consumer surplus that is more than offset by larger declines in profits and a much smaller reduction in externalities. Even if policy is motivated by raising revenue, a tax on sugar sweetened beverages is a more efficient means of raising revenue than the broader based alternative.

6 Summary and conclusions

A number of consumer goods, such as alcohol and tobacco, have long attracted excise duties; the most compelling justification for these duties is to correct for the external costs associated with their consumption. Increasingly, food products, particularly those high in sugar, are attracting similar tax treatment. These markets typically consist of many differentiated products offered by large multi-product firms, and therefore they are likely to be characterized by market power.

We consider corrective tax design in differentiated products markets in which equilibrium price-cost margins and externalities vary across products, and show optimal tax policy depends on the nature of externalities, price-cost margins on externality generating and substitute products, and consumer switching patterns. We combine this framework with a detailed empirical model of demand and supply in the drinks market to assess the design of taxes on sugar sweetened beverages. We simulate the optimal tax on the sugar in sweetened beverages and show that, despite substantial market power for these products, high margins on substitute goods leads to a positive optimal tax rate. We also show that the optimal sugar tax improves welfare by more than policy that ignores distortions from the exercise of market power and various other tax instruments used in practice, and that corrective tax policy is complementary with the degree of competition in the market.

In this paper, we focus on firms' strategic pricing response; however, firms may respond to the tax by also changing their product portfolios and changing the sugar content of existing products. The nature of such responses will depend on the structure of the tax, as well as production costs and the shape of demand. Most sugar sweetened beverage taxes are volumetric and therefore do not incentivize firms to remove sugar from their products. An exception is the recently adopted UK Soft Drink Industry Levy, which entails a tax schedule with notches in product

[56]Grummon et al. (2019) argue the health benefits of a sweetened beverage tax levied on sugar would be 30% larger than one levied volumetrically.

sugar content. This results in large discontinuities in tax rates that provide strong incentives for product reformulation. An important direction for future research is to consider the implications for corrective tax policy design of these other margins of firm response. The variation provided from such non-linear tax systems will provide an opportunity to do this.

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.

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

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

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

Bovenberg, L. A. and L. H. Goulder (2002). Environmental Taxation and Regulation. In *Handbook of Public Economics*, Volume 3, pp. 1471–1545.

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

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

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

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

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

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

Colchero, M. A., J. C. Salgado, M. Unar-Munguía, M. Molina, S. Ng, and J. A. Rivera-Dommarco (2015). Changes in Prices After an Excise Tax to Sweetened

Sugar Beverages Was Implemented in Mexico: Evidence from Urban Areas. *PLOS ONE 10*(12), e0144408.

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

Competition Commission (2013). AG Barr / Britvic merger inquiry. Technical report.

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

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? *IFS Working Paper W20/8*.

Einav, L., A. Finkelstein, and P. Tebaldi (2019). Market Design in Regulated Health Insurance Markets: Risk Adjustment vs. Subsidies. *Mimeo*.

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

Ganapati, S., J. Shapiro, and R. Walker (2019). Energy Cost Pass-Through in U.S. Manufacturing: Estimates and Implications for Carbon Taxes. *American Economic Journal: Applied Economics forthcoming.*

Global Food Research Program (GFRP) (2019). Sugary Drink Taxes around the World. http://globalfoodresearchprogram.web.unc.edu/multi-country-initiative/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, 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.

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

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

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

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*.

House of Commons Health Committee (2018). Childhood obesity: Time for action. Technical report.

Jacobsen, M. R., C. R. Knittel, J. M. Sallee, and A. A. van Benthem (2018). The Use of Regression Statistics to Analyze Imperfect Pricing Policies. *Journal of Political Economy forthcoming*.

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.

Kopczuk, W. (2003). A note on optimal taxation in the presence of externalities. *Economics Letters 80*(1), 81–86.

