How well targeted are soda taxes?

Pierre Dubois, Rachel Griffith and Martin O'Connell*

December 2018

Abstract

Soda taxes aim to reduce excessive sugar consumption. Their effectiveness depends on whether they target individuals for whom the harm of consumption is largest. We study individual level purchases made on-the-go, which account for around half of sugar from soft drinks. We estimate demand and account for supply-side equilibrium pass-through. We exploit longitudinal data to estimate individual preferences, which allows flexible heterogeneity that we relate to key individual characteristics. We show that soda taxes are relatively effective at targeting young consumers but not individuals with high total dietary sugar; they impose the highest monetary cost on poorer individuals, but are unlikely to be strongly regressive especially if we account for averted future costs from over consumption.

Keywords: preference heterogeneity, discrete choice demand, pass-through, soda tax

JEL classification: D12, H31, I18

Acknowledgements: The authors 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, the Open Research Area (ORA) grant number ES/I012222/1, ANR-10-ORAR-009-01 and ES/N011562/1 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 remained the responsibility of the authors.

*Correspondence: Dubois: Toulouse School of Economics, pierre.dubois@tse-fr.eu; Griffith: Institute for Fiscal Studies and University of Manchester, rgriffith@ifs.org.uk; O'Connell: Institute for Fiscal Studies, martin_o@ifs.org.uk

1 Introduction

Sugar consumption is far in excess of recommended levels in much of the developed world, is strongly linked with a range of diet-related diseases, including diabetes, cancers and heart disease, and is particularly detrimental to children (WHO (2015)). Soft drinks are an important contributor to excess sugar consumption (CDC (2016)) particularly in the young (Han and Powell (2013) and Cavadini et al. (2000)). "Soda taxes", taxes levied on soft drinks, have been proposed as a way to reduce sugar consumption, particularly for individuals whose consumption generates costs that are borne by others (externalities) or for whom the future costs of excess consumption are large and are partially ignored at the point of consumption (internalities). Internality correcting taxes have been advocated for unhealthy foods (O'Donoghue and Rabin (2006), Haavio and Kotakorpi (2011)), as the principal justification for high levels of cigarette taxation (Gruber and Koszegi (2004)), and in energy markets (Allcott et al. (2014)). A growing number of jurisdictions are adopting soda taxes.[1] Whether such measures will succeed in improving public health depends on how individuals' demand responses correlate with the size of any unanticipated costs that their consumption imposes on themselves in future and costs imposed on others.

Our contribution in this paper is to provide evidence on how well targeted soda taxes are; in particular, are they effective at lowering the sugar consumption of individuals for whom the consequences of high intake are most severe. We estimate consumer choice in the drinks market and simulate the introduction of a soda tax, accounting for pass-through to prices. We show that soda taxes are relatively effective at targeting young consumers, are less effective at targeting individuals with high total dietary sugar, impose somewhat higher monetary cost on poorer individuals, but are unlikely to be strongly regressive, especially if we account for averted future costs from over consumption. Relative to the existing literature we make two main advances.

First, we model consumer preferences as individual level parameters that we estimate. This departs from the standard approach to modeling consumer preference heterogeneity in discrete choice models, where preferences are treated as random effects drawn from a mixing distribution. The main advantage of our approach is that we do not need to make assumptions restricting or ruling out correlation in consumer level preferences with consumer attributes (including purchase behavior for other goods). We are therefore able to directly relate individual level predictions

[1] These include a number of US cities, including Philadelphia and San Francisco, as well as France, Mexico and the UK.

of the impact of the tax to consumer characteristics in a flexible way. This means that we can assess precisely which individuals respond to the tax and on whom the economic burden of the tax falls most heavily; in other words is a tax well targeted and how regressive is it?

Second, we study individual purchase decisions made for immediate consumption on-the-go using novel longitudinal data on a representative sample of British individuals (including teenagers and young adults). Around half of sugar from soft drinks is obtained on-the-go, making it an important part of the market on which we have little evidence on choice behavior. On-the-go purchases are made by individuals for immediate consumption – most of the literature on choice behavior studies purchases made in supermarkets and brought into the home for future consumption. A significant advantage of individual level on-the-go data is that they allow us to estimate individual level preferences, and individual level responses to tax, without the need to place strong restrictions on the intra-household preference structure (see, for example, Adams et al. (2014)). In addition, young adults are a particular group of interest and are typically not identified as a distinct group in data based on household purchases.

We are interested in how well targeted soda taxes are; that is do they lead to the largest reductions in sugar by individuals whose consumption is most likely to create social costs. The propensity for people to over consume sugar, the effects that excessive intake has on health and other future outcomes, and the role soft drinks play as a significant contributor to total dietary sugar are well established (see WHO (2015)). It is also well established that young people, and those for whom sugar represents a high share of the total calories that they purchase (high total dietary sugar individuals), tend to get particularly large amounts of sugar from soft drinks.[2]

These facts have motivated the implementation of soda taxes in many jurisdictions, and they suggest that soda taxes may be well targeted – sugar consumption is well above medical recommendations, products subject to soda taxes represents a substantial share of this, and their intake is especially high for the young and for individuals with high total dietary sugar – both groups for whom high soft drinks consumption is likely to be particularly costly (see Gortmaker et al. (2009), Currie (2009)).

However, the effectiveness of a soda tax depends not only on the extent to which individuals consume soft drinks prior to the introduction of the tax, but

[2]In Appendix A.1 we provide evidence of this based on dietary intake data from the UK and the US.

also on how strongly they switch away from the sugar in these products and what alternatives they switch to. To assess the targeting of the tax we need to know how demand responses vary across markers of likely harm from consumption (like age and total dietary sugar); to assess the redistributive consequences we need to know how they vary across the income distribution. We estimate a structural model of demand and supply that allows us to identify individual specific preference parameters and enables us to relate the effects of a soda tax in a flexible way to individual demographics, measures of their broader diet, and measures of income.

To model consumer choice we use a discrete choice framework in which consumer preferences are defined over product attributes. Like much of the literature on choice models (Berry et al. (1995), Nevo (2001), Train (2003)), we allow for consumer specific preference parameters. However, we depart from the standard approach by treating these preferences as consumer level parameters to be estimated (rather than random draws from a mixing – or random coefficient – distribution). This means that we can recover any arbitrary relationship between the individual preference parameters and functions of them, such as the predicted outcomes from a tax simulation, with any attributes of the individual consumers. In contrast, in standard random coefficient models, it is necessary to specify ex ante and with a particular functional form how preferences depend on exogenous characteristics; independence is assumed between the preference distribution and all other individual level attributes. While much more flexible in this dimensions, our approach entails estimating fixed effects in a non-linear model and therefore may suffer from an incidental parameters problem (Hahn and Newey (2004), Arellano and Hahn (2007)). We show robustness to this using the split sample jackknife bias correction procedure suggested in Dhaene and Jochmans (2015).

We find that preferences vary over consumer attributes in ways that would be difficult to capture by specifying a priori a random coefficient distribution. For instance, our estimates show that, on average, those aged 13-21 have stronger preferences for sugar than individuals aged 22-30, who in turn have stronger sugar preferences than older individuals. Among the youngest age group preferences over sugar and price are uncorrelated, but for older individuals they are positively correlated – those with strong sugar preferences tend to be the least price sensitive. We also show how the effects of a soda tax vary over the joint distribution of age, total dietary sugar and a proxy for income, while placing minimal restrictions on the joint distribution of preferences; we impose that an individual's preferences are stable over choice occasions, which allows us to use the long time dimension of repeated purchases to identify individual preferences.

We model tax pass-through assuming that drinks manufacturers compete by simultaneously setting prices in a Nash-Bertrand game. We abstract from modeling manufacturer-retailer relationships, but discuss how efficient vertical contracting would lead to such a price equilibrium. We allow possible differences in manufacturers' responsiveness to the tax across different retailers. The market demand curves faced by firms (and relevant for product pricing) depend both on behavior in the on-the-go segment of the market, and in the at-home segment. We use household level data to estimate at-home demand, thereby enabling us to take account of the effect of both segments on firm pricing. Our estimates suggest that an excise style tax on sugary soft drinks would be over shifted on to consumer prices and lead to marginally lower prices of diet products. Firms' pricing response would therefore amplify the price differential that the tax creates between sugary and diet varieties.

Our main interest in this paper is to use individual level data in the on-the-go segment to explore how well targeted soda taxes are (however, we show at-home household level responses are unlikely to undo individual level on-the-go patterns of response). We show that the sugary soda tax is relatively well targeted at young people. Those aged below 50 are considerably more likely to purchase soft drinks that older people. Reductions in sugar from the tax are highest in level terms for those aged 13-21, who exhibit an average reduction in sugar from drinks that is around 40% higher than those aged over 40. This is driven by young people being much more likely to obtain large quantities of sugar from soft drinks products – in percentage terms they actually lower their sugar consumption by less than older people.

The tax is less effective at targeting those people with a consistently high level of dietary sugar in their overall diets; despite those with high sugar diets being more likely to purchase soft drinks and to obtain relatively large amounts of sugar from it, their sugar intake from drinks responds less strongly in level terms (as well as percentage terms) than those with more moderate levels of sugar in their diets. This is driven by individuals with a high level of total dietary sugar tending to have a strong preference for sugar and be less price responsive.

If consumers fully internalized the future costs of excess sugar consumption, we could measure the full effect on consumer welfare using individuals' revealed preferences to compute their compensating variations. However, if some people do not fully account for the future costs at the point of consumption, then the tax will have a second effect on individual level welfare through averted future unanticipated costs (internalities). We estimate compensating variation, and show that it is highest among individuals with high total dietary sugar and among young consumers (especially young consumers with high total dietary sugar). While there is experimental evidence that people have behavioral biases with respect to food and drink consumption (see, for instance, Read and Van Leeuwen (1998) and Gilbert et al. (2002)), measuring the extent of these internalities is challenging, and not something we attempt to do in this paper. However, we can get an idea of the full effect on consumer welfare by computing how much internality per reduction in sugar is required to make people indifferent to the introduction of a soda tax. For the group of young consumers this number is around £1.11 per typical 330ml can of sugary soft drink.

A common criticism of excise style taxes is they are regressive;[3] the poor typically spend a higher share of their income on the taxed good, and so bear a disproportional share of the burden of the tax. However, if the tax plays the role of correcting an internality, then the distributional analysis is more complicated; if low income consumers also save more from averted internalities this may overturn the regressivity of the traditional economic burden of taxation (Gruber and Koszegi (2004)). These redistributive concerns become more subtle when income transfers are considered (Allcott et al. (2018a), Allcott et al. (2018b)). We show that compensating variation associated with a sugary soda tax is around 25% higher for those in the bottom half of the distribution of total annual expenditure (based on a wide set of food, drink and non-drink items) compared with those in the top half. However, the reduction in sugar is also larger for these individuals, which leaves open the possibility that they will also benefit more from averted internalities, and so the full effect on their welfare is likely to be less negative than the compensating variation suggests.

The rest of this paper is structured as follows. In Section 2 we introduce the data and describe the non-alcoholic drinks market. In Section 3 we describe our model of consumer demand and oligopoly pricing and summarize estimates of the demand model. In Section 4 we present results of the sugary soda tax simulation, discussing the impact on equilibrium pricing, how well targeted the measure is, the effects on consumer welfare and its distributional implications. In Section 5 we consider three possible concerns about the robustness of our conclusions; first, we incorporate broader patterns of consumer switching, including towards food, and show that our results are robust to inclusion of these additional margins of consumer response; second, we consider the impact of at-home demand for non-alcoholic drinks on the targeting of soda taxes; third, we consider possible bias that could arise due to the incidental parameters problem. A final section concludes.

[3]For instance, see Senator Sanders op-ed on the Philadelphia soda tax, Sanders (2016).

2 The non-alcoholic drinks market

In this paper we consider demand for non-alcoholic drinks. This includes *soft drinks* (i.e. carbonated drinks – often referred to as soda – with and without sugar, energy drinks, and other sugar-sweetened non-alcoholic drinks), *alternative sugary drinks* (naturally sweet non-alcoholic drinks such as fruit juice and milk based drinks such as shakes), and *bottled water*. "Soda taxes" are typically imposed on soft drinks products that contain sugar (and sometimes also on diet varieties).[4] Juices and milk based drinks are usually exempt. The existing literature on soda taxes study their impact on purchases made in grocery stores for future consumption at home; but close to half of sugar obtained from sugar sweetened soft drinks products is purchased for immediate consumption on-the-go.[5]

Our primary interest is in behavior when purchasing on-the-go. We focus on the on-the-go segment for two reasons. First, it is an important part of the market and an important source of sugar, particularly in children (Han and Powell (2013))), yet little attention has been paid to modeling choice behavior on-the-go, largely due to the lack of high quality data.[6] Second, studying on-the-go behavior provides the opportunity to model and exploit data on *individual* level purchases, including those made by teenagers and young adults. This provides an important opportunity to study the preferences of individuals, rather than the aggregate preferences of the household. Around three-quarters of the on-the-go segment of the market is comprised of purchases made from vending machines, convenience stores, larger grocery stores when consumed immediately and kiosks; the other quarter is from restaurants and bars. We study the former, our data do not include the latter.

While our principal interest is in behavior in the on-the-go segment of the market, we also model behavior in the at-home segment, exploiting data on the at-home purchases of the households that individuals in our on-the-go sample belong to. We do this to take account of linkages between the at-home and on-the-go segments. In particular, when modeling supply side responses to the introduction of a tax we use information on both segments. A product that is available for purchase for on-the-go and at-home consumption has a market demand curve that depends on preferences in both segments. There is also the possibility of linkages between the segments in the demand side – for instance, recent at-home household purchases may influence on-the-go purchases. We test for this possibility using reduced form analysis, which

[4] Note, confusingly soda taxes are usually imposed on soft drinks products, of which sodas are a subset.

[5] CDC (2016) and National Diet and Nutrition Survey England (2018).

[6] This is in contrast to the at-home segment, which has been studied in Bonnet and Réquillart (2013) and Wang (2015).

shows there is very little evidence for such demand linkage. It suffices therefore to account for the supply linkage through the influence of on-the-go and at-home preferences on market demand curves and hence firm pricing. A second advantage of modeling behavior in the at-home segment, is it allows us to assess whether our conclusions regarding the individual targeting of soda taxes could plausibly be undone by off-setting at-home preferences. In Section 5.2 we show that this is unlikely to be the case.

2.1 Purchases

We use data from the Kantar Worldpanel and the associated Food On-The-Go Survey. These data are collected by the market research firm Kantar. The Worldpanel data cover the at-home segment of the market. They track the grocery purchases made and brought into the home by a sample of households that are representative of the British population. The Food On-The-Go Survey covers the on-the-go segment of the market. These data track food and drink purchases people make on-the-go for immediate consumption. Individuals in the On-The-Go Survey are randomly drawn from households in the Worldpanel.

Households in the Worldpanel data scan the barcode of all grocery purchases made and brought into the home. These include all food, drink, alcohol, toiletries, cleaning produce and pet foods. This means that we have comprehensive information on the total grocery baskets of the households to which the individuals in our on-the-go sample belong. The Kantar Worldpanel (and similar data collected in the US by AC Nielsen) have been used in a number of papers studying consumer grocery demand (see, for instance, Aguiar and Hurst (2007), Kaplan and Menzio (2015) and Dubois et al. (2014)).

To our knowledge the accompanying Food On-The-Go Survey is unique. Participating individuals record all purchases of snacks and non-alcoholic drinks for consumption outside the home (with the exception of those made in bars and restaurants) using mobile phones.

In both the Worldpanel and Food On-The-Go Survey we know what products (at the barcode, or UPC, level) were purchased and the transaction price. We also observe information on the store of purchase, household and individual attributes and product attributes.

Our data include information on the on-the-go behavior of 5,554 individuals and the at-home behavior of 4,204 households. The on-the-go individuals are drawn from the at-home households; there are fewer households because in some cases multiple individuals from the same household are present in the on-the-go data. To estimate demand we use information only on the individuals and households that report purchasing soft drinks.[7] Our estimation sample contains 2,374 individuals and 3,314 households.

We have data over the period June 2009-December 2014. In demand estimation we exploit the panel structure of the data to estimate consumer specific preferences. On average, in the on-the-go estimation sample, we observe consumers making non-alcoholic drinks purchases on 152 separate days; in the at-home estimation sample we, on average, observe households making non-alcoholic drinks purchases on 91 different weeks. In Table 2.1 we provide more details of the distribution of observations per consumer. In both the on-the-go and at-home samples over 85% of consumers are observed for more than 25 choice occasions, and for around half of consumers we observe 75 or more choice occasions. On 90% of choice occasions (days) in the on-the-go segment and 83% of choice occasions (weeks) in the at-home segment the consumer elects to purchase either one soft drink product or an alternative drink. For the remaining choice occasions, the consumer chooses multiple (typically two) soft drinks products. In this case we randomly select one purchase and use this in demand estimation.

Table 2.1: *Time series dimension of estimation sample*

Number of choice occasions observed	Individuals on-the-go		Households at-home	
	N	%	N	%
<25	292	12.3	475	14.3
25-49	553	23.3	761	23.0
50-74	347	14.6	541	16.3
75-99	214	9.0	406	12.3
100+	968	40.8	1131	34.1
Total	2374	100.0	3314	100.0

Notes: The table shows the number of choice occasions on which we observe individuals (on-the-go) and households (at-home) making purchase choices. An on-the-go choice occasions is a day in which the individual purchases a soft drink or alternative drink; an at-home choice occasion is a week in which the households purchases a soft drink or alternative drink.

[7] Strictly speaking, we use individuals/households that purchase at least 15 non-alcohol drinks and at least 10 soft drinks over the 5 and half years period. In the on-the-go segment these individuals account for around 95% of sugar from soft drink and from non-alcoholic drink purchases.

Table 2.2: *Products I*

Firm	Brand	Product	On-the-go %	On-the-go price	At-home %	At-home price
Soft drinks						
CocaCola			43.34		24.10	
	Coke		28.58		18.39	
		Coca Cola 330	4.32	0.64	0.19	0.55
		Coca Cola 500	8.19	1.13	0.82	1.01
		Coca Cola Diet 330	5.30	0.64	0.14	0.56
		Coca Cola Diet 500	10.78	1.14	1.29	1.02
		Coca Cola multi can			3.12	3.37
		Coca Cola Diet multi can			4.87	3.33
		Coca Cola bottle			3.46	1.42
		Coca Cola Diet bottle			4.36	1.36
		Coca Cola multi bottle			0.15	4.39
	Dr Pepper		3.48		1.97	
		Dr Pepper 330	0.46	0.63	0.01	0.50
		Dr Pepper 500	2.83	1.10	0.25	1.00
		Dr Pepper Diet 500	0.20	1.07		
		Dr Pepper multi can			0.33	2.18
		Dr Pepper Diet multi can			0.15	2.16
		Dr Pepper bottle			0.90	1.28
		Dr Pepper Diet bottle			0.34	1.24
	Fanta		4.19		2.35	
		Fanta 330	0.62	0.61		
		Fanta 500	3.25	1.11	0.29	1.00
		Fanta Diet 500	0.32	1.08		
		Fanta multi can			0.32	2.08
		Fanta Diet multi can			0.32	2.30
		Fanta bottle			1.09	1.23
		Fanta Diet bottle			0.33	1.26
	Cherry Coke		2.99		0.95	
		Cherry Coke 330	0.46	0.63	0.01	0.43
		Cherry Coke 500	1.65	1.10	0.15	1.00
		Cherry Coke Diet 500	0.88	1.06	0.10	1.00
		Cherry Coke multi can			0.16	2.71
		Cherry Coke Diet multi can			0.15	2.71
		Cherry Coke bottle			0.25	1.32
		Cherry Coke Diet bottle			0.13	1.31
	Oasis		4.09		0.44	
		Oasis 500	3.87	1.11	0.44	0.99
		Oasis Diet 500	0.22	1.07		
Pepsico			11.03		14.47	
	Pepsi		11.03		14.47	
		Pepsi 330	1.06	0.59	0.07	0.38
		Pepsi 500	2.24	0.99	0.25	0.75
		Pepsi Diet 330	1.79	0.60	0.13	0.40
		Pepsi Diet 500	5.95	0.97	0.74	0.76
		Pepsi multi can			1.19	2.04
		Pepsi Diet multi can			3.61	2.10
		Pepsi bottle			2.35	1.05
		Pepsi Diet bottle			6.12	1.06

Notes: Market shares are based on transactions. Prices are the mean across all choice occasions. The table describes the market shares of products purchased by 2,374 individuals in the on-the-go segment and 3,314 households in the at-home segment between June 2009 and December 2014.