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

Levell, P., M. O'Connell, and K. Smith (2016). Chapter 9: Excise Duties. In *Green Budget 2016*, Volume 2016, pp. 201–231. London: IFS.

Mahoney, N. and E. G. Weyl (2017, October). Imperfect Competition in Selection Markets. *The Review of Economics and Statistics 99*(4), 637–651.

Miller, N. H. and M. C. Weinberg (2017, November). Understanding the Price Effects of the MillerCoors Joint Venture. *Econometrica 85*(6), 1763–1791.

Miravete, E. J., K. Seim, and J. Thurk (2018a). Market Power and the Laffer Curve. *Econometrica 86*(5), 1651–1687.

Miravete, E. J., K. Seim, and J. Thurk (2018b). One Markup to Rule Them All: Taxation by Liquor Pricing Regulation. *American Economic Journal: Microeconomics forthcoming*.

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.

O'Donoghue, T. and M. Rabin (2006). Optimal sin taxes. *Journal of Public Economics 90*(10-11), 1825–1849.

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*.

Roberto, C. A., H. G. Lawman, M. T. LeVasseur, N. Mitra, A. Peterhans, B. Herring, and S. N. Bleich (2019, May). Association of a Beverage Tax on Sugar-Sweetened and Artificially Sweetened Beverages With Changes in Beverage Prices and Sales at Chain Retailers in a Large Urban Setting. *JAMA 321*(18), 1799.

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

Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics 83*(2), 217–230.

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

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

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

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

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

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
April, 2020

A Optimal tax formulae

A.1 Many differentiated goods

Here we derive the many product optimal tax expression given by equation (2.3) in Proposition 1 in the main body of the paper. Suppose there are J goods; one subset, denoted $j \in \mathcal{S}$, contain an attribute associated with an externality, z_j; the remaining subset of products, denoted $j \in \mathcal{N}$, have $z_j = 0$.

The planner sets a single tax rate, $\tau_{\mathbb{f}}$, directly on a particular feature of the externality generating goods, which we denote by $\mathbb{f}_j$. The products are supplied imperfectly competitively and at constant marginal costs, so we can write:

$$\begin{aligned} p_j - \tau_{\mathbb{f}} \times \mathbb{f}_j - c_j &= \mu_j \quad \forall j \in \mathcal{S} \\ p_j - c_j &= \mu_j \quad \forall j \in \mathcal{N}. \end{aligned}$$

where μ_j denotes the equilibrium price-cost margin for product j.

The efficiency maximizing social planner chooses the tax rate to:

$$\max_{\tau_{\mathbb{f}}} v(\mathbf{p}) - \sum_{j \in \mathcal{S}} \phi z_j q_j + \sum_j (p_j - c_j) q_j. \tag{A.1}$$

The first order condition of the planner's problem is:

$$\sum_j (p_j - c_j - \phi z_j) \frac{dq_j}{d\tau_{\mathbb{f}}} = 0.$$

A tax on the externality generating attribute. Suppose the planner levies the tax on the externality generating attribute (i.e. $\mathbb{f} = z$). In this case we can re-express the planner's first condition as:

$$\begin{aligned}\tau_z^* =&\phi - \frac{\sum_j \mu_j \frac{dq_j}{d\tau_z}}{\sum_{j\in\mathcal{S}} z_j \frac{dq_j}{d\tau_z}} \\ =&\phi - \frac{1}{\sum_{j\in\mathcal{S}} z_j \frac{dq_j}{d\tau_z}}\left(\sum_{j\in\mathcal{S}}\mu_j \frac{dq_j}{d\tau_z} + \sum_{j\in\mathcal{N}}\mu_j \frac{dq_j}{d\tau_z}\right).\end{aligned}$$