Table 2.3: *Products II*

Firm	Brand	Product	On-the-go %	On-the-go price	At-home %	At-home price
Soft drinks *continued*						
GSK			7.77		4.20	
	Lucozade Energy		4.76		3.51	
		Lucozade Energy 380	2.57	0.94	0.22	0.75
		Lucozade Energy 500	2.18	1.16	0.35	1.01
		Lucozade Energy bottle			1.65	1.05
		Lucozade Energy multi bottle			1.29	2.79
	Ribena		3.01		0.69	
		Ribena 288	0.75	0.66	0.03	0.48
		Ribena 500	1.72	1.11	0.09	1.03
		Ribena Diet 500	0.54	1.08		
		Ribena multi			0.58	1.98
Other			19.14		22.03	
	Other		19.14		8.32	
		Other	16.32	1.10	2.07	1.17
		Other Diet	2.82	1.35	0.01	0.88
		Other big			3.16	1.07
		Other Diet big			1.49	1.11
		Other multi			1.08	2.12
		Other Diet multi			0.52	2.06
	Store		0.00		13.71	
		Store			5.17	0.45
		Store Diet			8.54	0.50
Alternative sugary drinks			8.25		23.16	
	Fruit juice		6.08		18.42	
		Fruit juice	6.08	1.07	2.78	1.60
		Fruit juice big			15.64	1.38
	Flavoured milk		1.41		3.96	
		Flavoured milk	1.41	0.98	2.78	0.72
		Flavoured milk big			1.17	1.05
	Fruit water		0.76		0.79	
		Fruit water	0.76	0.91	0.07	0.76
		Fruit water big			0.71	1.06
Outside			10.48		12.04	
	Bottled water		10.48		12.04	
		Bottled water	10.48	0.65	1.25	0.48
		Bottled water big			10.79	0.92

Notes: Market shares are based on transactions. Prices are the mean across all choice occasions. The table describes the market shares of products purchased by 2,374 individuals in the on-the-go segment and 3,314 households in the at-home segment between June 2009 and December 2014.

2.2 Brands, products and stores

In Tables 2.2 and 2.3 we describe the products available in the non-alcoholic drinks market, both in the on-the-go and at-home segments. Products classified as "soft drinks" are available in a number of large brands, owned by CocaCola, Pepsico and

GlasoSmithKline (GSK). There are a large number of small brands (with market shares below 2%). We aggregate these small brands into a composite "Other" brand. We also aggregate generic supermarket products into a composite "Store" brand (which is only available in the at-home segment). We additionally include a composite fruit juice, flavored milk, fruit (or flavored) water and bottled water brand. These together account for about 20% of on-the-go transactions and 25% of at-home transactions. Our counterfactual involves simulating the introduction of a "soda tax". This applies to the set of sugar sweetened soft drinks.[8] Diet varieties, as well as fruit juice, flavored milk and bottled water are alternative (non-taxed) goods, which consumers may choose to substitute towards. Each brand is available in a number of different sizes and container types.

Tables 2.2 and 2.3 make clear that in the on-the-go segment, individuals choose between single portion products (e.g. Coca Cola 330 refers to a 330ml, or 11oz, can and Coca Cola 500 refers to a 500ml, or 17oz bottle). These single portions are also available in the at-home segment, though they are significantly less popular than larger pack sizes. In the case of the brand Coke, large sizes available in the at-home segment alone include a multi-pack of cans, a large plastic bottle and multi-pack of bottles.[9]

For each transaction we observe in what type of store the consumer made its purchase. This means, in demand estimation, we can use retailer specific prices (Section 3.2 discusses how we exploit cross retailer price variation) and retailer specific choice sets (see Manski (1977), Goeree (2008) and Crawford et al. (2017) on how failing to account for heterogeneity in choice sets can lead to inconsistent demand estimates). These choice sets are based on all of the products we observe being purchased in a retailer type-year. We assume that in the at-home segment households that purchase single/multi portion options choose between the set of single/multi portion options available. Table 2.4 describes both the number and share of purchases of non-alcoholic drinks in both on-the-go and at-home samples that we observe across retailer types. In the on-the-go segment the largest share of purchases are made in branches of small national chains or independent stores (corner shops). The large national supermarket chains account together for around one-fifth of purchases, and vending machines account for around 8%. In the at-home segment most purchases are made at the large national supermarket chains (88%). These large supermarkets comprise the four large retailers that dominate the UK

[8] In the Appendix we show simulations of a broader tax that also applies to diet soft drinks.

[9] These multi portion products embed some aggregation. For instance, the product Coca Cole bottle comes in both a 1.25l and 2l variant, though typically only one of these two sizes is available in any given store.

grocery market – Asda, Morrisons, Sainsbury's and Tesco – as well as "Discounters", which aggregates together the low price Aldi and Lidl retailers. In demand we use national prices for the retailers that set national prices, and regional average prices for the set of small national and independent stores.

Table 2.4: *Retailer types in which non-alcoholic drinks purchases are observed*

	On-the-go		At-home	
Store	N	%	N	%
Small national or independent	258494	71.4	35697	11.8
Vending machines	28659	7.9		
Large national	74710	20.6	266686	88.2
Asda	10617	14.2	58475	21.9
Morrisons	7605	10.2	40678	15.3
Sainsbury's	15588	20.9	39395	14.8
Tesco	40393	54.1	110619	41.5
Discounters	507	0.7	17519	6.6
Total	361863	100.0	302383	100.0

Notes: The table describes the number and share of purchases made by 2,374 individuals in the on-the-go estimation sample and 3,314 households in the at-home estimation sample in each type of store type between June 2009 and December 2014.

Linkages between market segments

Product overlap across segments means that supply side pricing, even of products mainly purchased on-the-go, will depend on both on-the-go and at-home preferences. For instance, the market demand curve for the product Coca Cola 330 will depend on preferences over the product in the on-the-go and at-home segments. As it is the market demand curve that is relevant for firms' pricing decisions we refer to this as a supply-side link. We take account of this in counterfactual simulations.

Another possible link between segments would exist if individuals' on-the-go decisions are influenced by at-home purchases made by the households to which the individuals belong. We explore this possibility by checking for correlations in on-the-go purchases with recent at-home purchases. In particular, we create a data set that, for each of the 2,374 individuals in the on-the-go estimation sample, has one observation for every day (regardless of whether a non-alcoholic drink is purchased or not) between the individual's first and final day in the sample. We regress a dummy for whether the individual purchased a non-alcoholic drink on a particular day on dummy variables for whether the individual's household purchased non-alcoholic drinks in the at-home segment in each of the 4 preceding 7 day periods (column (1) of Table 2.5) and on each of the preceding 7 days (column (3) of Table

2.5). We also regress the volume of non-alcoholic drinks an individual buys on volume bought in the at-home segment in each of the 4 preceding 7 day periods (column (2) of Table 2.5) and on each of the preceding 7 days (column (4) of Table 2.5). In each case we include individual and year-month fixed effects.

The coefficient estimates in Table 2.5 indicate very little evidence of dependence between current on-the-go purchases and recent past at-home purchases. A number of the coefficients are statistically significant, however, this is driven by the very large sample size. The magnitude of the effects is very small. For instance, the average effect of purchases in the at-home segment in the past 4 weeks is associated with a *raised* probability of buying on-the-go of 0.005 (relative to a mean of 0.14) and *raised* volume purchased of 2ml (relative to a mean of 80ml). As well as being very small, the direction of these effects are opposite to what we would expect if consumers viewed on-the-go and at-home consumption as substitutes.

Our conclusion from this is that, once individual heterogeneity is accounted for, there is little evidence that demand linkages are of first order importance in the current context. While it would be interesting to study more broadly the interactions between household grocery demand and individual on-the-go grocery demand, we leave this for future work.

Switching across product types

An advantage of the long T dimension of our data is that we are able to distinguish between consumers that, when purchasing a non-alcoholic drink, only ever purchase soft drinks, from those that also sometimes choose alternative drinks (i.e. fruit juice, flavored milk or water). Similarly, we can distinguish between consumers who, when purchasing a non-water drink, either always choose sugary products, always choose diet products, or sometimes choose sugary and sometimes diet products. This enables us to identify consumers who effectively have infinite preferences for some product attributes (e.g. a consumer who only chooses diet products, in effect, has an infinitely negative preference for sugar) – see Section 3.1. In Table 2.6, for the on-the-go segment, we show what fraction of the estimation sample falls into each group.Consumers who at different points are observed buying both sugar sweetened soft drinks, diet varieties, and alternative sugary drinks account for 76.1% of the sample.

Table 2.5: *Relationship between purchases on-the-go and at-home*

	(1) Purchases	(2) Volume	(3) Purchases	(4) Volume
Purchased at home in last week	0.0055***	0.0002***		
	(0.0006)	(0.0001)		
Purchased at home 2 weeks ago	0.0044***	0.0002***		
	(0.0005)	(0.0001)		
Purchased at home 3 weeks ago	0.0029***	0.0002***		
	(0.0005)	(0.0001)		
Purchased at home 4 weeks ago	0.0025***	0.0001*		
	(0.0005)	(0.0001)		
Purchased at home yesterday			-0.0018	-0.0002
			(0.0013)	(0.0002)
Purchased at home 2 days ago			-0.0007	-0.0001
			(0.0012)	(0.0002)
Purchased at home 3 days ago			0.0035**	0.0001
			(0.0011)	(0.0001)
Purchased at home 4 days ago			0.0073***	0.0005**
			(0.0011)	(0.0002)
Purchased at home 5 days ago			0.0077***	0.0005***
			(0.0011)	(0.0001)
Purchased at home 6 days ago			0.0083***	0.0006***
			(0.0011)	(0.0002)
Purchased at home 7 days ago			0.0082***	0.0005**
			(0.0013)	(0.0002)
Constant	0.1402***	0.0818***	0.1435***	0.0827***
	(0.0044)	(0.0029)	(0.0044)	(0.0029)
N	3,420,627	3,420,627	3,488,632	3,488,632
Year-month effects	yes	yes	yes	yes
Individual effects	yes	yes	yes	yes

Notes: Column (1) reports coefficients of regression of a dummy for on-the-go purchase on dummies for at-home purchases in each of past 4 weeks. Column (2) reports coefficients of regression of volume of on-the-go purchase on volume of at-home purchases in each of past 4 weeks. Columns (3) and (4) repeat the analysis in (1) and (2) instead focusing on the effect of at-home purchases in each of the previous 7 days. There is an observation for every day between the first and final day in the sample for each of the 2,374 individuals in the on-the-go estimation sample.

Table 2.6: *Consumer specific product sets*

	Ever purchase fruit juice, flavored milk or bottled water?		
	Yes	No	Total
Individual chooses:			
Sugary and diet varieties	76.1	8.3	84.3
Only sugary varieties	10.4	3.1	13.5
Only diet varieties	1.3	0.8	2.2
Total	87.8	12.2	100.0

Notes: For each of the 2,374 on-the-go purchasers, across all their choice occasions, we distinguish between those that only buy soft drinks or buy soft drinks and other drinks and that only buy diet, sugary or both diet and sugary varieties. Numbers are % of the on-the-go estimation sample in each cell.

2.3 Demographics

The main justification for the adoption of soda taxes is that some consumers' intake of these products creates costs that are borne by others (externalities) or future costs they themselves will bear that they do not fully account for at the point of consumption (internalities). The more a soda tax can alter the consumption of those that create large externalities or internalities, the better targeted it will be. Our aim in this paper is to provide evidence on how well targeted soda taxes are. To do this we focus mainly on individuals' behavior in the on-the-go segment, using a demand framework that allows us to relate individual level consumption responses to measures of the likely harm associated with sugary soft drinks consumption. We focus on three measures of the likely magnitude of harm associated with sugary soft drinks consumption – age, total annual dietary sugar and income. The reason for this is twofold. First, policymakers have justified the introduction of soda taxes as a way of targeting excess consumption of sugar, and, in particular, excess sugar consumption among the young, while critics have expressed concern about the negative distributional effects of the policy. Second, there is evidence that social costs (externalities and internalities) associated with sugary soft drinks vary in important ways across these measures, although there is little evidence on the exact mapping between consumption and social costs.

Our focus on how the effects of the tax vary by age is motivated by a number factors: (i) young people are the stated target of policy (for instance, see CDC (2016), Public Health England (2015)); (ii) on average, the young get a relatively large fraction of their calories from sugar – in other words excess sugar consumption is more severe among this group (see details in Appendix A.1); (iii) there is a substantial literature that emphasizes that excess sugar consumption has negative consequences for children (e.g. Gortmaker et al. (2009) and Han and Powell (2013)) and in the long run these can be profound (Currie (2009) and Currie et al. (2010)); and (iv) it is likely that young people are less likely to take account of the long term consequences of poor dietary choices (for instance, Ameriks et al. (2007) show that the young suffer more from self-control problems than older people).

We also focus on the effects of the tax by total annual dietary sugar. The reason for this is due to the possibility of some convexity in how social costs arise from sugar consumption (e.g. at lower levels of sugar consumption the probability of developing type II diabetes is trivially small, but this probability may rise nonlinearly in sugar consumption). There is suggestive evidence of such effects arising from alcohol consumption – for example, there is evidence of a threshold effect with some diseases: at low levels of ethanol consumption, the risk of disease is not elevated, but this risk increases sharply above a certain point (see Lönnroth et al. (2008) for evidence on tuberculosis, and Rehm et al. (2010) for evidence on liver cirrhosis).

Our focus on how responses vary with a measure of income is motivated, in part, due to concerns that soda taxes are likely to be regressive. However, in addition to this, there is evidence that income might be (causally) related to excess consumption – for instance, Haushofer and Fehr (2014) and Mani et al. (2013) suggest that the stress and cognitive load of being in poverty leads people to be more likely to make unwise decisions and underweight the future. To the extent that this is true, if a soda tax achieves large reductions in consumption among poor consumers, this could be to the policy's advantage. As Allcott et al. (2018a) point out, this complicates assessment of whether the policy is regressive.

We observe individuals' age in our data. To construct a measure of total annual dietary sugar and income we use the Kantar Worldpanel (which records all grocery purchases made by the household each individual belongs to). We measure total annual dietary sugar as the share of total household calories that are from added sugar. As our income measure we construct each household's total annual equivalized grocery expenditure; we equivalize using the standard OECD modified equivalence scale. In Appendix A.4 we show that equivalized grocery expenditure is strongly correlated with current income, while expenditure is often viewed as a better proxy for lifetime income than current income is (e.g. Poterba (1989)).

In Tables 2.7 we show what fraction of the total on-the-go sample falls into 6 age categories. We also show, within each group, what fraction we observe purchasing soft drinks (we refer to these individuals as "soft drinks purchasers" in the table). The estimation sample comprises the group of individuals who are "soft drinks purchasers". The remaining rows of the table summarize various aspects of their purchase behavior. The table shows that young consumers (relative to older ones) are (i) more likely to be soft drink purchasers, (ii) conditional on being so, obtain more sugar from these products, and (iii) that this higher level of sugar is driven by purchasing more often and being more likely to choose sugary varieties (and is not driven by being more likely to buy the largest single portion size – the 500ml bottles).

Tables 2.8 and 2.9 show the same statistics for deciles of the distribution of total annual dietary sugar and total annual equivalized grocery expenditure. Those with more sugar in their total diet are both more likely to be soft drinks purchasers and, conditional on this, to get large quantities of sugar from these products. A similar pattern holds across the total annual equivalized grocery expenditure distribution;

those with lower total annual grocery expenditure are more likely to be soft drinks purchases and obtain a relatively high amount of sugar from these products.

Table 2.7: *Descriptive statistics by age groups*

	Age group					
	13-21	22-30	31-40	41-50	51-60	60+
% of sample	37.1	17.5	11.5	12	10.3	11.6
Fraction of soft drink purchasers	.42	.49	.52	.47	.37	.25
Conditional on purchase:						
Mean sugar from soft drinks per year (g)	2076	1777	1368	1402	1293	1090
Mean number of purchases per year	48.3	44	39.7	35.3	34.7	30.5
Fraction of sugary products	.76	.71	.66	.65	.67	.67
Fraction of 500ml bottles	.76	.73	.71	.71	.71	.69

Notes: Row 1 shows the fraction of individual-year observations in each age group. Row 2 shows the fraction of each age group that is ever observed purchasing soft drinks. The remaining rows show means for the set of soft drinks purchasers of; total annual sugar from these products, number of annual purchases, fraction of purchases for sugary rather than diet varieties, and fraction of purchases for the larger 500ml bottle size.

2.4 Weather data

Demand for soft drinks versus alternatives may fluctuate with weather conditions. To account for this in our demand model we use the mean temperature by day from the Met Office Historic station data. These data are reported monthly for 35 locations in the UK.[10]

[10]The data are available at https://www.metoffice.gov.uk/public/weather/climate-historic/#?tab=climateHistoric

Table 2.8: *Descriptive statistics by total annual dietary sugar*

	Decile of distribution of share of calories from added sugar									
	1	2	3	4	5	6	7	8	9	10
% of sample	8.5	10	11	11.9	12.8	13.8	14.9	16.2	18.3	24.7
Fraction of soft drink purchasers	.33	.4	.4	.4	.44	.44	.45	.47	.46	.47
Conditional on purchase:										
Mean sugar from soft drinks per year (g)	1070	1338	1247	1273	1292	1372	1560	1602	2009	1820
Mean number of purchases per year	38	39.2	37.3	35.6	36.3	37.5	38.6	39.3	44	39.8
Fraction of sugary products	.59	.63	.64	.66	.68	.71	.7	.7	.7	.74
Fraction of 500ml bottles	.67	.7	.69	.72	.73	.75	.73	.72	.72	.74

Notes: Row 1 shows the upper bound of the decile of total annual dietary sugar. Row 2 shows the fraction of each decile that is ever observed purchasing soft drinks. The remaining rows show means for the set of soft drinks purchasers of; total annual sugar from these products, number of annual purchases, fraction of purchases for sugary rather than diet varieties, and fraction of purchases for the larger 500ml bottle size.

Table 2.9: *Descriptive statistics by total annual equivalized grocery expenditure*

	Decile of distribution of total equivalized grocery expenditure									
	1	2	3	4	5	6	7	8	9	10
% of sample	.8	1.1	1.3	1.5	1.7	1.9	2.2	2.5	3.1	5.1
Fraction of soft drink purchasers	.47	.44	.43	.43	.46	.42	.39	.42	.42	.38
Conditional on purchase:										
Mean sugar from soft drinks per year (g)	1648	1494	1647	1447	1671	1477	1331	1326	1251	1461
Mean number of purchases per year	42.2	38.5	42.3	37.7	40	37.3	35.4	36.5	35.7	40
Fraction of sugary products	.7	.68	.66	.68	.7	.68	.71	.66	.66	.67
Fraction of 500ml bottles	.68	.7	.7	.71	.73	.72	.75	.73	.74	.73

Notes: Row 1 gives the upper bound of the decile, measured in £1000, of total annual equivalized grocery expenditure. Row 2 shows the fraction of each decile that is ever observed purchasing soft drinks.The remaining rows show means for the set of soft drinks purchasers of; total annual sugar from these products, number of annual purchases, fraction of purchases for sugary rather than diet varieties, and fraction of purchases for the larger 500ml bottle size.

3 Model

In this section we develop a model of consumer demand and firm pricing in the non-alcoholic drinks markets. What distinguishes our approach from previous work is (i) we focus on modeling the preferences of *individuals* using information on their purchases on-the-go, and (ii) we exploit the long panel nature of our data to estimate individual specific preference parameters, giving us the ability to relate individual specific preferences and counterfactual effects to any demographic or behavior measure of the individual.

While our focus is on on-the-go behavior we also estimate household level demand in the at-home segment. We therefore write down our demand model in a way that is consistent with estimation in either segment. As we argued in Section 2.2, the main linkage between the segments is their common effects on market level demand and hence firm pricing. We aggregate demand estimates in the two segments into market demand curves. At the consumer level we assume demand in the two segments is independent (an assumption supported by the descriptive analysis presented in Section 2.2). In Section 5.2 we discuss the implications of our estimates in the at-home segment for the analysis of how well targeted soda taxes are; we show accounting for what households bring into the home is unlikely to substantially alter the conclusions of our analysis.

We begin by considering consumer behavior in the non-alcoholic drinks market. This means that our model, in response to a simulation of the introduction of a soda tax, allows for the possibility of consumer switching to diet alternatives, or alternative drinks products. Ex ante, such switching to alternative drinks seems likely to be much more important than substitution towards foods, and there is some limited medical evidence that calories from liquids do not displace those from solids (see, for instance, DiMeglio and Mattes (2000)). However, to consider the possibility that consumers respond to the tax by switching from drinks to foods that contain sugar, we nest our model of demand for drinks within a two-stage demand model, in which, in a first stage, consumers choose between food and drinks. We present this in Section 5.1.

3.1 Demand model

We index consumers by $i \in \{1, \ldots, N\}$. In the on-the-go segment consumers are individuals; in the at-home segment they are households. Notationally, we distinguish between such consumers by indicating individuals as $i \in \mathcal{M}^{out}$ and households as $i \in \mathcal{M}^{in}$ (where $\mathcal{M}^{in} \cup \mathcal{M}^{out} = \{1, \ldots, N\}$). We observe each consumer on many choice occasions, indexed by $\tau = \{1, \ldots, \mathcal{T}\}$. A choice occasion τ refers to a consumer visiting a retailer r_τ at time t_τ and purchasing a drink.