Defining $\frac{dQ^{\mathcal{X}}}{d\tau_z} = \sum_{j\in\mathcal{X}} \frac{dq_j}{d\tau_z}$, $w_j^{\mathcal{X}} = \frac{dq_j}{d\tau_z} \Big/ \frac{dQ^{\mathcal{X}}}{d\tau_z}$ for $\mathcal{X} = \{\mathcal{S}, \mathcal{N}\}$ and $\frac{dZ}{d\tau_z} = \sum_{j\in\mathcal{S}} z_j \frac{dq_j}{d\tau_z}$, we can re-write this as:

$$\tau_z^* = \phi - \frac{1}{\frac{dZ}{d\tau_z}\Big/\frac{dQ^{\mathcal{S}}}{d\tau_z}}\left(\sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}\mu_j - \sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}\mu_j \times \frac{dQ^{\mathcal{N}}}{d\tau_z}\Big/\left(-\frac{dQ^{\mathcal{S}}}{d\tau_z}\right)\right), \tag{A.2}$$

which is the expression given by equation (2.3) in Proposition 1.

A volumetric tax. Suppose instead the tax is applied volumetrically to products in the set $\mathcal{S}$ i.e. $\mathbb{f}_j = \mathbb{1}_{z_j>0}$. In this case the optimal tax rate can be written:

$$\begin{aligned}\tau_v^* =&\phi\frac{\sum_{j\in\mathcal{S}} z_j \frac{dq_j}{d\tau_v}}{\sum_{j\in\mathcal{S}} \frac{dq_j}{d\tau_v}} - \frac{\sum_j \mu_j \frac{dq_j}{d\tau_v}}{\sum_{j\in\mathcal{S}} \frac{dq_j}{d\tau_v}} \\ =&\frac{dZ}{d\tau_v}\Big/\frac{dQ^{\mathcal{S}}}{d\tau_v} \times \phi - \left(\sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}\mu_j - \sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}\mu_j \times \frac{dQ^{\mathcal{N}}}{d\tau_v}\Big/\left(-\frac{dQ^{\mathcal{S}}}{d\tau_v}\right)\right).\end{aligned}$$

For the optimal volumetric tax $\frac{dZ}{d\tau}/\frac{dQ^{\mathcal{S}}}{d\tau}$ (which is the effect of a marginal change in the tax rate on intake of the externality generating attribute divided by its effect on consumption of the set of products that contain the attribute) pre-multiplies the externality; for the optimal tax on the externality generating attribute, z, the inverse of the expression pre-multiplies the equilibrium margin terms. This difference reflects the different bases of the two taxes (and that externalities are per unit of z and margins are pre volume of product). Note that the equilibrium margins μ_j, margin weights, w_j, and switching derivatives, $\frac{dQ^{\mathcal{N}}}{d\tau}/\left(-\frac{dQ^{\mathcal{S}}}{d\tau}\right)$ and $\frac{dZ}{d\tau}/\frac{dQ^{\mathcal{S}}}{d\tau}$ are all implicit functions of the tax and therefore will vary between the two forms of tax.

A.2 Extensions

Consider the case of a planner levying a tax on attribute z.

Heterogeneous externalities. Suppose that the marginal externalities are consumer specific and denoted by ϕ_i. It may be, for instance, that marginal externalities are constant in individual consumption of attribute z, $Z_i = \sum_j z_j q_{ij}$, but heterogeneous across individuals. Alternatively, it may be that marginal externalities are a non-linear function of individual consumption, $\Phi(Z_i)$, in which case, in the optimal tax formula ϕ_i should be interpreted as the marginal consumption externality of individual i at their equilibrium consumption level. The planner's first order condition is then:

$$\sum_j (p_j - c_j - \phi_i z_j)\frac{dq_j}{d\tau} = 0.$$

Defining $\frac{dZ_i}{d\tau} = \sum_{j\in\mathcal{S}} z_j \frac{dq_{ij}}{d\tau}$ (the impact of a marginal change in the tax rate on the equilibrium usage of the externality generating attribute by individual i), then optimal tax rate (levied on the externality generating product attribute) can be written:

$$\tau^* = \sum_i \phi_i \frac{dZ_i}{d\tau}\Big/\frac{dZ}{d\tau} - \frac{1}{\frac{dZ}{d\tau}\Big/\frac{dQ^{\mathcal{S}}}{d\tau}}\left(\sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}\mu_j - \sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}\mu_j \times \frac{dQ^{\mathcal{N}}}{d\tau}\Big/\left(-\frac{dQ^{\mathcal{S}}}{d\tau}\right)\right). \tag{A.3}$$

In the discussion in the main body of the paper we use ω_i to denote $\frac{dZ_i}{d\tau}/\frac{dZ}{d\tau}$.

Comparison of the expressions for the optimal tax formulation under homogeneous and heterogeneous externalities (equations (A.2) and (A.3)) suggest it is only the first term that differs between the two. This, however, is misleading, as equilibrium price-cost margins and all derivatives of equilibrium quantities depend implicitly on the tax rate. Therefore the numerical value of all parts of the expression are likely to vary depending on whether or not there is heterogeneity in externalities.

Externality leakage 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 therefore are also untaxed). It is useful to denote the products in set (ii) and (iii) by $\mathcal{N}$. Define $\frac{dZ^{\mathcal{S}}}{d\tau} = \sum_{j\in\mathcal{S}} z_j \frac{dq_j}{d\tau}$ and $\frac{dZ^{\mathcal{A}}}{d\tau} = \sum_{j\in\{\mathcal{S}\bigcup\mathcal{L}\}} z_j \frac{dq_j}{d\tau}$ 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. Suppose externalities are homogeneous. In this case the optimal rate of tax can be expressed:

$$\tau^* = \phi \times \frac{dZ^{\mathcal{A}}}{d\tau}\Big/\frac{dZ^{\mathcal{S}}}{d\tau} - \frac{1}{\frac{dZ^{\mathcal{S}}}{d\tau}\Big/\frac{dQ^{\mathcal{S}}}{d\tau}}\left(\sum_{j\in\mathcal{S}} w_j^{\mathcal{S}}\mu_j - \sum_{j\in\mathcal{N}} w_j^{\mathcal{N}}\mu_j \times \frac{dQ^{\mathcal{N}}}{d\tau}\Big/\left(-\frac{dQ^{\mathcal{S}}}{d\tau}\right)\right).$$

Full externality internalization. To illustrate the intuition behind the full externality internalization result, we return to the two product market setting outlined in Section 2.1. Suppose the planner wishes to ignore distortions associated with the exercise of market power and aims at full externality internalization. The planner will fail to maximize welfare as specified in equation (2.1). In particular, suppose the planner chooses to treat the equilibrium allocation (or equivalently price-cost margins) in the absence of tax as a reference point, with the aim of inducing agents to internalize externalities relative to this benchmark. In this case the planner will maximize a modified welfare function in which the marginal cost of each product is replaced with its equilibrium price when $\tau = 0$, which we denote by $\bar{p}_j$. Specifically, the planner will solve:

$$\max_{\tilde{\tau}} v(\mathbf{p}) - \phi q_1 + (p_1 - \bar{p}_1)q_1 + (p_2 - \bar{p}_2)q_2.$$

Define the equilibrium pass-through rate of the tax onto the consumer price of product 1 as $\rho = \frac{p_1 - \bar{p}_1}{\tau}$ and denote the change in equilibrium price for product 2 resulting from the tax as $\Delta p_2 = p_2 - \bar{p}_2$. We can then express the tax that maximizes the planner's modified welfare function as:

$$\tilde{\tau}^* = \frac{1}{\rho}\left(\phi + \Delta p_2 \times \frac{dq_2}{d\tilde{\tau}} \Big/ \left(-\frac{dq_1}{d\tilde{\tau}}\right)\right). \tag{A.4}$$