As outlined in Section 2.2, the choice sets facing consumers depend both on the retailer they shop with and whether they are shopping for single portion products (as in the on-the-go segment) or multi portion products (as is most commonly the case in the at-home segment). We denote the available set of products in retailer $r = \{1, \ldots, R\}$ and on single or multi portion choice occasion (denoted by $o = \{0, 1\}$) as Ω_{ro}. For instance, an on-the-go individual that visits Tesco will choose between all the single portion products available in that retailer.

We index the "inside" products (i.e. soft drinks) by $j = \{1, \ldots, j'\}$ and the alternative juice options by $j = \{j'+1, \ldots, J\}$. $j = 0$ denotes the option of selecting bottled water. These products are displayed in Tables 2.2 and 2.3. The choice set facing a consumer will contain a subset of the $J+1$ products. Each product belongs to a brand – we denote the brand that product j belongs to as $b(j)$. Products within a brand differ based on whether they are a sugary or diet variety and in their pack size.

For any product j, we assume the pay-off associated with selecting the product on choice occasion τ takes the form:

$$U_{ij\tau} = \alpha_i p_{jr_\tau t_\tau} + \beta_i s_j + \gamma_i w_j + \delta^z_{d(i)} z_j + \delta^h_{d(i)} h_{c(i)t_\tau} + \xi_{d(i)b(j)t_\tau} + \zeta_{d(i)b(j)r_\tau} + \epsilon_{ij\tau}, \quad (3.1)$$

where $\epsilon_{ij\tau}$ is an idiosyncratic shock distributed type I extreme value. $p_{jr_\tau t_\tau}$ denotes the price of product j, which varies over time (t) and cross-sectionally across retailers (indexed by r).[11] s_j is a dummy variable indicating whether the product is a sugary or diet variety and w_j is a dummy variable for whether the product is an inside product (i.e. a soft drink). We allow the preference parameters on these product attributes (α_i, β_i and γ_i) to be consumer specific.

We also include size-carton type effects (z_j), weather temperature effects in the county $c(i)$ the consumer lives at time t $(h_{c(i)t})$[12], time-varying brand effects $(\xi_{d(i)b(j)t})$ and retailer-brand effects $(\zeta_{d(i)b(j)r})$. In each case we allow the influence of these attributes to vary by demographic group – we denote these by $d \in \{1, \ldots, D\}$ and let $d(i)$ denote the group consumer i belongs to. For the on-the-go segment these

[11] Specifically, prices vary across Asda, Morrisons, Sainsbury's, Tesco, Discounters, other national supermarkets, other national convenience stores, vending machines, and independent stores in the north, midlands and south.

[12] Note, this effect is for inside (i.e. soft drinks) options only and captures how temperature affects the choice of soft drinks versus alternative drinks products.

groups are based on individual sex and age, for the at-home segment on whether the household contains children and the skill level of main shopper's occupation.[13]

We denote by $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)'$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_N)'$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)'$ the vectors of individual preference parameters. These individual level preferences enable our model to capture within individual correlation in choices across choice occasions. We do not place any a priori restriction on the joint distribution of these variables. We use the large $\mathcal{T}$ dimension of our data to recover estimates of individual specific parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$, while the large N dimension allows us to identify nonparametrically the joint probability distribution function $f(\alpha_i, \beta_i, \gamma_i)$ using the empirical probability distribution function of estimated $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$. We can also construct the distribution of preferences conditional on observable consumer characteristics, X; $f(\alpha_i, \beta_i, \gamma_i|X)$. These observable characteristics can be demographic variables or measures of the overall diet or grocery purchasing behavior of the household to which the individual belongs.

A number of papers (see, for instance, Berry et al. (1995), Nevo (2001) and Berry et al. (2004)) show that incorporating consumer level preference heterogeneity is important for enabling choice models to capture switching patterns across products,[14] while a few papers have used non-parametric methods to relax parametric restrictions on random coefficients.[15] Like these papers we model consumer specific preferences, however, in contrast to them, we treat the preferences as parameters to be estimated and thereby avoid having to make independence assumptions to integrate out the density. This allows us to flexibly relate the preference parameters and individual specific effects of simulations to observable attributes of consumers. Unlike in a random coefficient approach we do not need to a priori specify how the preference distribution depends on exogenous attributes of consumers, and we can relate individual specific effects to attributes of consumers (like other aspects of their grocery purchasing behavior).

[13]Specifically, in the on-the-go segments we let these preferences vary over four groups based on individual sex and whether they are aged below 40 or not. In the at-home segment we allow them to vary over five groups based on whether the household has no children, is a pensioner household, or contains children, and for non pensioner households, whether the main shopper's occupation is classified as high or low skilled.

[14]Lewbel and Pendakur (2017) show similar results apply in non-linear continuous choice models, with the incorporation of random coefficients resulting in their model much more effectively capturing the distributional impacts of taxation.

[15]Burda et al. (2008) exploit Bayesian Markov Chain Monte Carlo techniques and Train (2008) uses an expectation-maximization algorithm to estimate the random coefficient distribution. Train (2008) applies the method either with a discrete random coefficient distribution or with mixtures of normals. Bajari et al. (2007) discretize the random coefficient distribution and use linear estimation techniques to estimate the frequency of consumers at each fixed point.

One potential concern is that our estimates may be subject to an incidental parameter problem that is common in non-linear panel data estimation. Even if both $N \to \infty$ and $\mathcal{T} \to \infty$, asymptotic bias may remain, although it shrinks as the sample size rises (Hahn and Newey (2004), Arellano and Hahn (2007)). The long $\mathcal{T}$ dimension of our data helps lower the chance that the incidental parameter problem leads to large biases. We implement the split sample jackknife procedure suggested in Dhaene and Jochmans (2015) and in Section 5.3 show that our maximum likelihood and jackknife estimates are similar and that the bias correction does not affect our results.

Another benefit of having large $\mathcal{T}$ for each individual is that we can allow for consumers who may have sufficiently strong distaste for some product sets that they endogenously will never choose to buy them. In contrast to standard logit discrete choice models, we allow some consumers to have zero probability of purchasing certain products by using the long time dimension of our data to identify consumers that never purchase products with particular characteristics.

In particular, we identify consumers that only ever purchase inside products (i.e. soft drinks; those with $w_j = 1$), and never purchase alternative drinks (i.e. fruit juice, flavored milk or bottled water; those with $w_j = 0$). Such consumers can be thought of as having negative infinite preferences for non-inside products (which we denote by $\gamma_i = \infty$; an infinite preference for inside products). Consumers that sometimes purchase inside products and other times purchase alternative drinks have $\gamma_i \in (-\infty, \infty)$.

Similarly, we distinguish between consumers that, when buying either an inside or alternative juice product only ever select sugary options (i.e. those for which $s_j = 1$), those that only ever purchase non-sugary option (i.e. diet soft drinks or bottled water; for which $s_j = 0$), and those that we observe sometimes purchasing sugary products and at other times non-sugary products. The three groups, respectively, have sugar preferences given by $\beta_i = \infty$, $\beta_i = -\infty$, and $\beta_i \in (-\infty, \infty)$.

To specify the set of products that consumers have non-zero probabilities for, it is useful to define the product sets Ω_{ws}, Ω_{wn}, Ω_{as} and Ω_{an}, which denote respectively the sets of sugar sweetened soft drinks, diet soft drinks, alternative sugar drinks and water. We can then define consumer i specific sets of products with non-zero purchase probabilities, denoted by Ω_i, as

$$\Omega_i = \begin{cases} \Omega_{ws} \cup \Omega_{wn} \cup \Omega_{as} \cup \Omega_{an} & \text{if} \quad \beta_i \in (-\infty, \infty) \text{ and } \gamma_i \in (-\infty, \infty) \\ \Omega_{wn} \cup \Omega_{an} & \text{if} \quad \beta_i = -\infty \text{ and } \gamma_i \in (-\infty, \infty) \\ \Omega_{ws} \cup \Omega_{as} \cup \Omega_{an} & \text{if} \quad \beta_i = +\infty \text{ and } \gamma_i \in (-\infty, \infty) \\ \Omega_{ws} \cup \Omega_{wn} & \text{if} \quad \beta_i \in (-\infty, \infty) \text{ and } \gamma_i = \infty \\ \Omega_{wn} & \text{if} \quad \beta_i = -\infty \text{ and } \gamma_i = \infty \\ \Omega_{ws} & \text{if} \quad \beta_i = +\infty \text{ and } \gamma_i = \infty. \end{cases}$$

The share of consumers in each group is given in Table 2.6. We assume that the consumer level products sets Ω_i are measured exactly due to the large $\mathcal{T}$ dimension of observed consumer level choices. However, our sample is finite and thus a finite sample measurement error is introduced on Ω_i. We ignore this measurement error; Monte Carlo simulations show that such error is negligible in our application where $\mathcal{T}$ is relatively large.[16]

We define:

$$\begin{aligned} v_{ijr_\tau t_\tau} \equiv & \alpha_i p_{jrt_\tau} + \beta_i s_j 1_{\{\beta_i \in (-\infty,\infty)\}} + \gamma_i w_j 1_{\{\gamma_i \in (-\infty,\infty)\}} \\ \eta_{ijr_\tau t_\tau} \equiv & \delta^z_{d(i)} z_j + \delta^h_{d(i)} h_{c(i)t_\tau} + \xi_{d(i)b(j)t_\tau} + \zeta_{d(i)b(j)r_\tau} \end{aligned}$$

such that equation (3.1) can be written

$$U_{ij\tau} = v_{ijr_\tau t_\tau} + \eta_{ijr_\tau t_\tau} + \epsilon_{ij\tau}.$$

The assumption that $\epsilon_{ij\tau}$ is an idiosyncratic shock distributed type I extreme value means that the consumer level choice probabilities are given by the multinomial logit formula, such that the choice probability of consumer i on choice occasion τ purchasing any good $j \in \Omega_{r_\tau}$ can be written[17]

$$P_{i\tau}(j) = \frac{1_{\{j \in \Omega_i\}} \exp(v_{ijr_\tau t_\tau} + \eta_{ijr_\tau t_\tau})}{\sum_{k \in \Omega_i \cap \Omega_{r_\tau}} \exp(v_{ikr_\tau t_\tau} + \eta_{ikr_\tau t_\tau})} \tag{3.2}$$

If we denote consumer i's sequence of choices across all choice occasions as $\boldsymbol{y}_i = (y_{ir_1t_1}, ..., y_{ir_\mathcal{T}t_\mathcal{T}})$. The probability of observing $\boldsymbol{y}_i$ is given by:

$$\mathcal{L}_i(\boldsymbol{y}_i) = \prod_{\tau=1}^{\mathcal{T}} P_{i\tau}(y_{ir_\tau t_\tau})$$

[16]Further details available from authors on request.

[17]Of course the probability that consumer i at occasion τ purchases a good $j \notin \Omega_{r_\tau}$ is zero.

and, denoting the demographic specific preference parameters $\boldsymbol{\eta}$, the associated log-likelihood function is:

$$l(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\eta}) = \sum_i \ln \mathcal{L}_i(\boldsymbol{y}_i), \tag{3.3}$$

which is globally concave with respect to all parameters.

3.2 Identification

Our main identification challenge is to pin down the causal impact of price on demand. Our strategy for doing this relies on two sources of price variation. First, conditional on brand-time and retailer-drink type effects, we exploit cross-retailer price variation. We observe individuals making purchases in different retailers (and thereby facing different price vectors). We assume the retailer choice is not driven by shocks to demand for specific drinks products, but rather is driven by daily life in which individuals move between home, school, leisure or work. Second, we exploit variation in prices within brand across different containers and sizes. While there may be some aggregate shock to demand for a specific brand (that manufacturers observe and change prices as a consequence of), we assume that there are not aggregate shocks within brand for different container types. We discuss each source of variation in turn. In Appendix A.3 we provide descriptive statistics that suggest that individuals face price variation, and that average prices across transactions reflect actual variation in underlying prices.

The price vector an individual faces at the point of purchase depends on which retailer they visited. These retailers include a set of large national retailers that price nationally, smaller retailers with regionally varying prices and vending machines (see Table 2.4). We include demographic group specific time varying brand effects $\xi_{d(i)b(j)t}$ and retailer effects, interacted with the set of soft drinks, the alternative sugary drinks and the bottled water, $\xi_{d(i)b(j)r}$. The former capture aggregate (demographic specific) fluctuations in brand demand over time (e.g. driven by national advertising) and the latter capture any differential propensity of consumers to choose different drink types across retailers. Conditional on these, the cross-retailer differences in prices provides a useful source of price variation.

There are two main concerns with exploiting this type of price variation. First, an issue would arise if individual level demand shocks to specific soft drinks products drive store choice for the on-the-go segment; for instance, if a consumer that has a demand shock that leads them to want Coca Cola visits a retailer that happens to temporarily have a low price for that product, and, if instead they had a demand shock that led them to want Pepsi they would have selected a retailer with a relatively low Pepsi price. Such behavior would occur either if consumers could predict fluctuating relative prices across retailers or if they visited several retailers in search of a low price draw for the product they are seeking. We find either scenario highly unlikely in the case of soft drinks, which make up only a very small fraction of total grocery spending.

Second, an issue would arise if differential changes in the prices of different soft drinks across retailers are driven by retailer-time varying demand shocks for soft drinks. In the UK the vast majority of soft drinks advertising is done nationally and by the manufacturer. There is little retailer or regional advertising of specific drinks products. In the at-home segment stores are predominately national and set national prices. For the small stores price variation that make up the majority of on-the-go transactions, differential within drink type price variation will be driven by local discounts related to excess stock.

The second source of price variation we exploit is non-linear pricing across container sizes. This price variation is not collinear with the size effects and the extent of non-linear pricing varies over time and retailers. This source of identifying variation would be invalid if there were systematic shocks to consumers' valuation of container sizes that were differential across brand after conditioning on time varying brand effects and container size and type effects. It seems more plausible that such tilting of brand price schedules is driven by cost variations that are not proportional to pack size, differential pass-through of cost shocks and differences in how brand advertising affects demands for different pack sizes. This identification argument is similar to that in Bajari and Benkard (2005). In an application to the computer market, they assume that, conditional on observables, unobserved product characteristics are the same for products that belong to the same model. We assume that, conditional on time-varying brand characteristics, unobserved size specific attributes do not vary differentially across brands.

The main source of variation in the sugar content of products is between sugary and diet varieties (with most brands being available in each variety). We identify consumer specific preferences parameter for sugary versus diet products (rather than a preference for a marginal increase in sugar quantity). We assume that the brand effects are common across sugary and diet varieties of the same brand, and that the taste for the sugary variety is additively separable. This means that, for example, we do not allow the individual sugary taste to be different for Coke versus Pepsi. Table 2.6 shows that there are many individuals who purchase both sugary and diet varieties.

3.3 Pass-through of a tax on sugary soft drinks

We consider the impact of a tax levied on sugary soft drinks. Such taxes have recently been introduced in a number of locations. These policies are often referred to as "soda taxes", though they typically apply to broader set of products than carbonates (or sodas). These taxes are typically volumetric (i.e. levied per liter or ounce) and levied either on soft drinks that contain sugar, or on all soft drinks (including diet varieties). We consider a tax levied on sugary soft drinks and in Appendix C show results for a tax on all soft drinks. A number of US cities have recently legislated for the introduction "soda taxes"[18], the UK introduced a tax on sugary soft drinks in 2018 and France and Mexico have had taxes in place since 2012 and 2014. We model a tax of 25 pence per liter, (or 33 US cents per liter, which is 1.2 cents per ounce – similar to the US taxes of 1-1.5 cents per ounce).

The degree of pass-through of the tax to consumer prices will depend on the nature of competition in the market. We model tax pass-through by assuming that drinks manufacturers compete by simultaneously setting prices in a Nash-Bertrand game. We consider a mature market with a stable set of products, and we therefore abstract from entry and exit of firms and products from the market. We use our demand estimates and an equilibrium pricing condition to infer firms' marginal costs (see Berry (1994) or Nevo (2001)) in order to then simulate the effect of a tax on consumer prices.

Let $f = \{1, \ldots, F\}$ index manufacturers and F_f denote the set of products owned by firm f. We assume that prices are set by manufacturers and abstract from modeling manufacturer-retailer relationships. Such an outcome would be achieved by efficient vertical contracting (Villas-boas (2007), Bonnet and Dubois (2010)).[19] Bonnet and Dubois (2010) show that in the French grocery market, price equilibria correspond to the case where manufacturers and retailers do use non-linear contracts in the form of two part tariffs. Testing for deviations from efficient contracting in UK manufacturer-retailer relations is an interesting question that we leave for future research.

We index markets by m. Markets are both temporal and, given the efficient contracting assumption, vary across retailer type. We denote the size of the on-the-go segment in market m by M_m^{out} and the size of the at-home segment by M_m^{in}.

[18] A tax of 1.5 cent per ounce on regular and diet soft drinks is effective in Philadelphia as of January 2017; Berkeley, San Francisco, Oakland, Albany California and Boulder Colorado all legislated for sugary soft drinks taxes of 1 cent per ounce (2 cents in Albany) implemented in 2017-18; a sugary soft drinks tax of 1 cent per ounce was effective in Cook County, Illinois (which includes Chicago) as of June 2017, but was soon repealed thereafter.

[19] Efficient contracting entails non-linear contracts and side transfers between manufacturers and retailers to reallocate profits, and avoids the double marginalization problem.

Aggregating across consumer level purchase probabilities we obtain the market level demand function for product j:

$$q_{jm}(\mathbf{p}_m) = \underbrace{M_m^{out} \sum_{\tau/t_\tau=m, i\in\mathcal{M}_m^{out}} P_{i\tau}(j)}_{\equiv q_{jm}^{out}(\mathbf{p}_m)} + \underbrace{M_m^{in} \sum_{\tau/t_\tau=m, i\in\mathcal{M}_m^{in}} P_{i\tau}(j)}_{\equiv q_{jm}^{in}(\mathbf{p}_m)}$$

for each product j and where $P_{i\tau}(j)$ follows equation (3.2).

If product j is available only in the at-home segment (e.g. if it is a large multi portion product), then $P_{i\tau}(j) = 0$ for all $i \in \mathcal{M}_m^{out}$. However, for products available in both on-the-go and at-home segments the market demand curve depends on purchase probabilities (and hence preferences) in both segment.

Firm f's (variable) profits in market m are given by:

$$\Pi_{fm} = \sum_{j\in F_f} (p_{jm} - c_{jm}) q_{jm}(\mathbf{p}_m) \tag{3.4}$$

and the firm's first order conditions are:

$$q_{jm}(\mathbf{p}_m) + \sum_{k\in F_f} (p_{km} - c_{km}) \frac{\partial q_{km}(\mathbf{p}_m)}{\partial p_{jm}} = 0 \quad \forall j \in F_f. \tag{3.5}$$

Under the assumption that observed market prices are an equilibrium outcome of the Nash-Bertrand game played by firms, and given our estimates of the demand function, we can invert the first order conditions to infer marginal costs c_{jm}.

The introduction of a tax creates a wedge between post-tax prices, $\mathbf{p}$, and pre-tax prices, which we denote $\tilde{\mathbf{p}}$. The volumetric tax, π, on sugary soft drinks implies pre-tax and post-tax prices are related by:

$$p_{jm} = \begin{cases} \tilde{p}_{jm} + \pi l_j & \forall j \in \Omega_{ws} \\ \tilde{p}_{jm} & \forall j \in \Omega_{wd} \bigcup \Omega_{as} \bigcup \Omega_{an} \end{cases}$$

where l_j is the volume of product j.

In the counterfactual equilibrium, prices satisfy the conditions:

$$q_{jm}(\mathbf{p}_m) + \sum_{k\in F_f} (\tilde{p}_{km} - c_{km}) \frac{\partial q_{km}(\mathbf{p}_m)}{\partial p_{jm}} = 0 \quad \forall j \in F_f. \tag{3.6}$$

for all firms. We solve for the new equilibrium prices as the vector that satisfies the set of first order conditions (equation (3.6)) when $\pi = 0.25$.[20] Tax pass-through describes how much of the tax is shifted through to post-tax prices, for products

[20]We solve for a new equilibrium price for each of the products belonging to the main soda brands; we assume there is no change in the producer price of the composite other soft drinks

$j \in \Omega_{ws}$, we measure this as the difference in the post-tax and pre-tax equilibrium consumer price over the amount of tax levied, πl_j.[21]

3.4 Demand estimates

3.4.1 On-the-go

Preference distribution and elasticities In Table 3.1 we summarize the parameter estimates for the distribution of consumer specific preference parameters for on-the-go behavior obtained by maximizing the likelihood function (equation 3.3). We report the means, standard deviations, skewness and kurtosis for the price, soft drinks and sugar preference parameters, as well as the covariance between them. These numbers are based on the finite portion of the joint preference distribution. In Appendix B.1 we report coefficients estimates on brand, size and weather effects.