$\tilde{\tau}^*$ differs from the Pigovian prescription ($\tau = \phi$) for two reasons. First, the tax rate depends inversely on the equilibrium pass-through rate. Second, the tax depends positively (negatively) on any increase (decrease) in the equilibrium price of the substitute good. The strength of this effect is in proportion to how much of any reduction of equilibrium consumption of product 1 in response to a marginal tax switches to product 2. If $\frac{dq_2}{d\tilde{\tau}} / \left(-\frac{dq_1}{d\tilde{\tau}}\right) = 1$, then $\tilde{\tau}^* = \frac{1}{\rho}(\phi + \Delta p_2)$; while if $\frac{dq_2}{d\tilde{\tau}} / \left(-\frac{dq_1}{d\tilde{\tau}}\right) = 0$, then $\tilde{\tau}^* = \frac{1}{\rho}\phi$. In each case, in equilibrium, the difference in equilibrium prices for the taxed and substitute good equals the price difference in the absence of any tax plus the marginal externality, $p_1 - p_2 = \bar{p}_1 - \bar{p}_2 + \phi$.

In the many product case the planner's first order condition is:

$$\sum_j (p_j - \bar{p}_j - \phi z_j)\frac{dq_j}{d\tilde{\tau}} = 0.$$

The tax rate that solves this condition is implicitly defined by:

$$\tilde{\tau} = \frac{1}{\sum_{j \in \mathcal{S}} \rho_j \varpi_j}\left(\phi + \sum_{j \in \mathcal{N}} w_j^{\mathcal{N}} \Delta p_j \times \frac{dQ^{\mathcal{N}}}{d\tilde{\tau}} \Big/ \left(-\frac{dZ}{d\tilde{\tau}}\right)\right). \tag{A.5}$$

$\sum_{j\in\mathcal{S}} \rho_j \varpi_j$ is the weighted average pass-through rate across products. The weights are the contribution product j makes to the derivative of equilibrium consumption of attribute z with respect to the tax rate; $\varpi_j = \frac{dZ_j}{d\tilde{\tau}} / \frac{dZ}{d\tilde{\tau}}$. This expression is a natural generalization of the two good formula (equation A.4).

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 planner's first order condition then takes the form:

$$\sum_j (p_j - c_j - \overline{\phi}_i z_j) \frac{dq_j}{d\tau} = 0.$$

where $\overline{\phi}_i = \phi_i + \varphi_i$ is the sum of the marginal externality and internality. The optimal tax can then be written as in equation A.3, but with $\overline{\phi}_i$ 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 $\Omega = \mathcal{S} \bigcup \mathcal{N}$ 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

In Table B.1 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). Across the regular variants there is variation in sugar contents. This variation in sugar content means a tax levied directly on sugar will have different implications to one levied volumetrically (i.e. per liter of product sold). 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.1: *Brands, sugar contents and sizes*

			Sugar	Number of sizes	
Firm	Brand	Variant	(g/100ml)	At-home	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.

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 non-alcoholic drink on that day, (ii) how much they purchased, and (iii) the share of volume of drinks purchased on sale. To account for households who do not record purchasing any groceries for a sustained period of time (for instance, because they are on holiday), we construct "purchase strings" for each households. These are periods that do not contain a period of non-reporting of any grocery purchases longer than 3 or more weeks.

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

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

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

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

If stockpiling behavior is present we would expect that $\beta^{\text{inv,pp}} < 0$ and $\beta^{\text{inv,q}} < 0$; when a household's inventory is high it is less likely to purchase, and conditional on purchasing it will buy relatively little. The first two columns of Table C.2 summarize the estimates from these regressions. There is a small positive relationship between inventory and purchase probability and quantity purchased, conditional on buying. An increase in inventory of 1 liter leads to an increase in the probability of buying of 0.001, relative to a mean of 0.23, and an increase in the quantity purchased, conditional on buying a positive amount, of 0.013, relative to a mean of 3.925. These effects are both very small and go in the opposite direction to that predicted by Hendel and Nevo (2006b) if stockpiling behavior was present.

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

We check for this by estimating:

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

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

Stockpiling behavior should lead to $\beta^{\text{lead}} > 0$ and $\beta^{\text{lag}} < 0$. Columns (3) and (4) of Table C.2 summarize the estimates from these regressions. We estimate that purchasing on sale is associated with an increase of 0.14 days to the next purchase and 0.23 days less since the previous purchase. The sign of these effects are consistent with stockpiling, however their magnitudes are 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:

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

where $\text{brandswitch}_{i\tau}$ is a dummy variable equal to 1 if the household purchased a brand that they did not buy the last time they visited the store, $\text{sizeswitch}_{i\tau}$ is a dummy variable equal to 1 if the household purchased a size that they did not buy the last time they visited the store, and $\text{packtypeswitch}_{i\tau}$ is a dummy variable equal to 1 if the household purchased a pack type that they did not buy the last time they visited the store.

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 at $\tau-1$; 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 at $\tau-1$; 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 at $\tau-1$. 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

We estimate the demand model using simulated maximum likelihood. We allow all parameters to vary by consumer group and estimate the choice model separately by groups.[57] 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.

[57] In the at-home segment, for each group, we use a random sample of 1,500 households and 10 choice occasions per households; 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.

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	× Price	0.129	0.110	0.125	0.181	0.180	0.120
with low income		(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; mark-ups are margins over price. 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 of the market. 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).[58] 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, 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.[59]

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

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

where p_{jrt} denotes the price per liter of product j in retailer r in week t,[60] τ_t are week effects, ξ_j are product fixed effects, and ρ_r are retailer fixed effects.

[58] 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.

[59] 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.

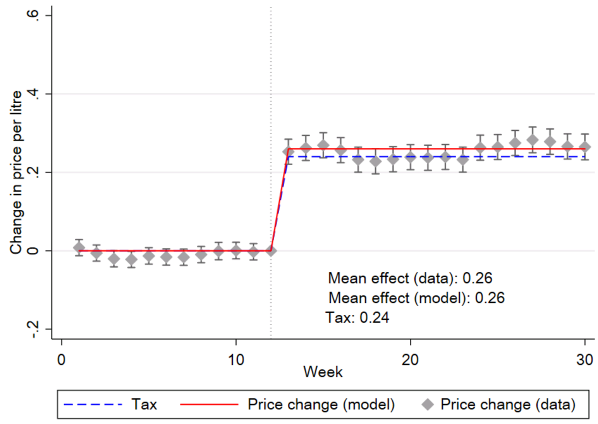
[60] This is the VAT-exclusive price per liter.

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.

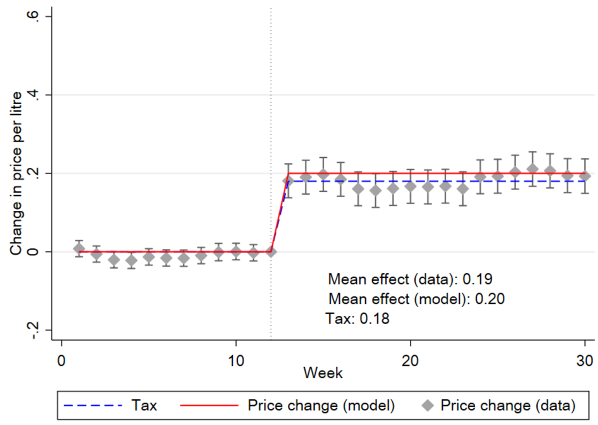
We simulate the introduction of the SDIL using our estimated model of demand and supply in the non-alcoholic drinks market (based on product sugar contents when the SDIL was implemented). The red lines plot the average price increase for each of the three group predicted by our model. These match very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