In Figure 3.1 we plot the marginal preference distributions for price, and the soft drinks and sugar product attributes for the on-the-go segment. These are based on individual level preference estimates, so we have a measure of statistical significance for each individual; this is represented by the shading, which indicates consumers with negative, positive and indifferent (i.e. not statistically significantly different from zero) preferences for each attribute. Table 3.1 shows that moments of each of these distributions are estimated with a high degree of statistical significance. Figure 3.1 makes clear that the univariate preference distributions depart significantly from normality (which is typically imposed in random coefficient models) – this is apparent both in the negative (for price and sugar) and positive (for soft drinks) skew in the preference distributions, and also in the infinite portions of the soft drinks and sugar preference distributions.

The estimates of the consumer specific preference parameters (on price, sugar and soft drinks) reveal a large degree of heterogeneity in preferences across individuals – the standard deviation for price preferences is 2.7 (with a coefficient of variation of 0.9), while the standard deviation for sugar and soft drinks is 1.6 and 1.8. Price sensitive consumers tend to have relatively strong soft drinks preferences (the correlation coefficient between price and soft drinks preferences is -0.33), while those with strong preferences for sugar tend to be less price sensitive (the correla-

brand (which aggregates together many very small soft drinks brands). We also assume no pricing response for the set of outside products.

[21] We solve for separate price equilibrium in each retailer-time period. Instead of solving a new price equilibrium in every month (which would imply 737 separate markets), for each retailer type, we solve for equilibrium prices in each year, which entails solving for equilibrium prices in 66 markets.

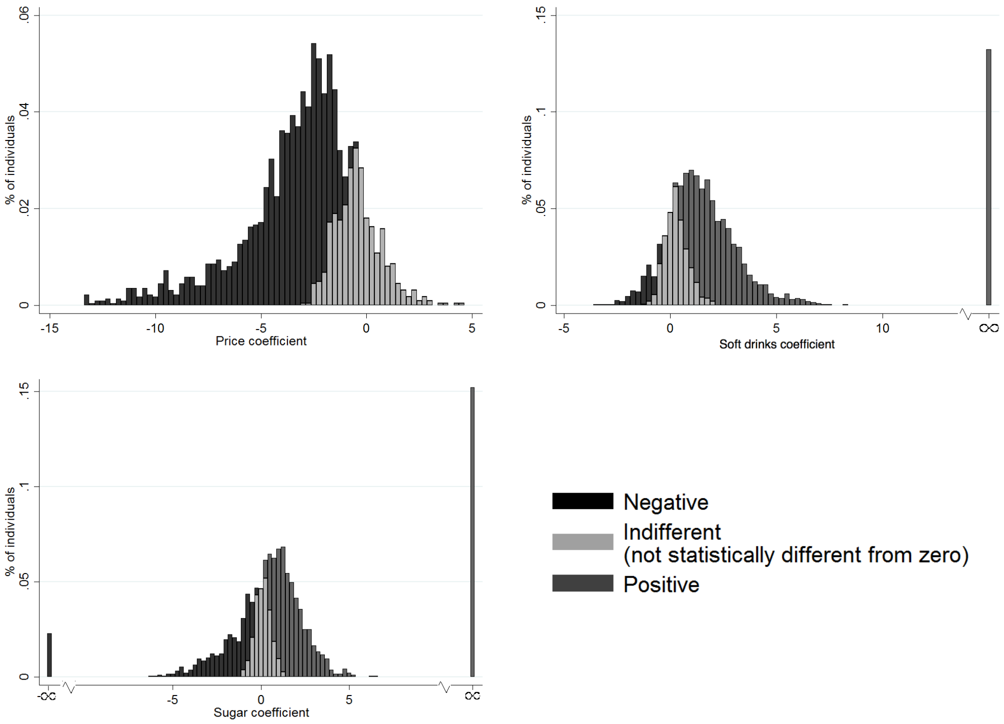
tion coefficient between price and sugar preferences is 0.15). We show contour plots of the bivariate preference distributions in Appendix B.1.

Table 3.1: *Demand model estimates - on-the-go*

Moments of distribution of consumer specific preferences			
Variable		Estimate	Standard error
Price (α_i)	Mean	-3.0737	0.0287
	Standard deviation	2.6825	0.0210
	Skewness	-0.9247	0.0462
	Kurtosis	4.3175	0.1117
Soft drinks (γ_i)	Mean	1.4297	0.0421
	Standard deviation	1.6065	0.0153
	Skewness	0.5001	0.0415
	Kurtosis	3.6833	0.1307
Sugar (β_i)	Mean	0.4244	0.0104
	Standard deviation	1.8058	0.0141
	Skewness	-0.4838	0.0407
	Kurtosis	3.5801	0.1026
Price-Soft drinks	Covariance	-1.4058	0.0463
Price-Sugar	Covariance	0.7413	0.0439
Soft drinks-Sugar	Covariance	-0.6585	0.0311
Demographic specific carton-size effects ($\delta^z_{d(i)}$)		Yes	
Demographic specific weather effects ($\delta^h_{d(i)}$)		Yes	
Time-demographic-brand effects ($\xi_{d(i)b(j)t}$)		Yes	
Retailer-demographic-brand effects ($\zeta_{d(i)b(j)r}$)		Yes	

Notes: We estimate demand on a sample of 2,374 individuals who we observe on 361,863 on-the-go choice occasions. Estimates of the consumer specific preferences are summarized in the table. Moments of distribution are computed using estimates of consumer specific preference parameters. These moments are based on consumers with finite parameters and omit the top and bottom percentile of each distribution. Standard errors for moments are computed using the delta method.

Figure 3.1: *Univariate distributions of consumer specific preference parameters - on-the-go*



Notes: Distributions are based on individual level preference parameter estimates based for the 2,374 individuals in the on-the-go estimation sample. We trim the top and bottom percentile of the finite portion of each distribution. The shading denote statistical significance of individual level preference estimates at the 95% level.

We report a selection of price elasticities in Table 3.2. The top panel of the table reports elasticities for products that belong to the two most popular brands, Coca Cola and Pepsi.[22] In column 1 we report the percent change in demand for the product when its price increases by 1%. Columns 2-4 report how demand for alternative products (sugary soft drinks, diet soft drinks and alternative sugary drinks) would change and a final column reports what would be the overall change in demand for non-alcoholic drinks. For example, a 1% increase in the price of the most popular sugary product, Coca Cola 500 (a 500ml bottle of Coca Cola), would result in a reduction in demand for that product of 2.27%. Demand for alternative sugary soft drinks would rise by around 0.41%, demand for diet soft drinks would rise by 0.18% and demand for alternative sugary drinks would rise by 0.25%. Demand for soft drinks and alternative sugary drinks as a whole would fall by 0.06%.

A couple of interesting patterns are apparent. First, consumers are more willing to switch from sugary soft drinks products to alternative sugary soft drinks and from diet products to diet alternatives, than they are between sugary and diet products. Second, the price elasticities for the 500ml products are smaller in magnitude than

[22] We show elasticities for all products in Appendix B.2.

for the 330ml versions; consumers that choose to buy the larger bottle variants rather than smaller cans, tend to be less willing to switch away from their chosen product in response to a price increase. This is precisely the opposite pattern from what we would get in a logit choice model without preference heterogeneity, in which the functional form imposes that own price elasticities are approximately proportional to price and therefore the higher price bottles would be more price elastic than cans.

The bottom panel of Table 3.2 reports the effect on demand of a marginal increase in the price of all sugary soft drinks and in the price of all soft drinks (i.e. both sugary and diet). The own price elasticity for soft drinks is -0.43. This is smaller than the own price elasticity of any individual soft drink product. The own price elasticity for sugary soft drinks is -0.89. This is larger than for all soft drinks, reflecting that some consumers respond to an increase in the price of sugary soft drinks by switching to diet alternatives.

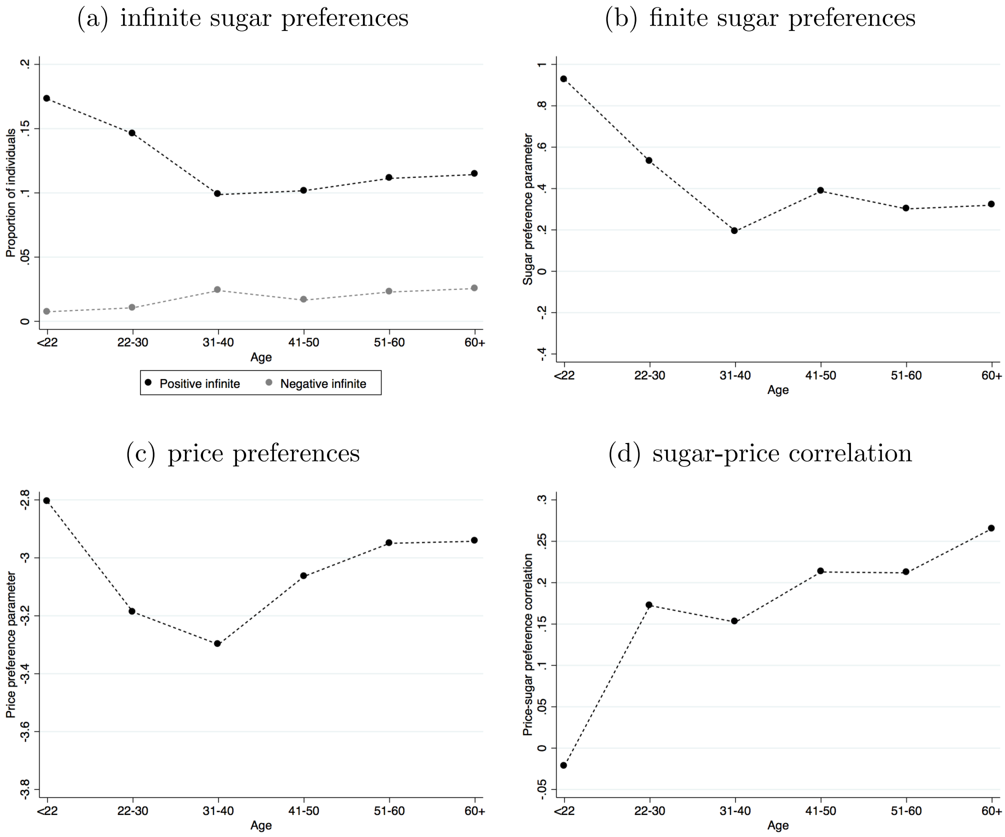
Table 3.2: *Price effects – on-the-go*

	Effect of 1% price increase on:				
	Own	cross demand for:			Total
	demand	sugary soft drinks	diet soft dirnks	sugary alternatives	demand
Coca Cola 330	-2.91	0.16	0.08	0.07	0.02
Coca Cola 500	-2.27	0.41	0.18	0.25	-0.06
Coca Cola Diet 330	-2.90	0.06	0.29	0.02	0.02
Coca Cola Diet 500	-2.72	0.14	0.53	0.09	-0.05
Pepsi 330	-3.08	0.06	0.03	0.03	0.01
Pepsi 500	-2.76	0.17	0.08	0.10	-0.03
Pepsi Diet 330	-3.41	0.02	0.13	0.01	0.01
Pepsi Diet 500	-3.28	0.06	0.25	0.03	-0.03
Soda	-0.43			1.36	-0.30
Sugary soda	-0.89		0.86	1.09	-0.17

Notes: For each of the eight products listed we compute the change in demand for that product, for alternative sugary and diet options and for total demand resulting from a 1% price increase. We also compute demand response for a 1% increase in the price of all soft drink products and all sugary soft drink products. Numbers are means across time.

Relationship with individual attributes A key feature of our model is that it allows us to flexibly relate preference parameters to characteristics of consumers. This enables us to address the question of how well targeted soda taxes are, and to what extent they disproportionately impact the young and the poor.

Figure 3.2: *Preference variation with age*



(a) infinite sugar preferences

(b) finite sugar preferences

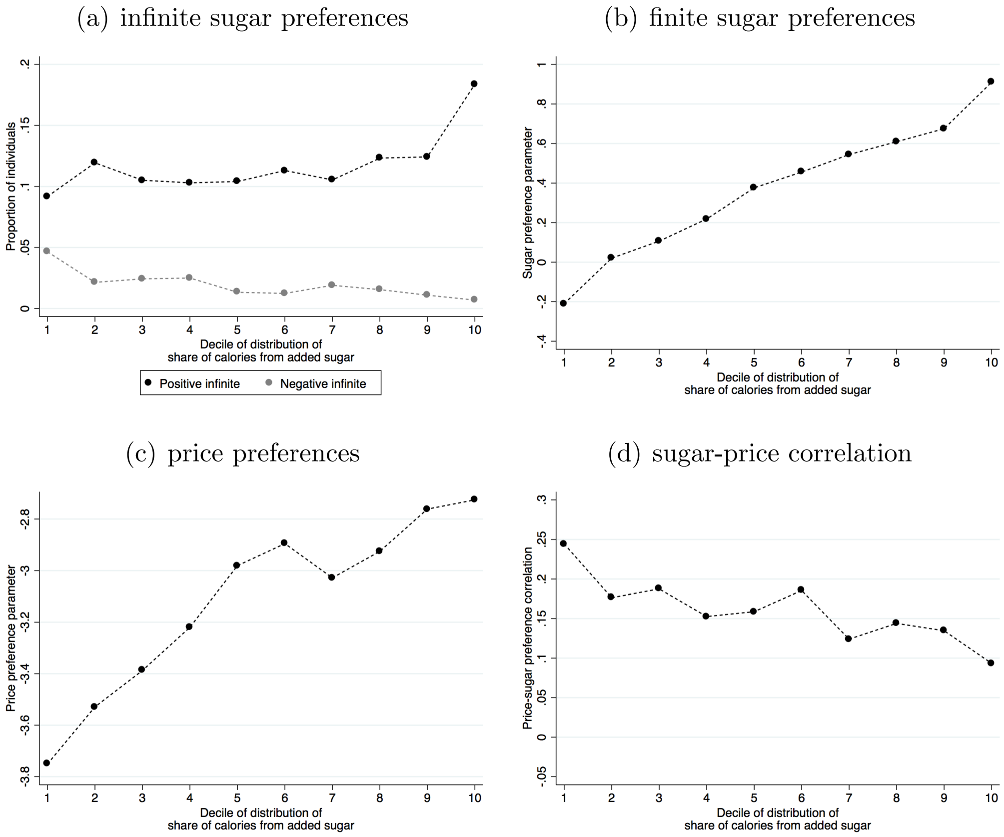
(c) price preferences

(d) sugar-price correlation

Notes: Figure shows how the share consumers with infinite sugar preferences, the mean of finite sugar preferences, the mean of price preferences and the correlation between sugar and price preferences vary by age groups.

In Figure 3.2 we show how features of the preference distribution vary with age. Panel (a) shows how the fraction of consumers with infinitely negative and positive sugar preferences varies across groups – a higher fraction of individuals aged below 30 have infinitely positive sugars preferences (i.e. only buy sugary varieties) than older individuals. Panel (b) shows that, for those individuals with finite sugar preferences, the mean preference for sugar varies with age, with the youngest group of individuals (aged 13-21) having stronger sugar preferences than older individuals. Panel (c) shows that the youngest group of consumers tend to have slightly less negative price preferences than older individuals, though the difference is small. Panel (d) shows how the *within* age group correlation in sugar and price preferences varies across age groups. Among older groups sugar and price preferences are positively correlated – those with the strongest sugar preferences are also the least price sensitive. However, among the youngest group this correlation in price and sugar preferences is close to zero. These preference patterns are important in determining the shape of demand and in driving how the responses to a soda tax vary across the age distribution.

Figure 3.3: *Preference variation with total annual dietary sugar*



Notes: Figure shows how the share consumers with infinite sugar preferences, the mean of finite sugar preferences, the mean of price preferences and the correlation between sugar and price preferences vary by deciles of the distribution of total annual dietary sugar.

Figure 3.3 shows how price and sugar preferences vary across deciles of the distribution of total annual dietary sugar (measured as the share of a households' total at-home calories from added sugar). Preferences governing on-the-go drinks demand are strongly related to total annual dietary sugar. Those in the top decile of the added sugar distribution are considerably more likely than other individuals to have infinitely positive sugar preferences. For those individuals with finite sugar preferences, being in a higher decile of the added sugar distribution is strongly associated with a stronger sugar preferences when buying drinks on-the-go. Price preferences vary more strongly with total annual dietary sugar than with age, with those in the top of the top half of the added sugar distribution being considerable less price sensitive than those in the bottom half. On the other hand the within group correlation in sugar and price preferences varies less starkly across deciles than it does across age groups.

In Appendix B.1 we show how preferences vary across deciles of the distribution of total equivalized grocery expenditure (a proxy for income). There is a clear gradient for both sugar and price preference parameters; poorer individuals typically have stronger sugar preferences and more negative preferences for price than richer individuals.

3.4.2 At-home

In Table 3.3 we summarize estimates of the household specific preference parameters governing at-home demand. In Appendix B.1 we report estimates of the demographic specific preference parameters. It is difficult to directly compare the moments of the preference distribution reported in Table 3.3 with their on-the-go counterparts because the set of products over which demand is estimated is different. However, two noticeable differences are that in the at-home segment the difference between the mean soft drinks and mean sugar preferences is much larger – when purchasing for at-home consumption households' preference for soft drinks relative to sugary drinks is considerably stronger than for individuals purchasing for on-the-go consumption. A second noticeable difference is that in the at-home segment the correlation in preferences is less strong than in the on-the-go segment. In Section 5.2 we show these differences do not reverse our results on the effect of soda taxes on individual level consumption.

Table 3.3: *Demand model estimates – at-home*

Moments of distribution of consumer specific preferences			
Variable		Estimate	Standard error
Price (α_i)	Mean	-2.1051	0.0141
	Standard deviation	1.3932	0.0115
	Skewness	-1.3236	0.0398
	Kurtosis	5.3699	0.2014
Soft drinks (γ_i)	Mean	7.6998	0.3607
	Standard deviation	1.8600	0.2609
	Skewness	0.0561	0.2646
	Kurtosis	2.9723	0.2385
Sugar (β_i)	Mean	-0.4608	0.0086
	Standard deviation	1.5219	0.0106
	Skewness	0.0084	0.0378
	Kurtosis	3.1882	0.1023
Price-Soft drinks	Covariance	-0.2336	0.0517
Price-Sugar	Covariance	-0.1076	0.0157
Soft drinks-Sugar	Covariance	-0.0844	0.0853
Demographic specific carton-size effects ($\delta^z_{d(i)}$)		Yes	
Demographic specific weather effects ($\delta^h_{d(i)}$)		Yes	
Demographic specific time-brand effects ($\xi_{d(i)b(j)t}$)		Yes	
Demographic specific retailer-brand effects ($\zeta_{d(i)b(j)r}$)		Yes	

Notes: We estimate demand on a sample of 3,314 households who we observe on 302,383 at-home choice occasions. Estimates of the consumer specific preferences are summarized in the table. Moments of distribution are computed using estimates of consumer specific preference parameters. These moments are based on consumers with finite parameters and omit the top and bottom percentile of each distribution. Standard errors for moments are computed using the delta method.

4 The effects of a soda tax

We use our demand estimates, along with the supply side model outlined in Section 3.3, to simulate the introduction of a tax levied on sugary soft drinks.[23] To compute supply side responses to the tax we use both the on-the-go and at-home demand estimates. In this section we focus on how individual level on-the-go demand changes in response to a tax. In Section 5.2 we discuss how responses vary across households in the at-home segment, showing, in particular, that at-home responses are unlikely to undo the individual level effects that the on-the-go segment enables us to study.

4.1 Market equilibrium

We consider the introduction of a tax of 25 pence per liter. This tax is similar to what has been implemented in some cities and counties in the US and also implies equilibrium price changes that are of a roughly similar order to the observed price changes in our data period. By construction, for soft drink brands with sugar, such a tax will be larger for larger sized products, imposing more tax on a 500ml bottle than on a 330ml can. We simulate the introduction of the tax allowing for price re-optimization of the set of branded sugary and diet soft drinks products (the former are subject to the tax, the latter are not). We hold fixed the pre-tax price of products belonging to the aggregate "other" soft drink brand, as well as for the alternative sugar drinks (fruit juice, flavored milk and flavored water) and bottled water.

In Table 4.1 we report the mean tax levied per product, price change and change in share of the on-the-go segment of the drinks market due to the tax. We report these for the set of "inside products" – the sugary and diet soft drinks – and for the alternative (non soft drink) sugary products and bottled water. The average tax liable on sugary soft drinks is 10.65 pence – for products with 500ml the tax liable is 12.5 pence, while for those with 330ml it is 8.25 pence. On average, the price of sugary soft drinks rises by 13.15 pence – average equilibrium pass-through of the tax is therefore around 120%. Important in driving this over shifting of the tax is that it is imposed in all stores and on a broad set of products owned by competitor firms. For instance, if we impose a sugary soda tax only on the products owned by the largest firm in the market, Coca Cola, the average pass-through of the tax onto its products is less than 100%.

Pass-through rates vary across products; the larger 500ml bottled products typically have rates in of 130-140% and smaller 330ml canned products have rates of

[23]In Appendix C we show results for a tax levied on all soft drinks.

around 100%. This means manufacturers respond to the tax by increasing margins on the 500ml products and maintaining them at around the pre-tax level for the 330ml cans. Our demand estimates imply that the bottled products have less elastic demands than the cans. By raising margins on these products, firms sacrifice some marginal consumers, who switch to alternatives, but earn more profits on the infra-marginal consumers who still buy bottles. Nevertheless, profits on the bottled products fall, while profits on the canned products, in some cases, rise as some consumers respond to the tax by downsizing (i.e. switching from bottles to cans).

Table 4.1: *Effects of sugary soda tax at product level on products available in the on-the-go segment*

	Tax (pence)	Δ price (pence)	Δ share (p.p.)
Sugary soda	10.65	13.15	-6.34
Coca Cola 330	8.25	8.12	-0.08
Coca Cola 500	12.50	17.48	-1.94
Dr Pepper 330	8.25	8.11	-0.01
Dr Pepper 500	12.50	17.33	-0.32
Fanta 330	8.25	8.24	-0.02
Fanta 500	12.50	17.29	-0.35
Cherry Coke 330	8.25	8.09	0.00
Cherry Coke 500	12.50	17.01	-0.26
Oasis 500	12.50	17.00	-0.41
Pepsi 330	8.25	8.66	-0.08
Pepsi 500	12.50	16.19	-0.88
Lucozade Energy 380	9.50	10.83	-0.17
Lucozade Energy 500	12.50	18.33	-0.36
Ribena 288	7.20	7.11	0.03
Ribena 500	12.50	18.07	-0.23
Other soft drinks	12.50	12.50	-1.27
Diet soda	0.00	-1.37	3.96
Coca Cola Diet 330	0.00	-0.77	0.63
Coca Cola Diet 500	0.00	-2.12	1.26
Dr Pepper Diet 500	0.00	-2.02	0.21
Fanta Diet 500	0.00	-1.92	0.23
Cherry Coke Diet 500	0.00	-1.84	0.17
Oasis Diet 500	0.00	-1.97	0.27
Pepsi Diet 330	0.00	-0.52	0.21
Pepsi Diet 500	0.00	-1.05	0.49
Ribena Diet 500	0.00	-1.54	0.13
Other Diet soft drinks	0.00	0.00	0.37
Sugary alternatives	0.00	0.00	1.09
Bottled water	0.00	0.00	1.28

Notes: Panels 2 and 4 show the mean effect of the sugary soda tax on price and market share of products in the on-the-go segment. Panels 1, 3 and 4 show the mean effects of the tax on all sugary soft drinks, all diets soft drinks and on alternative drinks.

The tax on sugary soft drinks thus increases equilibrium prices for sugary soft drinks, doing so by more for the larger sized products due to a higher tax rate and over shifting. The market share of sugary soft drinks in the on-the-go segment falls by 6.34 percentage points. Soft drink manufacturers also optimally respond to the tax by lowering the price of diet products. The average reduction in price is 1.37 pence, however, the 500ml bottle products see larger price reductions of around 2 pence, with smaller changes in the equilibrium price of the smaller 330ml canned products. The pricing response of soft drink manufacturers therefore acts to magnify the price differential that the tax creates between sugary and diet products. Relative to the case in which producers simply increase consumer prices by an amount exactly equal to the tax (so pass-through of tax is 100%), firms' equilibrium pricing response induces more switching away from sugary soft drinks and more towards diet soft drinks; the share of diet soft drinks increases by 3.96 percentage points. Alternative sugary drinks and bottled water also see increases in market share of 1.09 and 1.28 percentage points.

A number of papers use observed tax changes to estimate pass-through of taxes to prices. These include Besley and Rosen (1999), which exploits variation in state and local sales taxes in the US and looks at the impact on prices of a number of products and finds over shifting for soda products, Delipalla and O'Donnell (2001), which analyzes the incidence of cigarette taxes in several European countries and Kenkel (2005), which uses data on how the price of alcoholic beverages changed in Alaska. Results from the literature vary, but typically these papers find complete or over shifting of specific taxes, which broadly accord with our pass-through results.

Evidence from papers that study recently implemented taxes imposed on soft drinks is mixed; comparing taxed and non-tax products, Grogger (2015) finds that prices rose by more than the amount of the tax following the adoption of the Mexican soda tax in 2014, while Cawley and Frisvold (2017) find under-shifting of the Berkeley soda tax, which they rationalize as due to the ease with which consumers can avoid the tax by shopping in neighboring municipalities.[24] Cawley et al. (2018a) use a differences-in-differences approach and find evidence of average pass-through of around 100% for the Philadelphia soda tax (applied to diet drinks as well as sugary ones) that was introduced in 2017.[25] In an ex ante study of the effects of a sugary soda tax in France, Bonnet and Réquillart (2013) find pass-through that

[24]Taylor et al. (2018) show evidence that soda sales changed due to the campaign attention and election outcome in Berkeley, California, months before the tax was adopted, suggesting that the media coverage and the election may have had an important impact on purchasing behavior.

[25]Cawley et al. (2018b) also uses a difference-in-difference approach based on consumers within and outside the city, to show a reduction in consumption of taxed drinks, with larger effects for adults than for children.

exceeds 100% and also reductions in the prices of diet products. The empirical literature on pass-through of cigarette taxes is similarly mixed; Harding et al. (2012) find taxes in the US are under-shifted and that avoidance opportunities have a sizeable effect on purchases, while Lillard and Sfekas (2013) find evidence of over shifting once the implicit tax in state lawsuits is taken account off.

There is also a related literature that looks at pass-through of cost shocks. Much of this finds under-shifting (see, for instance, Goldberg and Hellerstein (2013) and Nakamura and Zerom (2010)). An important reason for incomplete pass-through of cost shocks is that often not all cost components are affected by the shock. For instance, exchange rate movements do not directly impact the cost of non-traded inputs (Goldberg and Hellerstein (2013)). In a context where firms' marginal costs are observable (in the wholesale electricity market), Fabra and Reguant (2014) find changes in marginal costs are close to fully shifted to prices.

4.2 How well targeted is the tax?

Our tax simulation suggests that, on average, soft drinks purchasing consumers will lower the total amount of sugar they purchase from soft drinks on-the-go by around 195g per annum, which represents an average reduction of 18% of sugar from soft drinks. However, some of this reduction is off set by switching to alternative (non-taxed) drinks that contain sugar. The average reduction in sugar from drinks is around 170g. However, the distribution of reductions in sugar is right skewed with the 75th, 90th, 95th and 99th percentiles being 201g, 451g, 706g and 1388g.

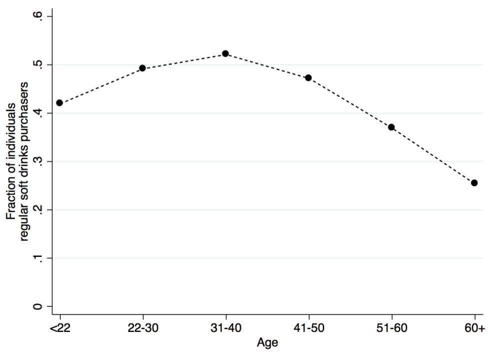
Key to understanding the effectiveness of a soda tax is whether it successfully achieves reductions in sugar amongst the targeted groups of consumers – the young and those with high total annual dietary sugar. In Figure 4.1 we show how the effects of the tax vary across these dimensions. The left hand graphs show how the fraction of individuals who are soft drinks purchasers (and therefore are affected by the tax) varies with age (panel (a)), total annual dietary sugar deciles (panel (c)) and jointly with these variables (panel (e)). The graphs show those aged below 40 (and to a lesser extent those at the top of the dietary sugar distribution) are considerably more likely to be impacted by the tax. The right hand graphs show how soft drink purchasers alter their sugar purchases as a result of the tax. Panels (b) and (d) show how the fall in sugar from soft drinks and sugar from all drinks varies with age and total annual dietary sugar. As all groups respond to some extent to the tax by switching to sugary alternatives, the reductions in sugar from soft drinks are larger than from all drinks. Panel (f) shows how the fall in sugar from all drinks varies jointly with age and dietary sugar. The level reductions in sugar are larger for young than for older individuals – the youngest group exhibit a reduction in sugar from drinks that is around 40% as large as those aged over 40. In contrast, soft drink purchasing individuals at the top of the added sugar distribution actually lower their sugar consumption by less than those at the bottom of the distribution.

The figure shows the policy is relatively effective at targeting young consumers – they are both more likely to be impacted by the policy and, conditional on this, exhibit bigger level responses than older groups. Note, however, that the average percent reduction in sugar from drinks is actually lowest (14% vs 18% across all individuals) for those aged below 22. Key to the tax influencing the sugar purchases of this group by the most is that they begin by obtaining a relatively large amount of sugar from products targeted by the tax. However, the policy is less able to target individuals that have high total annual dietary sugar. Such individuals are more likely to be soft drinks purchasers and therefore be impacted by the policy than those lower down the dietary sugar distribution, but, conditional on being affected by the policy, their response, on average, is smaller in level terms (and much smaller in percentage terms – for instance the reduction for the top decile of the dietary sugar distribution is 10 percentage points below that for the bottom decile). This is driven by those with high levels of dietary sugar both having strong sugar preferences and being relatively price insensitive. In contrast, while the young have strong preferences for sugar, they are not particularly price insensitive.

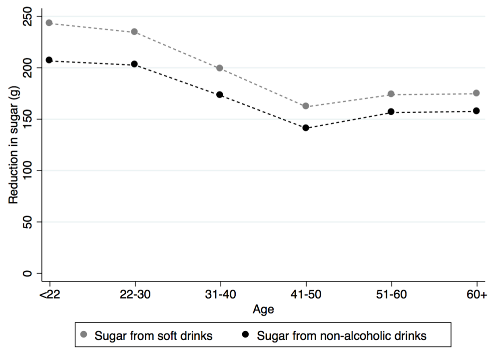
Figure 4.1: *Reductions in sugar*

Effect by age

(a) soft drinks purchasers

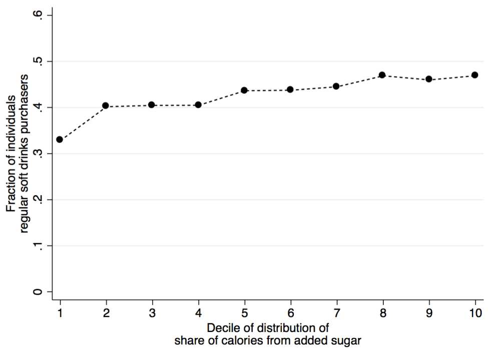


(b) fall in sugar from soft and all drinks

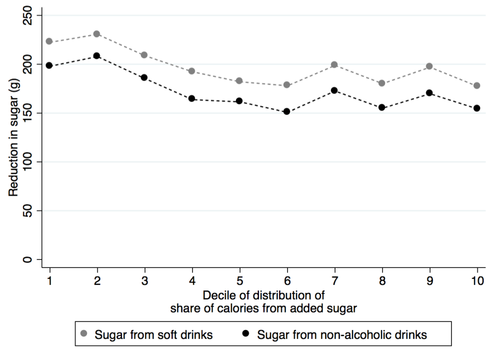


Effect by total annual dietary sugar

(c) soft drinks purchasers

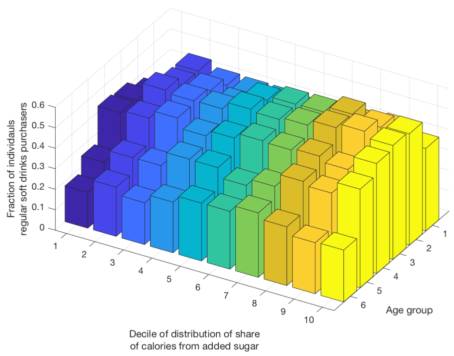


(d) fall in sugar from soft and all drinks

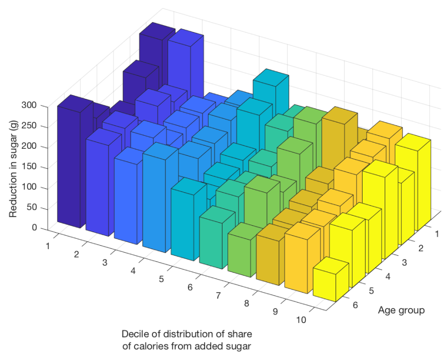


Effect by age and total annual dietary sugar

(e) soft drinks purchasers



(f) fall in sugar from all drinks



Notes: Numbers are for the on-the-go segment. Panels (a) and (c) show how the fraction of individuals in the estimation sample (i.e. that are soft drinks purchases) varies across age groups and deciles of the distribution of share of calories from added sugar. Panel (e) shows the joint variation across these dimensions Panels (b), (d) and (f) show variation in the reduction in sugar conditional on being a soft drinks purchaser. In panel (e) and (f) age groups are 1=<22, 2=22-30, 3=31-40, 4=41-50, 5=51-60, 6=60+.

4.3 Consumer welfare

Higher taxes, to the extent they raise prices, impose an economic burden on consumers; after a tax is introduced consumers can obtain less produce for a given amount of expenditure than before. In the case of a tax on sugary soft drinks, consumers that buy sugary soft drinks will incur a welfare loss through this channel. Those consumers that never buy soft drinks will see no change in their welfare (we assume that the prices of non soft drinks are unaffected by the tax), while those individuals that drink diet soft drinks may actually benefit slightly as the optimal pricing response to the tax is to lower the price of diet soft drinks.

In Figure 4.2 we describe this effect; we use the preference estimates to compute compensating variation – the monetary amount an individual would require to be paid to be indifferent to the imposition of the tax based on their estimated preferences. Letting p_{jrt} and p'_{jrt} denote the retailer r time t price of product j prior to and following the introduction of the tax, the expected compensating variation for individual i on a choice occasion is given by (Small and Rosen (1981)):

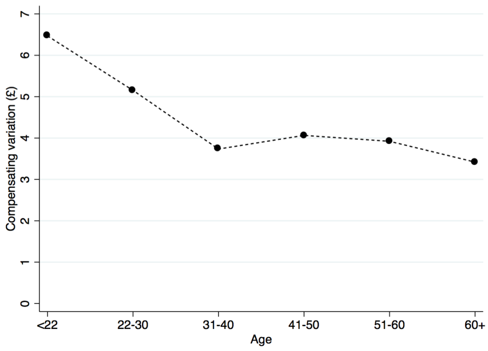
$$cv_{i\tau} = \frac{1}{\alpha_i}\left[\ln\left(\sum_{k\in\Omega_i\cap\Omega_{r_\tau}} \exp(v_{ijr_\tau t_\tau} + \eta_{ijr_\tau t_\tau} - \alpha(p_{krt_\tau} - p'_{krt_\tau})\right) - \right.$$
$$\left. \ln\left(\sum_{k\in\Omega_i\cap\Omega_{r_\tau}} \exp(v_{ijr_\tau t_\tau} + \eta_{ijr_\tau t_\tau}\right)\right]$$

where $v_{ijr_\tau t_\tau}$, and $\eta_{ijr_\tau t_\tau}$ are defined in Section 3.1. Summing $cv_{i\tau}$ over all the consumer's choice occasions in the year gives their annual compensating variation. We show, for the set of soft drinks purchasers, how this varies by an individual's age (panel (a) of Figure 4.2), position in the distribution of total annual dietary sugar (panel (b)), and jointly with these variables (panel (c) of Figure 4.2). Compensating variation is falling across age groups and rising across total dietary sugar deciles; on average, both the youngest and those with high total dietary sugar have the highest compensating variations. This is because both groups are more likely to purchase large quantities of soft drinks. Panel (c) of Figure 4.2 shows that the highest compensating variations are concentrated about young individuals regardless of their position in the total annual dietary sugar distribution.

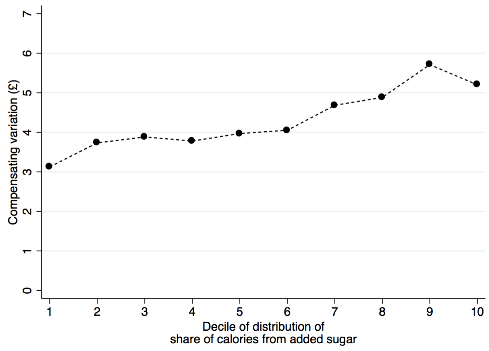
If consumers fully took account of all the costs associated with their soft drink consumption, then compensating variation would capture the total effects of the tax on consumer welfare and we could conclude that the tax makes all purchasers of sugary soft drinks worse off, with the largest effects being among the young and those with high levels of dietary sugar. However, if sugary soft drinks consumption is associated with future costs that are not taken account of by the individual at the point of consumption (internalities), or with costs imposed on others (externalities), then compensating variation measured based on revealed preference captures only part of the total consumer welfare effect of the tax.

Figure 4.2: *Revealed consumer welfare effect*

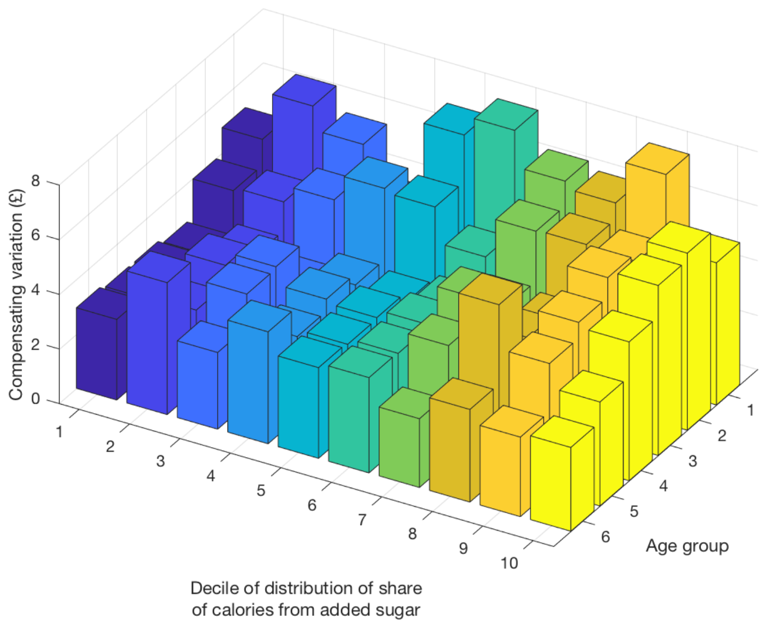
(a) by age



(b) by total dietary sugar



(c) by age and total dietary sugar



Notes: Numbers are for the on-the-go segment and are based on estimation sample of 2,374 individuals. Numbers show how the mean compensating variation varies by age and deciles of the distribution of share of calories from added sugar. In panel (c) age groups are 1=<22, 2=22-30, 3=31-40, 4=41-50, 5=51-60, 6=60+.

The potential consequences of consuming sugar in excess are well established. It may be that some individuals fully internalize these future costs when deciding whether to consume sugary soft drinks. However, there is a large theory literature that posits that not all individuals fully account for future costs of consumption (for a survey see Rabin (1998)) and there is evidence this is particularly relevant for food, both experimental (for instance Read and Van Leeuwen (1998) and Gilbert et al. (2002)) and circumstantial, through the existence of a multi-billion pound dieting industry (Cutler et al. (2003)).

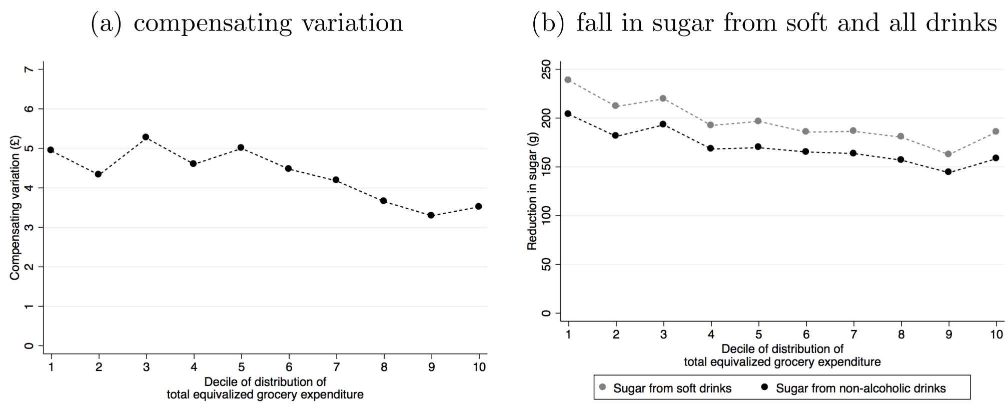
The young are particularly susceptible to suffer from internalities from excess sugar. The consequences of poor nutrition early in life are profound: with excess sugar associated with poor mental health and school performance in children, and poor childhood nutrition thought to be an important determinant of later life health, social and economic outcomes and of persistent inequality. (see, for instance Cawley (2010)).[26] Few would argue that these significant costs are fully taken account of by children and young adults when making consumption decisions. The average compensating variation for individuals aged 13-21 and who are soft drink purchasers is around £6.50, while the average reduction in sugar for this group is around 205g. If the internality associated with drinking the amount of sugar in a can of Coca Cola is above £1.11, then, for the average person aged 13-21, the soda tax will be welfare improving.