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

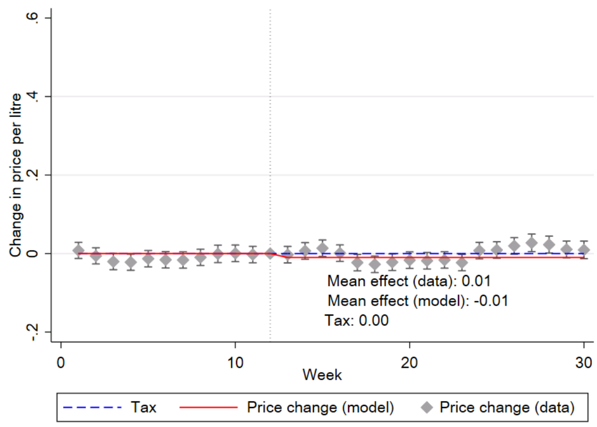
(a) High treatment group



(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 Empirical implementation of optimal tax problem

Let $\mathbf{p}_m = (p_{1m}, \ldots, p_{Jm})$ denote the equilibrium price vector in market m, $\mathbf{q}_m = (q_{1m}, \ldots, q_{Jm})$ denote the equilibrium vector of quantities, and $\mathbf{c}_m = (c_{1m}, \ldots, c_{Jm})$ denote marginal costs. Equilibrium prices and quantities depend on the level of any tax rate levied on the products. Denote by Y_m total consumer income in market m; total spending on the numeraire good is then $X_m = Y_m - \sum_j p_{jm}q_{jm}$. We denote the price-cost mark-up on the numeraire good by $\tilde{\mu}$.

The planner sets a tax rate τ on the product attribute z. Assume there is a marginal externality of ϕ associated with 1 unit of attribute z. We denote the subset of products for which $z_j > 0$ by $\mathcal{S}$. The planner's problem is:

$$\max_{\tau} \sum_m \left(v(\mathbf{p}_m) - \sum_{j \in \mathcal{S}} \phi z_j q_{jm} + \sum_j (p_{jm} - c_{jm}) q_{jm} + \tilde{\mu} X_m \right),$$

and first order condition is:

$$\sum_m \sum_j (p_{jm} - c_{jm} - \phi z_j) \frac{dq_{jm}}{d\tau} + \tilde{\mu} \frac{dX_m}{d\tau} = 0.$$

We compute the optimal tax rate by searching for the τ that solves this implicit non-linear equation. In order to do this, for each candidate tax rate, we must solve for the equilibrium prices and their tax derivative. To find the equilibrium price vector we solve the system of equations defined by firms' first order conditions, discussed in Section 4.3 and repeated here: $\forall\ j$

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

To solve for the derivative of equilibrium prices with respect to the tax we solve the system of equations defined by the derivative of firms' first order conditions with respect to the tax rate: $\forall\ j$

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

G Additional optimal tax results

Here we present optimal tax rates and welfare changes and their confidence intervals. We create confidence intervals by drawing from the joint normal asymptotic distribution of the parameter estimates and, for each draw, computing the statistic of interest. We then use the resulting distribution across draws to compute Monte Carlo 95% confidence intervals. These intervals need not be symmetric.

In Table G.1 we report the various tax rates we discuss in the paper, along with their confidence intervals. In Table G.2 for all of these tax policies (with the exception of the optimal sugar tax) we report the difference between their impact on welfare and its components and the impact under the optimal sugar tax. This means the associated confidence intervals tell us whether these tax regimes have an impact on welfare that is statistically significantly different than the impact of the optimal sugar tax. The welfare gain from each of these tax regimes is statistically significantly different from the gain achieved by the optimal sugar tax.

In column (2) of Table G.3 we report the optimal sugar tax under the counterfactual market structures of single brand firms and joint profit maximization. In the remaining columns we report the difference between the impact of the optimal sugar tax under these ownership structures and the impact of the optimal sugar tax under the true ownership structure. The confidence intervals on columns (3)-(6) therefore show that the welfare changes under these conterfactual ownership structures are statistically significantly different from the changes achieved by optimal tax under the true ownership structure.