4.4 Redistribution

A common criticism of excise taxes is that they are regressive. This is typically based on the observation that those with lower incomes tend to be relatively heavy consumers of the taxed product (which, for a small change in price, is a good approximation to the revealed consumer welfare cost). Table 2.9 confirms that, in the case of sugary soft drinks, poorer individuals (those with low total annual equivalized grocery expenditure) are more likely to be soft drink purchasers and to get more sugar from these products; those in the bottom half of the distribution are around 10% more likely to be soft drinks purchasers than those in the top half, and conditional on being one, on average obtain 15% more sugar from these products. Based on our demand and supply model we estimate the true revealed welfare cost from the tax – Panel (a) of Figure 4.3 shows how this varies across deciles of the equivalized grocery expenditure distribution for soft drink purchasers. Compensating variation is around 25% higher, on average, for individuals in the bottom half of total equivalized grocery expenditure distribution than for those in the top half.

[26] Cavadini et al. (2000) document an increase in soft drink consumption in the US for 11-18 years old of almost 300% for boys, and over 200% for girls between 1965 and 1996. Nielsen and Popkin (2004) document a contemporaneous fall in the share of calories children get from milk. Medical evidence suggests that exposure to sweetened beverages early in life can establish strong lifelong preferences for these products (Mennella et al. (2016)).

Figure 4.3: *Effects by total equivalized grocery expenditure*



Notes: Numbers are for the on-the-go segment and are based on estimation sample of 2,374 individuals. Panel (a) shows how compensating variation and panel (b) shows how reductions in sugar, from tax varies across deciles of the distribution of total equivalized grocery expenditure. .

However, if some consumers impose internalities on themselves, then the standard revealed consumer welfare loss (compensating variation) provides an incomplete picture of the redistributive effects of the tax (a point made by Gruber and Koszegi (2004) in the case of cigarette taxation). Panel (b) of Figure 4.3 shows that mean sugar reductions from the tax are somewhat higher on average among those towards the bottom of the equivalized grocery expenditure distribution compared to those further up (183g for the bottom half of the distribution versus 157g for the top). Therefore, if the prevalence of internalities is broadly constant across the expenditure distribution, the larger reductions in sugar among low spending individuals may be enough to offset much of the compensating variation difference.

In addition, there is some evidence that low income people are more likely to exhibit behavior that creates internalities. For example, Haushofer and Fehr (2014) and Mani et al. (2013) suggest that the stress and cognitive load of being in poverty means people are more likely to make unwise decisions and underweight the future. Focusing on asset accumulation Bernheim et al. (2015) argue that poverty can perpetuate itself by undermining the capacity for self-control: low initial wealth precludes self-control, and hence asset accumulation, creating a poverty trap. Banerjee and Mullainathan (2010) take an alternative approach by assuming that "temptation goods" are inferior goods, which leads to a similar conclusion that self-control problems give rise to asset traps. Any propensity for self-control problems, or other sources of internality generating behavior, that are concentrated among poorer individuals is likely to result in a soda tax being less regressive.

5 Robustness

5.1 Substitution to sugar in food

Our analysis so far has considered the impact of a soda tax, incorporating rich patterns of consumer switching across drinks. We have thus far not modeled the possibility that consumers respond to the tax by switching from sugary soft drinks to foods that contain sugar. In this section we explore how important consumer switching from sugar in soft drinks to sugar in food is likely to be. It would be numerically difficult to estimate our model with all food on-the-go items being simultaneous choices. Instead we embed our drinks model into a two stage food on-the-go choice model. We assume that the idiosyncratic unobserved shocks that affect the choice of which drink to consume are unknown in the first stage, thereby allowing us to reduce the dimensionality of parameters generating substitution between drinks and non drinks, whilst still taking account of the heterogeneity in consumer preferences for drinks. In the first stage the consumer takes expectations over the second stage i.i.d. extreme value shocks.

Specifically, suppose the choice model of Section 3 is a second stage of a two-stage decision process, which governs which drink to select, conditional on choosing to purchase a drink. Consider a first stage in which the consumer chooses between chocolate products, a non-sugary snack and a drink. Let $k = \{\varnothing, 1, \ldots, K, \mathcal{D}\}$ denote first stage options. $k = \varnothing$ denotes the first stage outside option of a non-sugary snack, $k = 1, \ldots, K$ indexes chocolate products and $k = \mathcal{D}$ indexes choosing a drink (with the specific drinks product determined by the second stage of the decision problem). Suppose utility from these options takes the form:

$$\begin{aligned} V_{i\varnothing\tau} &= \varepsilon_{i\varnothing\tau} \\ V_{ik\tau} &= \mu_c + W_{ik\tau} + \varepsilon_{ik\tau} \quad \text{for all } k \in \{1, ..., K\} \\ V_{i\mathcal{D}\tau} &= \mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}\tau} + \varepsilon_{i\mathcal{D}\tau}, \end{aligned}$$

where $W_{i\mathcal{D}\tau}$ is the expected utility from choosing the preferred drink product and can be computed using estimates of the second stage choice model and where

$$W_{ik\tau} = \alpha_i p_{krt_\tau} + \beta_i s_k + \vartheta_{b(k)}$$

is the product specific utility from choosing chocolate product k. We assume that the error terms, $(\varepsilon_{i0\tau}, \varepsilon_{i1\tau}, \ldots, \varepsilon_{iK\tau}, \varepsilon_{i\mathcal{D}\tau})$ are distributed i.i.d. extreme value. This extends our choice model to capture switching between drinks, chocolates and non-sugar snacks and allows us to estimate the strength of switching between non-alcoholic drinks and chocolate (see Appendix D for further details).

We estimate the extended choice model allowing both the constant in the drinks pay-off, $\mu_{i\mathcal{D}}$, and the parameter on the expected second stage utility from drinks, $\psi_{i\mathcal{D}}$, to vary across the six age groups across which we describe results in Section 4.2. Table D.1 in Appendix D shows that for each age group the coefficient estimate is positive and statistically significant indicating that an increase in the price of soft drinks (and thus a fall in the expected utility from choosing the preferred drink) does induce some switching away from drinks and towards foods. However, the strength of this switching to food between the baseline model (results presented in Section 4) and the extended two-stage model is relatively small. Taking account of switching to food sources of sugar dampens the mean overall reduction in sugar by between 9% (for those aged 30-39) to 13% (for those aged 21 or under) and has no bearing on the qualitative relationship that sugar reductions are considerably larger for younger individuals. More broadly, none of our conclusions about the impact of a soda tax are materially affected by accounting for the (limited) switching to sugar in food. Appendix D provides further details.

5.2 Effects in the at-home segment

Our main interest in this paper is in the on-the-go segment of the market, which has been much less well studied than the at-home segment. Our counterfactual simulations of price equilibria after the introduction of a tax account for supply linkages across the two segments, but are focused on individual level outcomes in the on-the-go segment of the market. It is difficult to say much about individual level outcomes with household level at-home data without placing considerably more structure on how purchases are shared within the household. Our use of on-the-go data enables us to avoid this, however, there might be concern that at-home responses off set our conclusions about the targeting of the tax based on the on-the-go segment.

In order to check the robustness of our results on the targeting of a policy that aims to reduce the consumption of sugary soft drinks, we simulate the effect on drinks purchases of a 10% increase in the price of sugary soft drinks and compare percentage responses in the on-the-go and at-home segment. To this point our focus has been on level response because this is what is relevant for whether the tax is able to tackle internalities and externalities associated with excess sugar consumption. However, as we do not know how drinks are distributed across individuals in the household, here we instead compare percent responses across the two segments.

Table 5.1 shows the percent changes in sugar from drinks for each of the age groups in the on-the-go and at-home sectors. In the on-the-go segment the young exhibit the smaller percentage changes (though of course, they exhibit the largest level changes). For the at-home segment we report percent reductions by age based on the age of the individual from the household that is in the on-the-go sample. In the at-home segment percent changes are close to monotonic in the age groups, with the households that the youngest individuals belong to exhibiting the largest percent reductions. The final column of the table gives the within age group correlation in percent changes in sugar in the two segments. For each group the correlation is positive, indicating that conditional on age groups, individuals with relatively large on-the-go responses are from households with relatively large at-home responses. Together we believe this evidence suggests it is unlikely that at-home response would undermine our conclusions on the individual targeting of soda taxes. An interesting step for future research would be to couple our demand and supply analysis with a collective model of within household consumption behavior to extent the targeting analysis to household purchase decisions.

Table 5.1: *Effects of 10% price increase on sugar from drinks*

Age goup	% change in sugar		On-the-go – at-home correlation
	On-the-go	At-home	
<22	-5.32	-8.52	0.18
22-30	-6.44	-7.43	0.30
31-40	-7.39	-7.29	0.19
41-50	-5.79	-7.52	0.24
51-60	-6.59	-7.67	0.18
>60	-7.53	-6.87	0.19

Notes: We simulate the effect of a 10% increase in the price of sugary soda. Columns 1 and 2 show the changes in sugar from drinks in the on-the-go and at-home segments. Column 3 shows the correlation in percent consumption responses between the two segments. The analysis is based on the full sample of 5,554 individuals.

5.3 Bias correction for incidental parameters problem

In our non-linear model with fixed effects, maximum likelihood estimates of the parameters may suffer from an incidental parameters problem, noted by Neyman and Scott (1948). Even if both $N \to \infty$ and $T \to \infty$, if N and T grow at the same rate ($\frac{N}{T} \to \rho$ where ρ is a non zero constant), our fixed effect estimator will be asymptotically biased (Arellano and Hahn (2007)). Bias correction methods exist that reduce the bias from being of order $1/T$ to $1/T^2$.

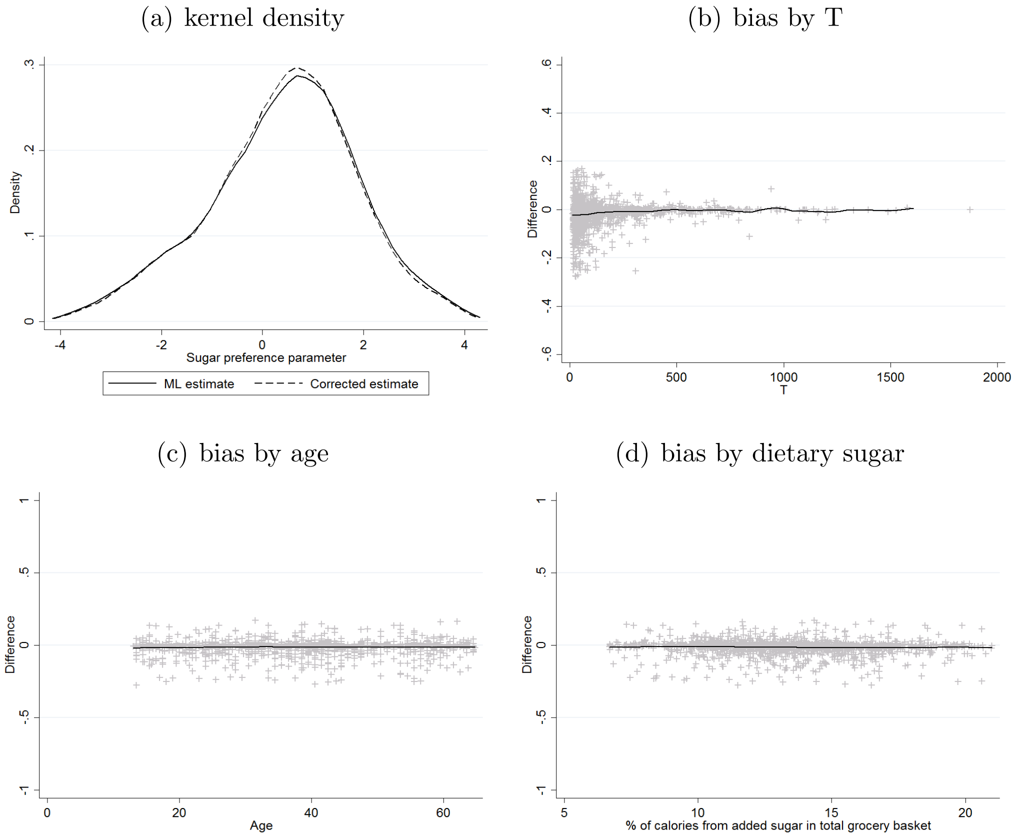
A range of bias correction methods exist (see surveys in Arellano and Hahn (2007), Arellano and Bonhomme (2011)). We use panel jackknife methods (Hahn and Newey (2004)), employing the split sample procedure suggested in Dhaene and Jochmans (2015). This entails obtaining estimates of the model parameters $\theta = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\eta})$ based on splitting the sample into two non overlapping random sub-samples. Each sub-sample contains one half of the choice occasions for each individual. We denote the maximum likelihood estimate for the full sample $\widehat{\theta}$ and the estimate for the two subsamples $\widehat{\theta}_{(1,T/2)}$ and $\widehat{\theta}_{(T/2,T)}$. The jackknife (bias corrected) estimator is:

$$\widetilde{\theta}_{split} = 2\widehat{\theta} - \frac{\widehat{\theta}_{(1,T/2)} + \widehat{\theta}_{(T/2,T)}}{2}.$$

In Figure 5.1 we graph the difference between the jackknife (bias corrected) and maximum likelihood sugar preference parameters for the on-the-go segment. Panel (a) shows the distribution of estimates (for those with finite sugar preferences) for the maximum likelihood and jackknife estimates. Panel (b) shows how the difference in these estimates relates to the time a consumer is in the sample. Panels (c) and (d) show how the difference relates to consumers' age and total dietary sugar.

The figure shows that the difference between the two estimates is small; the standard deviation of the sugar preference parameter estimates is 1.8, while the average absolute difference between the jackknife and maximum likelihood estimates is 0.06. The difference is decreasing in T; those in the sample for a relatively small number of choice occasions tend to have higher differences than those in the sample relatively many times. However, conditional on T, the average difference between the jackknife and maximum likelihood estimates is zero – a positive difference is equally likely as a negative difference. Indeed, the distribution of the maximum likelihood and jackknife estimates of the preference parameters are almost indistinguishable and the difference between the jackknife and maximum likelihood estimates is completely unrelated to individuals' age or total dietary sugar.

Figure 5.1: *Sugar preference parameters*



Notes: Graphs are based on preferences estimates in the on-the-go segment. In panels (b)-(d) markers represent consumer level differences. Lines are local polynomial regressions.

In Appendix B.3 we show that similar conclusions to those for sugar hold for estimated price and soft drink preferences; the maximum likelihood and jackknife procedures yield almost identical preference distributions, any individual level differences are relatively small and are equally likely to be positive or negative and there is no systematic relationship with the key variables that we relate our demand effects to. For instance, the standard deviation of the price preference parameter estimates is around 2.7, while the average absolute difference between the jackknife and maximum likelihood estimates is 0.2. For the soft drink preferences the numbers are 1.6 and 0.1. As a consequence, our results regarding the effectiveness of soda taxes are completely robust to the bias correction procedure.

6 Summary and conclusion

Corrective taxes have traditionally been applied to alcohol, tobacco and gambling. Recently there has been a drive to extend them to cover some types of foods, with soda taxes being at the vanguard of this move. The principal economic rationale for such taxes is that they discourage consumption that generates costs not taken account by individuals at the point of consumption. In the case of sugar, there is clear medical evidence that excess consumption can lead to large future costs, while almost all individuals exceed official recommendations on how much to consume. It is plausible that, at least for some consumers, these health costs are not factored in at the point of consumption. This is most obviously true for young people, but is also likely to be the case for some individuals with high total dietary sugar and who therefore are at elevated risk of suffering health problems. The efficacy of a soda tax relies on to what extent it can encourage these groups to avoid internalities and at what cost to consumers in terms of welfare loss associated with higher prices.

Our results show that young consumers would lower their sugar consumption by more than older individuals in response to a soda tax. The tax does therefore succeed in achieving relatively large reductions in sugar among one group most likely to suffer from internalities. However, the young also loose out most in terms of direct consumer surplus loss due to higher prices. The relatively large internalities some young people impose on themselves makes it likely that the gain from averted internalities will outweigh this. The performance of the tax in terms of reducing the sugar intake of those with the most sugary diets is less good – those with high total dietary sugar are relatively price inelastic and therefore fail to lower their sugar consumption in response to the tax by more than more moderate sugar consumers. Nevertheless, if internalities are sufficiently convex in total sugar, this group may still benefit from the tax. The redistributive properties of the tax are more attractive than one based purely on traditional economic tax incidence. While the traditional economic burden of the tax falls, to a moderate extent, disproportionately on the poor, the poor also lower their sugar consumption to a somewhat larger extent and therefore are likely to benefit by more than better off consumers due to averted internalities.

In our analysis we have taken account both of consumer demand responses and the equilibrium pricing response of soft drink manufacturers. In the longer run we would expect firms to also respond to the tax by changing their product portfolios and changing the sugar content of existing products. Our results therefore provide a picture to the short-medium run impact of soda taxes. An important direction for future work will be to incorporate how firm portfolio choice will be effected by such policies.

References

Adams, A., L. Cherchye, B. De Rock, and E. Verriest (2014). Consume now or later? Time inconsistency, collective choice, and revealed preference. *American Economic Review 104*(12), 4027–4070.

Aguiar, M. and E. Hurst (2007). Life-Cycle Prices and Production. *American Economic Review 97*(5), 1533–1559.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2018a). Regressive Sin Taxes, with an Application to the Optimal Soda Tax.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2018b). Should we tax soda? an overview of theory and evidence. *Journal of Public Economics*.

Allcott, H., S. Mullainathan, and D. Taubinsky (2014, April). Energy policy with externalities and internalities. *Journal of Public Economics 112*, 72–88.

Ameriks, J., A. Caplin, J. Leahy, and T. Tyler (2007). Measuring Self-Control Problems. *American Economic Review 97*(3).

Arellano, M. and S. Bonhomme (2011). Nonlinear Panel Data Analysis. *Annual Review of Economics 3*(1), 395–424.

Arellano, M. and J. Hahn (2007). Understanding bias in nonlinear panel models: Some recent developments. In *Econometric Society Monographs*, Volume 43, pp. 381.

Bajari, P. and C. L. Benkard (2005). Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy 113*(6), 1239–1276.

Bajari, P., J. T. Fox, and S. P. Ryan (2007). Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients. *American Economic Review 97*(2), 459–463.

Banerjee, A. and S. Mullainathan (2010, May). The Shape of Temptation: Implications for the Economic Lives of the Poor. Working Paper 15973, National Bureau of Economic Research.

Bernheim, B. D., D. Ray, and Ş. Yeltekin (2015). Poverty and Self-Control. *Econometrica 83*(5), 1877–1911.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from combination of micro and macro data: The new car market. *Journal of Political Economy 112*(1), 68–105.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics 25*, 242–262.

Besley, T. J. and H. S. Rosen (1999). Sales Taxes and Prices: An Empirical Analysis. *National Tax Journal 52*(2), 157–78.

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.

Burda, M., M. Harding, and J. Hausman (2008). A Bayesian mixed logit-probit model for multinomial choice. *Journal of Econometrics 147*(2), 232–246.

Cavadini, C., A. M. Siega-Riz, and B. M. Popkin (2000, July). US adolescent food intake trends from 1965 to 1996. *Archives of Disease in Childhood 83*(1), 18–24.

Cawley, J. (2010). The economics of childhood obesity. *Health Affairs 29*(3), 364–371.

Cawley, J., D. Frisvold, A. Hill, and D. Jones (2018a, September). The Impact of the Philadelphia Beverage Tax on Prices and Product Availability. Working Paper 24990, National Bureau of Economic Research.

Cawley, J., D. Frisvold, A. Hill, and D. Jones (2018b, September). The Impact of the Philadelphia Beverage Tax on Purchases and Consumption by Adults and Children. Working Paper 25052, National Bureau of Economic Research.

Cawley, J. and D. E. Frisvold (2017, March). The Pass-Through of Taxes on Sugar-Sweetened Beverages to Retail Prices: The Case of Berkeley, California. *Journal of Policy Analysis and Management 36*(2), 303–326.

CDC (2016, September). Cut Back on Sugary Drinks. http://www.cdc.gov/nutrition/data-statistics/sugar-sweetened-beverages-intake.html.

Crawford, G., R. Griffith, and A. Iaria (2017, October). Preference estimation with unobserved choice set heterogeneity using sufficient sets. *CEPR DP 11675*.

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature 47*(1), 87–122.

Currie, J., S. DellaVigna, E. Moretti, and V. Pathania (2010, August). The Effect of Fast Food Restaurants on Obesity and Weight Gain. *American Economic Journal: Economic Policy 2*(3), 32–63.

Cutler, D., E. Glaeser, and J. Shapiro (2003). Why Have Americans Become More Obese? *Journal of Economic Perspectives 17*(3), 93–118.

Delipalla, S. and O. O'Donnell (2001). Estimating tax incidence, market power and market conduct: The European cigarette industry. *International Journal of Industrial Organization 19*(6), 885–908.

Dhaene, G. and K. Jochmans (2015, July). Split-panel Jackknife Estimation of Fixed-effect Models. *Review of Economic Studies 82*(3), 991–1030.

DiMeglio, D. P. and R. D. Mattes (2000). Liquid versus solid carbohydrate: Effects on food intake and body weight. *International journal of obesity 24*(6), 794.

Dubois, P., R. Griffith, and A. Nevo (2014). Do Prices and Attributes Explain International Differences in Food Purchases? *American Economic Review 104*(3), 832–867.

England, P. H. (2018, March). National Diet and Nutrition Survey (NDNS). https://www.gov.uk/government/collections/national-diet-and-nutrition-survey.

Fabra, N. and M. Reguant (2014). Pass-Through of Emissions Costs in Electricity Markets. *American Economic Review 104*(9), 2872–2899.

Gilbert, D. T., M. J. Gill, and T. D. Wilson (2002, May). The Future Is Now: Temporal Correction in Affective Forecasting. *Organizational Behavior and Human Decision Processes 88*(1), 430–444.

Goeree, M. S. (2008). Limited Information and Advertising in the U.S. Personal Computer Industry. *Econometrica 76*(5), 1017–1074.

Goldberg, P. K. and R. Hellerstein (2013). A structural approach to identifying the sources of local currency price stability. *Review of Economic Studies 80*(1), 175–210.

Gortmaker, S. L., M. W. Long, and Y. Wang (2009, November). The negative impact of sugar-sweetened beverages on children's health. A Research Synthesis.

Grogger, J. (2015, May). Soda Taxes and the Prices of Sodas and Other Drinks: Evidence from Mexico. Working Paper 21197, National Bureau of Economic Research.

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.

Haavio, M. and K. Kotakorpi (2011, May). The political economy of sin taxes. *European Economic Review 55*(4), 575–594.

Hahn, J. and W. Newey (2004). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica 72*(4), 1295–1319.

Han, E. and L. M. Powell (2013, January). Consumption patterns of sugar sweetened beverages in the United States. *Journal of the Academy of Nutrition and Dietetics 113*(1), 43–53.

Harding, M., E. Leibtag, and M. F. Lovenheim (2012). The Heterogeneous Geographic and Socioeconomic Incidence of Cigarette Taxes: Evidence from Nielsen Homescan Data. *American Economic Journal: Economic Policy 4*(4), 169–198.

Haushofer, J. and E. Fehr (2014, May). On the psychology of poverty. *Science 344*(6186), 862–867.

Kaplan, G. and G. Menzio (2015). The morphology of price dispersion. *International Economic Review 56*(4), 1165–1206.

Kenkel, D. S. (2005). Are alcohol tax hikes fully passed through to prices? Evidence from Alaska. *American Economic Review: Papers & Proceedings 95*(2).

Lewbel, A. and K. Pendakur (2017). Unobserved Preference Heterogeneity in Demand Using Generalized Random Coefficients. *Journal of Political Economy 125*(4).

Lillard, D. R. and A. Sfekas (2013). Just passing through: The effect of the Master Settlement Agreement on estimated cigarette tax price pass-through. *Applied economics letters 20*(4), 353–357.

Lönnroth, K., B. G. Williams, S. Stadlin, E. Jaramillo, and C. Dye (2008). Alcohol use as a risk factor for tuberculosis–a systematic review. *BMC Public Health 8*(1).

Mani, A., S. Mullainathan, E. Shafir, and J. Zhao (2013, August). Poverty Impedes Cognitive Function. *Science 341*(6149), 976–980.

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision 8*(3), 229–254.

Mennella, J. A., N. K. Bobowski, and D. R. Reed (2016, June). The development of sweet taste: From biology to hedonics. *Reviews in Endocrine and Metabolic Disorders 17*(2), 171–178.

Nakamura, E. and D. Zerom (2010). Accounting for Incomplete Pass-Through. *Review of Economic Studies 77*(3), 1192–1230.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Neyman, J. and E. L. Scott (1948, January). Consistent Estimates Based on Partially Consistent Observations. *Econometrica 16*(1), 1.

Nielsen, S. J. and B. M. Popkin (2004, October). Changes in beverage intake between 1977 and 2001. *American Journal of Preventive Medicine 27*(3), 205–210.

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

Poterba, J. M. (1989). Lifetime Incidence and the Distributional Burden of Excise Taxes. *American Economic Review 79*(2), 325–330.

Public Health England (2015). *Sugar Reduction: The Evidence for Action.*

Rabin, M. (1998). Psychology and Economics. *Journal of Economic Literature 36*(1), 11–46.

Read, D. and B. Van Leeuwen (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational behavior and human decision processes 76*(2), 189–205.

Rehm, J., B. Taylor, S. Mohapatra, H. Irving, D. Baliunas, J. Patra, and M. Roerecke (2010). Alcohol as a risk factor for liver cirrhosis: A systematic review and meta-analysis. *Drug and alcohol review 29*(4), 437–445.

Sanders, B. (2016, April). Bernie Sanders Op-Ed: A Soda Tax Would Hurt Philly's Poor.

Small, K. A. and H. S. Rosen (1981). Applied Welfare Economics with Discrete Choice Models. *Econometrica 49*(1), 105–130.

Taylor, R., S. Kaplan, S. Villas-Boas, and K. Jung (2018). Soda Wars: Effect of a Soda Tax Election on Soda Purchases.

Train, K. E. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press.

Train, K. E. (2008). EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling 1*(1), 40–69.

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

WHO (2015). Sugars intake for adults and children.

Appendix

How well targeted are soda taxes?

Pierre Dubois, Rachel Griffith and Martin O'Connell

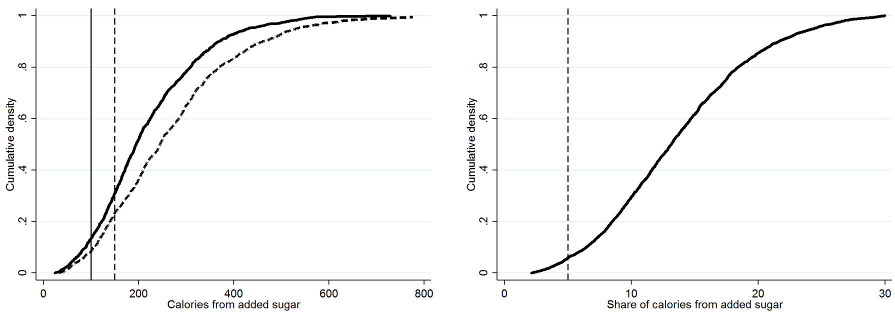
A Data appendix

A.1 Patterns of sugar consumption

In this appendix we use data from the National Diet and Nutrition Survey 2008-2011, which is an intake study of a representative sample of 3,073 UK adults and children. A.1 shows that most people consume more than the recommended amount of sugar. Figure A.2 shows that those people who consume the most sugar (as a share of calories) get a lot of the sugar they consume from soft drinks. These two facts motivate why soda taxes might be well targeted at reducing sugar consumption.

Figure A.1: *Sugar*

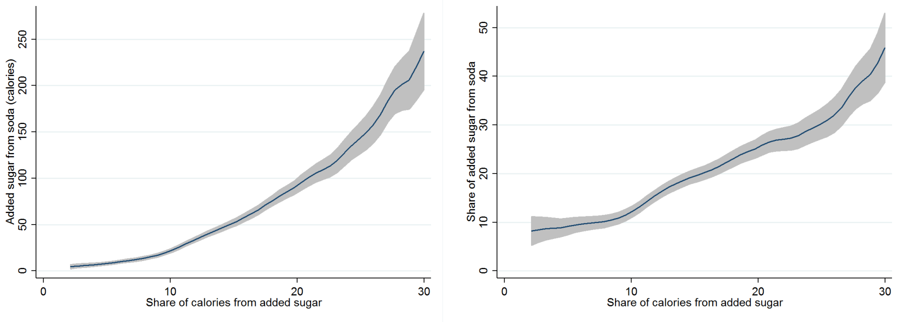
(a) Calories from sugar (b) % of calories from sugar



Notes: Numbers using National Diet and Nutrition Survey 2008-2011. Shaded areas denote 95% confidence intervals.

Figure A.2: *Sugar from soft drinks*

(a) Sugar calories from soft drinks (b) % of sugar calories from soft drinks

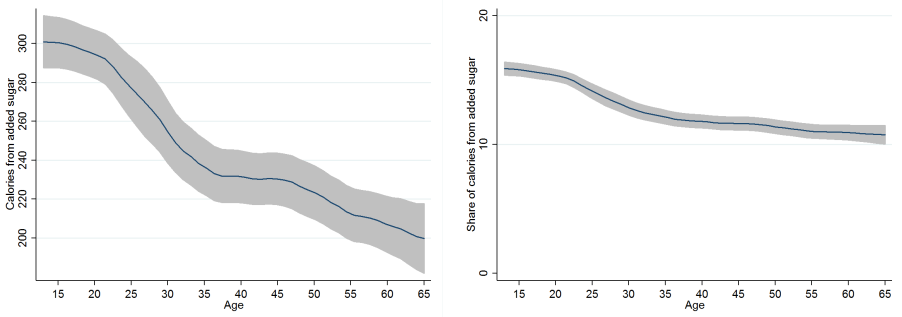


Notes: Numbers using National Diet and Nutrition Survey 2008-2011 for a representative sample of 3,073 UK adults and children. Shaded areas denote 95% confidence intervals.

Figure A.3 shows that consumption of sugar is higher amongst younger individuals. Figure A.4 shows that younger people get a higher share of their sugar from soft drinks.

Figure A.3: *Sugar, by age*

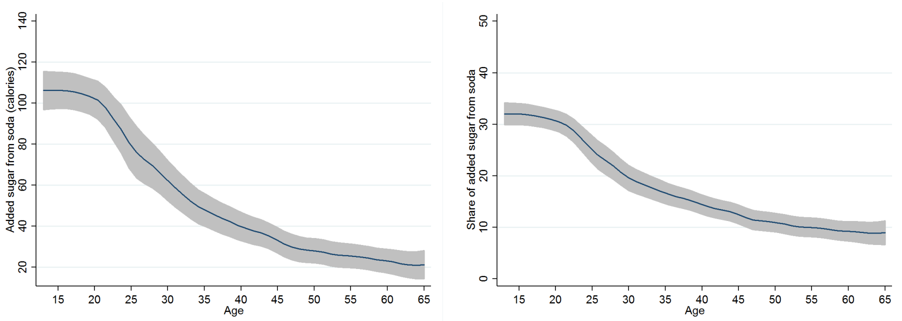
(a) Calories from sugar (b) % of calories from sugar



Notes: Numbers using National Diet and Nutrition Survey 2008-2011 for a representative sample of 3,073 UK adults and children. Shaded areas denote 95% confidence intervals.

Figure A.4: *Sugar from soft drinks, by age*

(a) Sugar calories from soft drinks (b) % of sugar calories from soft drinks



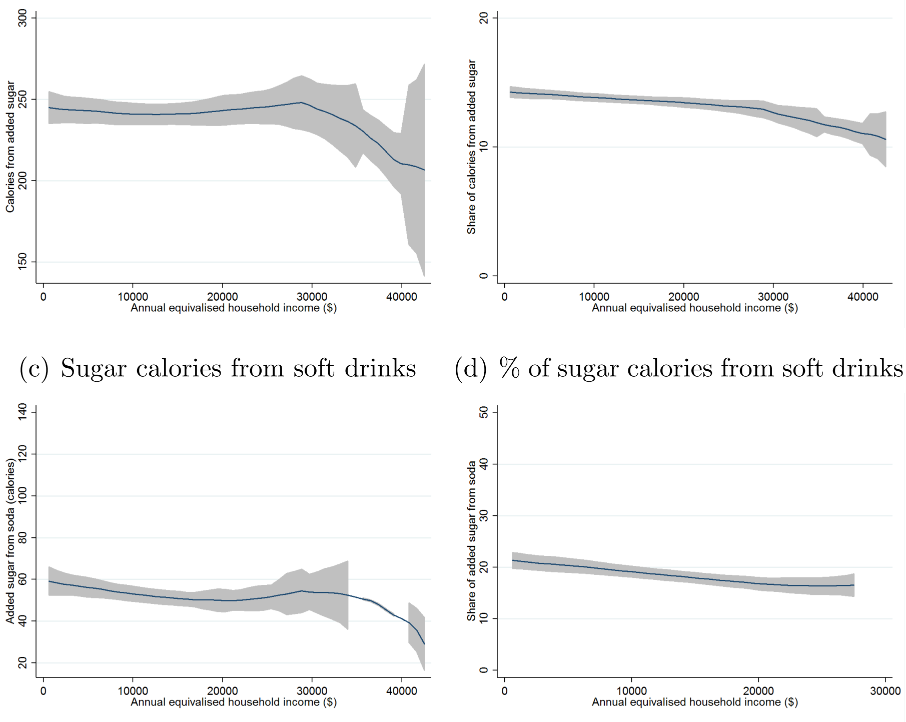
Notes: Numbers using National Diet and Nutrition Survey 2008-2011 for a representative sample of 3,073 UK adults and children. Shaded areas denote 95% confidence intervals.

Figure A.3 shows that relationship with income is less strong.

Figure A.5: *Sugar, by age*

(a) Calories from sugar (b) % of calories from sugar

(c) Sugar calories from soft drinks (d) % of sugar calories from soft drinks



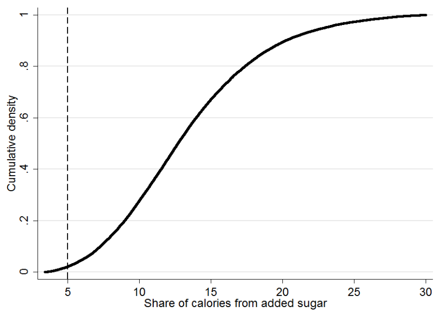
Notes: Numbers using National Diet and Nutrition Survey 2008-2011 for a representative sample of 3,073 UK adults and children. Shaded areas denote 95% confidence intervals.

A.2 Purchase patterns in US

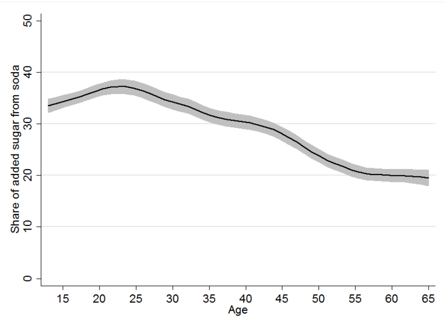
Using the National Health and Nutrition Examination Study over 2007-2014, a sample of 39,189 adults and children, we show similar patterns hold in the US. In Figure A.6 we use these data to show this. Panel (a) shows in the US, like the UK, the majority of the population get more calories from added sugar than the WHO guideline. Panels (b) and (c) show that, like in the UK, younger consumers get a higher share of added sugar from soft drinks and that people show consume the most added-sugar get a lot of that sugar from soft drinks.

Figure A.6: *Added sugar and soft drinks (US)*

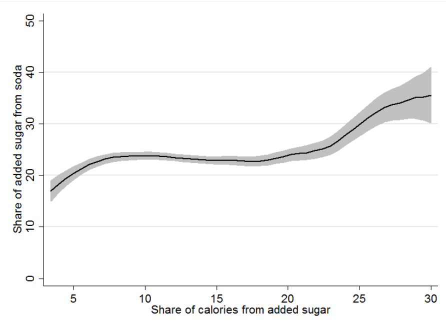
(a) Share of calories from added sugar



(b) Soft drinks and age



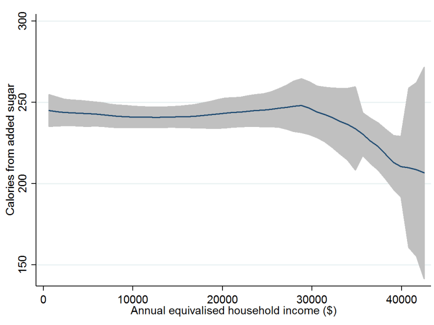
(c) Soft drinks and added sugar



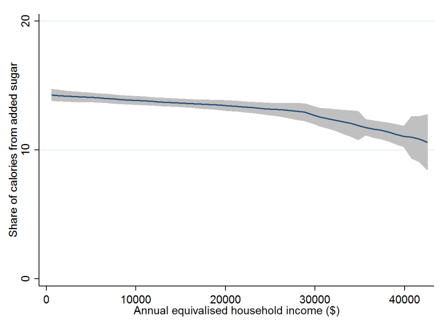
Notes: Numbers based on National Health and Nutrition Examination Study over 2007-2014. Vertical line in panel (a) denotes the WHO target of no more than 5% of calories from added sugar. Shaded areas in panels (b) and (c) denote 95% confidence intervals.

Figure A.7: *Sugar, by income*

(a) Calories



(b) % of calories



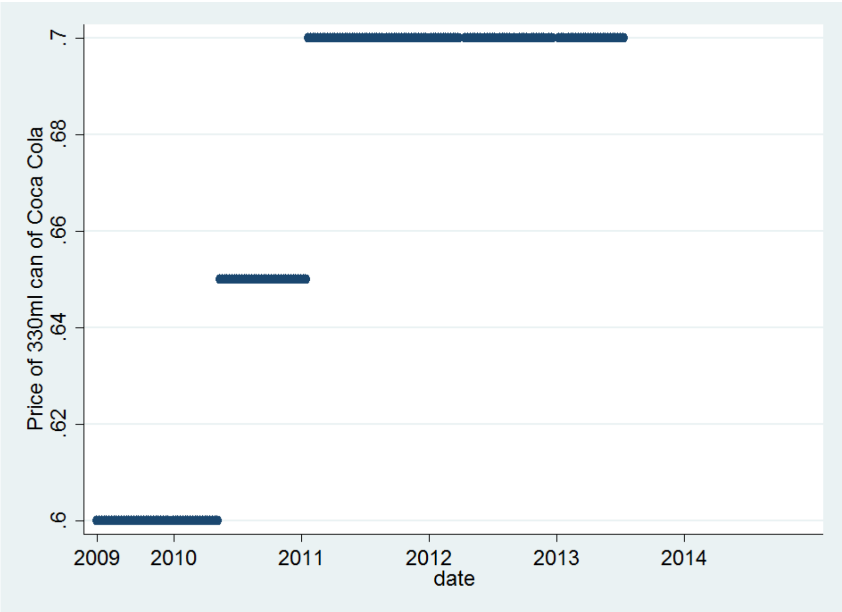
Notes: Numbers based on National Health and Nutrition Examination Study over 2007-2014.

A.3 Variation in prices

Product prices vary over time and across retail outlets. We compute the mean monthly price for each product in each retail outlet and use this in demand estimation. On each choice occasion we observe where an individual shops, we assume that this is independent of demand shocks, and we assume that the consumer faces the vector of prices for products in the retailer that we observe them shopping in. We exploit two sources of price variation, see Section 3.2.

In order to confirm that the variation originates from real variation over time within consumer and across stores in this Appendix we provide some additional description of individual level variation in transaction prices by given product over time and across stores by each individual. To be concrete, Figure A.8 shows an example of the raw data. This is a scatter plot of the 838 observed transaction prices for this one individual who has purchased a 330ml can of Coca Cola from a vending machine on a regular basis. We see that the price rose on two occasions, from 60p to 65p on 12 May 2010, and from 65p to 70p on 20 January 2011.

Figure A.8: *Observed transaction prices for one individual from a vending machine*



Notes: The figure shows all observed transaction prices for a single individual for purchases of 330ml can of coke from vending machines. .

This is obviously one example. To show the full variation in our data we describe the variation in observed transaction prices for each individual over time within a store. For each individual and product we compute the coefficient of variation of transaction prices within each retailer type. This captures variation in prices within retailer (if the consumers shop at different shops within the same retailer type, for example, visits different vending machines, then we will also pick up this variation, but this is small relative to within retailer time series variation).

Table A.1 shows quantiles of the individual consumer level coefficients of variation of prices within product and retailer (in the on-the-go segment of the market) over time. A coefficient of 0.10 means that the standard deviation is 10% of the average price.

The quantiles reported by product in Table A.1 show that for each product, more than half and sometimes more than 75% of the consumer-retailer level coefficient of variation of prices is strictly positive. We find some cases of zero variation, reflecting individuals with non-frequent purchasers of the product in the same retailer, so we only observe a few transaction prices. As prices will vary even when consumers do not purchase, the share of consumer-retailer observations exposed to true zero price variation is necessarily even lower. The table also shows that most of the consumers-retailer observations are exposed to substantial variation in prices before aggregation.