Table G.1: *Alternative tax rates*

Optimal sugar tax	Pigovian tax	Rate aiming at at full internalization	Optimal volumetric tax on SSBs	Optimal ad valerom tax on SSBs	Rev. equiv volumetric tax on soft drinks
1.74	4.00	3.64	0.13	0.19	0.06
[1.62, 1.85]	[4.00, 4.00]	[3.54, 3.68]	[0.12, 0.14]	[0.17, 0.20]	[0.06, 0.07]

Notes: The table shows the tax rates discussed in Sections 5.1 and 5.5 of the paper. 95% confidence intervals are given in square brackets..

Table G.2: *Welfare differences relative to optimal sugar tax*

	Welfare components					
	Cons. surplus	Tax rev.	Drinks profits	Num. profits	Ext. costs	Total welfare
Pigovian tax	-654	346	-230	48	-381	-108
	[-677, -636]	[337, 358]	[-240, -223]	[39, 59]	[-397, -372]	[-124, -92]
Rate aiming at full internalization	-562	308	-198	39	-334	-81
	[-582, -547]	[300, 317]	[-207, -192]	[32, 49]	[-347, -325]	[-94, -66]
Optimal vol. tax on SSBs	106	-61	30	-14	115	-54
	[101, 112]	[-63, -58]	[27, 32]	[-16, -12]	[110, 120]	[-58, -50]
Optimal ad val. tax on SSBs	152	321	-336	-6	200	-69
	[134, 175]	[314, 328]	[-350, -322]	[-10, -2]	[190, 212]	[-81, -58]
Rev. equiv. vol. tax on soft drinks						
	[179, 212]	[-84, -71]	[-40, -24]	[104, 119]	[287, 308]	[-116, -87]

Notes: The table shows the impact of the tax policy shown in the first column on welfare and its components relative to the impact of the optimal sugar tax. Rows (1) and (2) correspond to row (2)-(1) and row (3)-(1) of Table 5.2. Rows (3), (4) and (5) correspond to rows (2)-(1), (3)-(1) and (4)-(1) of Table 5.4. Num. profits refers to profits outside the drinks market (i.e. from the numeraire good), Ext. costs refers to externality costs. Numbers are in £million per year and are reported for $\phi = 4$ and $\tilde{\mu} = 0.4$. 95% confidence intervals are given in square brackets.

Table G.3: *Welfare effect of optimal tax under different market structures*

		Change in:					
		Welfare components					
	Optimal tax rate	Cons. surplus	Tax rev.	Drinks profits	Num. profits	Ext. costs	Total welfare
Single brand firms	1.86	-112	76	33	-30	-56	23
	[1.75, 2.00]	[-158, -69]	[49, 103]	[19, 2106]	[-1540, -27]	[-86, 560]	[-74, 27]
Joint profit maximization	1.35	324	-213	-31	43	186	-64
	[1.10, 1.58]	[276, 378]	[-265, -166]	[-75, 14]	[37, 50]	[142, 232]	[-74, -56]

Notes: Column (2) shows the optimal sugar tax rate for the market structure reported in column (1). The remaining columns show the impact of the optimal sugar tax under these counterfactual market structures on welfare and its components relative to the impact of the optimal sugar tax under the true ownership structure. Row (1) corresponds to the terms in paranthesis in the middle panel of Table 5.3 minus the numbers in paranthesis in the top panel. Row (2) corresponds to the terms in paranthesis in the bottom panel of Table 5.3 minus the numbers in paranthesis in the top panel. Num. profits refers to profits outside the drinks market (i.e. from the numeraire good), Ext. costs refers to externality costs. Numbers are in £million per year and are reported for $\phi = 4$ and $\tilde{\mu} = 0.4$. 95% confidence intervals are given in square brackets.