Table A.1: *Variation over time: distribution of individual-product-retailer coefficient of variation of transaction prices*

Product	Mean	Q10	Q25	Q50	Q75	Q90
Coca Cola 330	.102	0	.0139	.0812	.156	.236
Coca Cola 500	.109	0	.0401	.0897	.159	.236
Dr Pepper 330	.0886	0	0	.0456	.118	.26
Dr Pepper 500	.0924	0	.00905	.0716	.137	.223
Fanta 330	.0811	0	0	.0454	.136	.231
Fanta 500	.0878	0	.00652	.0661	.131	.216
Cherry Coke 330	.0818	0	0	.0527	.121	.22
Cherry Coke 500	.078	0	.00711	.0604	.117	.195
Oasis 500	.0931	0	.019	.0787	.138	.209
Pepsi 330	.131	0	0	.0867	.202	.349
Pepsi 500	.151	0	.0444	.118	.232	.345
Lucozade Energy 380	.104	0	0	.0823	.164	.258
Lucozade Energy 500	.0966	0	.016	.0792	.145	.221
Ribena 288	.0992	0	0	.0741	.161	.241
Ribena 500	.0955	0	0	.072	.139	.228
Other	.215	0	.0773	.183	.325	.458
Other Diet	.174	0	.0345	.135	.263	.404
Fruit juice	.23	0	.0842	.212	.346	.474
Flavoured milk	.189	0	.0351	.155	.302	.42
Fruit water	.103	0	0	.0679	.169	.254
Water	.218	0	.0621	.184	.326	.481

Notes: Quantiles of the distribution across individuals of the coefficient of variation over time of individual transaction prices, that is the ratio of the standard deviation over time of prices paid for that product by an individual consumer within a retailer divided by the mean price paid by that individual for that product in that retailer.

To quantify the variation that each consumer faces future, in Table A.2 we show the share of retailer-product level time series that an individual consumer faces are zero. If prices of one product do not vary over time, so the coefficient of variation for that retailer-product are zero, the consumer could still face relative price variation if the prices of other products vary. The table shows that for the majority of individuals (59%) all of the retailer-product price vectors show variation over time, and for 95% less than 10% of the price series they face have no variation. For no individual do more than half of the price series that they face not vary over time.

Table A.2: *% of price series where an individual faces no variation*

% price series with no variation over time	Frequence	%	Cumulative %
0%	1,441	59.32	59.32
5%	667	27.46	86.78
10%	219	9.02	95.80
14%	62	2.55	98.35
19%	25	1.03	99.38
24%	11	0.45	99.84
29%	2	0.08	99.92
33%	1	0.04	99.96
48%	1	0.04	100.00
All	2,429	100.00	

Notes: .

Tables A.3 and A.4 show the same statistics for the coefficient of variation using the average monthly prices that we use for estimation of the demand model (based on chosen options and all options in individuals' choice sets). While the means are lower than in Table A.1, it shows that there is still considerable variation over time within individual, product and retailer. In estimation we control for brand time effects and other covariates that will absorb some of this variation.

Table A.3: *Variation over time: distribution of individual-product-retailer coefficient of variation of average monthly prices of products purchased*

Product	Mean	Q10	Q25	Q50	Q75	Q90
Coca Cola 330	.0335	.00307	.0114	.0263	.0478	.0698
Coca Cola 500	.0343	.00343	.0116	.0313	.0519	.0686
Dr Pepper 330	.048	.00181	.0199	.0431	.0686	.098
Dr Pepper 500	.0372	.00294	.0103	.0294	.0528	.0836
Fanta 330	.0486	.00246	.0159	.0396	.0653	.103
Fanta 500	.03	.00152	.00826	.0247	.0446	.0653
Cherry Coke 330	.0399	.011	.0201	.036	.0541	.0757
Cherry Coke 500	.0284	.00213	.00759	.021	.0424	.0675
Oasis 500	.0304	.00226	.00931	.0237	.047	.0657
Pepsi 330	.057	.00572	.0219	.048	.0795	.117
Pepsi 500	.0313	.00365	.0141	.0301	.0426	.0599
Lucozade Energy 380	.031	.00141	.0112	.0252	.0463	.0646
Lucozade Energy 500	.0327	.00189	.0102	.0287	.0489	.0709
Ribena 288	.0426	.00485	.018	.0377	.061	.0816
Ribena 500	.0367	.00167	.01	.0266	.0492	.085
Other	.0473	.00643	.02	.0416	.0645	.0946
Other Diet	.055	.00572	.0211	.0519	.0786	.109
Fruit juice	.064	.00654	.0252	.0597	.0937	.125
Flavoured milk	.0579	.00242	.0177	.047	.0869	.129
Fruit water	.0385	.00112	.0141	.0322	.0527	.0819
Water	.0526	.00564	.0195	.0458	.077	.106

Notes: Quantiles of the distribution across individuals of coefficient of variation across stores of smoothed prices, that is the ratio of the standard deviation over time of prices paid for that product by a consumer divided by the mean price of that product.

Table A.4: *Variation over time: distribution of individual-product-retailer coefficient of variation of average monthly prices of all options*

Product	Mean	Q10	Q25	Q50	Q75	Q90
Coca Cola 330	.0442	.00453	.0211	.0429	.0651	.0822
Coca Cola 500	.0361	.00236	.0154	.037	.0551	.0658
Dr Pepper 330	.0654	.0251	.0448	.0594	.0885	.107
Dr Pepper 500	.0483	.004	.0243	.0453	.0712	.0933
Fanta 330	.0516	.0000187	.00148	.0526	.0727	.114
Fanta 500	.0346	.00225	.0177	.0339	.0499	.0645
Cherry Coke 330	.0527	.023	.0361	.0503	.0646	.0863
Cherry Coke 500	.0337	.00327	.0153	.0316	.051	.0655
Oasis 500	.0353	.00365	.0171	.0363	.0513	.0635
Pepsi 330	.089	.00582	.0374	.07	.128	.206
Pepsi 500	.0337	.00409	.0221	.0347	.0457	.0574
Lucozade Energy 380	.0449	.0162	.0306	.0453	.0596	.0695
Lucozade Energy 500	.0426	.0139	.0277	.0435	.0576	.068
Ribena 288	.0543	.0201	.0347	.0543	.0735	.0874
Ribena 500	.0436	.00387	.0256	.039	.0584	.0825
Other	.0586	.0189	.0357	.0521	.0797	.106
Other Diet	.0703	.0307	.0518	.0694	.0871	.111
Fruit juice	.0756	.0256	.0467	.0735	.102	.126
Flavoured milk	.0802	.0242	.0446	.0764	.106	.139
Fruit water	.0606	.022	.0354	.0559	.0811	.109
Water	.0632	.0206	.0392	.0617	.0831	.107

Notes: Quantiles of the distribution across individuals of coefficient of variation across stores of smoothed prices, that is the ratio of the standard deviation over time of prices paid for that product by a consumer divided by the mean price of that product.

We exploit the fact that individuals face different vectors of prices when they (exogenously) purchase from a different retailer. Table A.5 gives an example of this price variation for the same individual shown in Figure A.8. On 25 June 2009 we observe this individual purchasing a 330ml can of Coca Cola from a vending machine. On 10 October 2009 we observe this same individual purchasing a 380ml bottle of Lucozade Energy drink from a small corner store. Comparing the vectors or prices they faced the can of Coca Cola was relatively cheaper on 25 June (Coke was 58% the price of Lucozade) than on 10 October (when Coke was 60% of the price of Lucozade). We include time-vary brand effects and retailer-brand effects (along with other covariates) which absorb some of this variation, but there remains considerable variation.

Table A.5: *Price vectors for one individual on two different days*

Product	Vending machine 25 June 2009	Corner store 10 October 2009
Coca Cola 330	**0.60**	0.54
Coca Cola 500	1.09	1.03
Dr Pepper 330		0.57
Dr Pepper 500		0.95
Fanta 330	0.59	0.53
Fanta 500	1.05	1.09
Cherry Coke 500	1.00	1.01
Oasis 500	0.98	0.94
Pepsi 330	0.60	0.60
Pepsi 500	1.02	0.98
Lucozade Energy 380	1.03	**0.90**
Ribena 500		1.03
Other	1.03	1.02
Other Diet	1.30	1.35

Notes: .

As a final exercise we regress the monthly price on all covariates included in the demand model and describe the coefficient of variation for each individual as above. This is shown in Table A.6.

Table A.6: *Variation over time: distribution of individual-product-retailer coefficient of variation of residualised average monthly prices of all options*

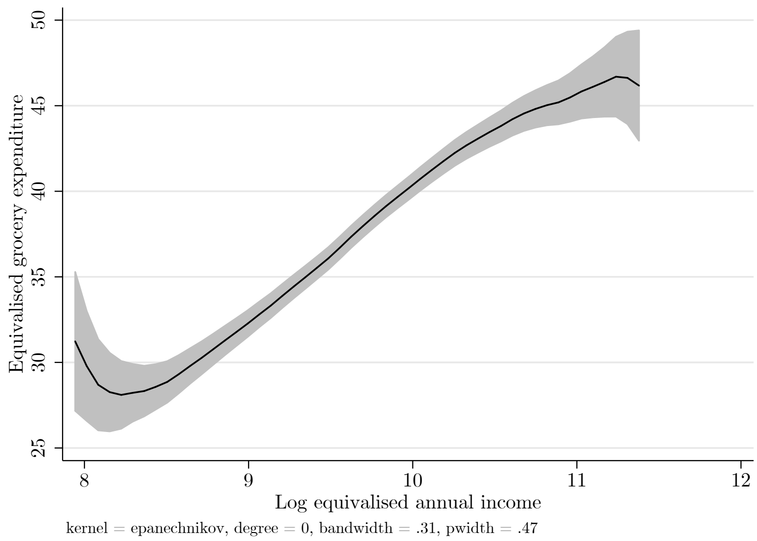
Product	Mean	Q10	Q25	Q50	Q75	Q90
Coca Cola 330	.0293	.0159	.0195	.0262	.0372	.0483
Coca Cola 500	.0234	.0157	.0189	.0225	.0278	.0326
Dr Pepper 330	.0426	.0153	.0289	.0419	.0566	.0677
Dr Pepper 500	.0413	.0195	.0243	.0376	.054	.0705
Fanta 330	.0311	.00565	.00971	.0326	.0433	.0571
Fanta 500	.0312	.0161	.0195	.0291	.04	.0504
Cherry Coke 330	.0319	.0129	.0216	.0312	.0387	.0488
Cherry Coke 500	.028	.0159	.0192	.0221	.0353	.0502
Oasis 500	.0243	.0153	.0178	.0204	.0306	.0393
Pepsi 330	.0375	.0182	.0227	.0305	.0464	.0727
Pepsi 500	.0335	.0185	.0231	.034	.0418	.0486
Lucozade Energy 380	.0389	.0165	.0266	.0391	.0497	.06
Lucozade Energy 500	.0406	.0156	.027	.041	.0537	.0652
Ribena 288	.0335	.0121	.0217	.0345	.045	.0526
Ribena 500	.0485	.0195	.0308	.0485	.0605	.0733
Other	.0637	.0227	.042	.0582	.0831	.112
Other Diet	.0933	.0385	.0659	.0934	.118	.147
Fruit juice	.0738	.0281	.0481	.0747	.0946	.113
Flavoured milk	.0686	.0279	.042	.0611	.0861	.126
Fruit water	.0454	.0158	.027	.042	.062	.0787
Water	.0327	.0105	.0166	.0288	.0406	.0599

Notes: Quantiles of the distribution across individuals of coefficient of variation across stores of smoothed prices, that is the ratio of the standard deviation over time of prices paid for that product by a consumer divided by the mean price of that product.

A.4 Relationship between equivalized expenditure and income

We use total household grocery expenditure to proxy for household income. The Living Costs and Food Survey (LCFS) is an expenditure survey that collects data on spending for a repeated cross-section of households. It also contains information on household income. Figure A.9 shows that there is a strong relationship between households' annual equivalized income and equivalized weekly grocery spending.

Figure A.9: *Relationship between household income and grocery expenditure*



Notes: Figure drawn using data on 4937 households in the Living Costs and Food Survey 2011. The horizontal axis shows logged equivalized annual income of the household, and the vertical axis shows equivalized weekly grocery expenditure. Figure trims the 5th and 95th percentiles of the logged equivalized annual income distribution. We equivalise using the standard OECD modified equivalence scale (see **?***).*

B Further details of demand estimates

In Table 3.1 and 3.3 we summarize moments of the distribution of estimated consumer specific preferences. Table B.1 and B.2 provide details of the estimated demographic group specific preference parameters. Table B.1 is for the on-the-go segment. For each demographic group the effect of temperature is negative – this indicates that during warm periods individuals are more likely than normal to purchase an alternative (non soft drink) product than a soft drink. The table also reports size effects (where the omitted group is a 500ml bottle) and brand effects. As we include a (consumer specific) soft drinks preferences, we omit both a Coca Cola brand effect and an outside option (bottled water) effect. Table B.2 shows estimates for the at-home segment. As with on-the-go purchases, higher temperature is associated with a shift from buying soft drinks towards alternative drinks. We allow for bottle vs. can and multi vs. single portions effects, as well as size fixed effects (not reported in the table) to affect utility. The table also reports brand effects, which we allow to vary across multi- and single-portion variants of the products.[27]

[27]For both the on-the-go and at-home segments we do not report the time varying brand effects or the retailer effects. These are available upon request.

Table B.1: *Demand model estimates: demographic specific preferences - on-the-go*

	Female - <40		*Female - 40+*		*Male - <40*		*Male - 40+*	
Variable	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Temperature	-0.0039	0.0042	-0.0018	0.0039	-0.0031	0.0041	-0.0102	0.0038
288ml carton	-1.0890	0.0454	-1.7468	0.0547	-2.1261	0.0493	-2.2692	0.0544
330ml can	-2.7766	0.0293	-2.7350	0.0295	-2.7673	0.0252	-2.9435	0.0275
380ml bottle	-0.2643	0.0384	-0.3738	0.0354	-0.5636	0.0294	-0.3636	0.0311
Dr Pepper	-1.5706	0.0468	-2.7651	0.0531	-1.6008	0.0445	-2.1560	0.0413
Fanta	-1.6302	0.0467	-1.9179	0.0486	-2.0054	0.0454	-1.6862	0.0396
Cherry Coke	-1.9810	0.0482	-2.5402	0.0518	-2.0021	0.0458	-1.8900	0.0405
Oasis	-1.5748	0.0469	-1.6399	0.0479	-1.5841	0.0449	-1.6256	0.0399
Pepsi	-1.3705	0.0520	-1.3766	0.0494	-1.4033	0.0482	-1.7350	0.0478
Lucozade	-1.7856	0.0672	-1.3013	0.0586	-1.2529	0.0570	-1.4363	0.0525
Ribena	-2.2299	0.0651	-2.0530	0.0574	-2.1882	0.0569	-2.1148	0.0521
Other soda	-0.7566	0.0569	-0.3382	0.0524	0.0370	0.0462	-0.3866	0.0446
Fruit juice	3.5904	0.0996	3.8352	0.0914	3.0910	0.1030	3.7545	0.0963
Flavoured milk	-0.9563	0.0914	-1.4490	0.0848	-1.1970	0.0972	-1.7485	0.0916
Flavoured water	-1.7236	0.0927	-1.7726	0.0844	-2.6850	0.1037	-2.2541	0.0929
Time-demographic-brand effects ($\xi_{d(i)b(j)t}$)					Yes			
Retailer-demographic-brand effects ($\zeta_{d(i)b(j)r}$)					Yes			

Notes: We estimate demand on a sample of 2,374 individuals who we observe on 361,863 on-the-go choice occasions. The table shows estimates are of the demographic specific preference parameters. Brand effects are shown for a baseline period. We allow these to vary across years and quarters.

Table B.2: *Demand model estimates: demographic specific preferences - at-home*

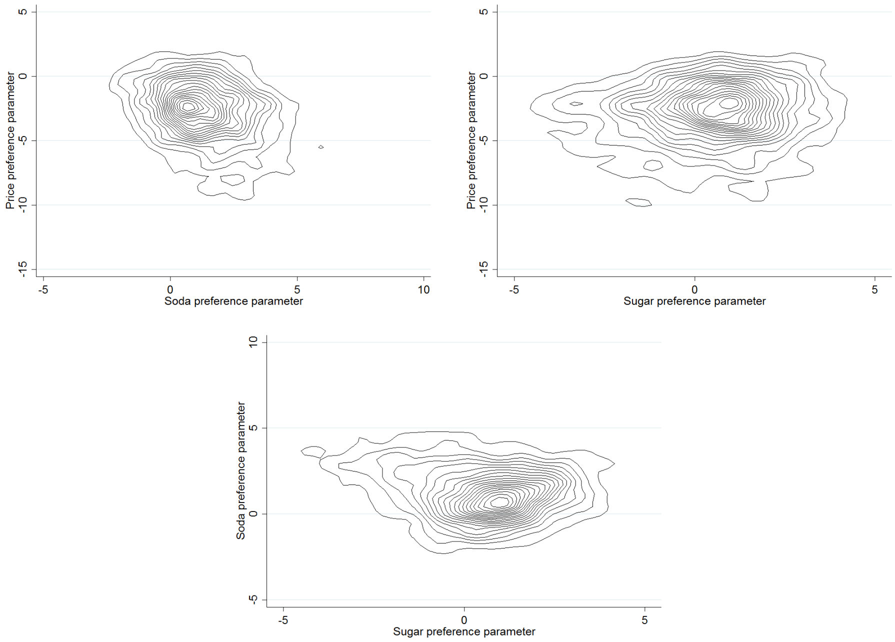
Variable	*No children High skill* Estimate	Standard error	*No children Low skill* Estimate	Standard error	*Pensioner* Estimate	Standard error	*Children High skill* Estimate	Standard error	*Children Low skill* Estimate	Standard error
Temperature	-0.0154	0.0034	-0.0202	0.0046	-0.0140	0.0060	-0.0163	0.0030	-0.0124	0.0038
Bottle	-0.5381	0.0533	-0.3967	0.0649	-1.2049	0.1153	-0.1330	0.0503	-0.0262	0.0586
Multi	0.6467	0.0399	0.8504	0.0508	0.6867	0.0754	0.5146	0.0391	0.9113	0.0454
Dr Pepper	-0.8396	0.1015	-0.1632	0.1189	-1.9871	0.3203	-1.0520	0.0987	-0.3353	0.1036
Fanta	-0.8499	0.1001	-0.2847	0.1192	-0.5397	0.2388	-1.0118	0.0944	-0.6571	0.1073
Cherry Coke	-2.3434	0.1155	-0.7415	0.1035	-3.5360	0.3644	-2.1864	0.0990	-1.3569	0.1044
Oasis	-0.5637	0.0930	-0.0131	0.1105	-0.0225	0.2257	-0.4299	0.0820	-0.3762	0.0990
Pepsi	-1.0908	0.0660	-0.6783	0.0789	-1.1230	0.1494	-1.3468	0.0633	-0.8733	0.0720
Lucozade	-0.4492	0.1034	0.2759	0.1228	0.6160	0.2121	-0.9169	0.1064	-0.1951	0.1113
Ribena	-2.1418	0.1732	-0.9797	0.1787	-1.4762	0.3467	-1.5003	0.1303	-1.2729	0.1542
Other soda	-6.1458	1.0045	-3.2747	0.3028	-3.4049	0.6095	-4.1219	0.3884	-4.9368	1.0053
Store brand soda	-9.3745	1.0102	-5.4311	0.3298	-6.8853	0.6488	-6.8693	0.4012	-6.9107	1.0122
Fruit juice	4.0229	0.0878	3.1303	0.1177	4.1080	0.1581	3.3660	0.0788	2.8604	0.0980
Flavoured milk	9.0774	1.0049	5.8969	0.3068	7.3214	0.6069	7.6291	0.3886	8.6367	1.0058
Flavoured water	5.4735	1.0188	2.4549	0.3519	4.9148	0.6315	4.5249	0.4025	4.7067	1.0221
Coke*single	-7.3220	1.0042	-3.4033	0.3009	-5.4663	0.6066	-5.2858	0.3878	-5.9317	1.0054
Dr Pepper*single	-8.3424	1.0078	-4.7342	0.3161	-6.7532	0.6690	-5.8229	0.3964	-7.0296	1.0085
Fanta*single	-8.2664	1.0074	-4.5395	0.3157	-7.2744	0.6263	-5.8772	0.3949	-6.9469	1.0087
Cherry Coke*single	-7.2033	1.0086	-4.7429	0.3083	-6.0316	0.6924	-5.6750	0.3959	-7.1402	1.0086
Pepsi*single	-6.7511	1.0042	-3.2255	0.3004	-5.1496	0.6062	-4.5223	0.3884	-5.7567	1.0053
Lucozade*single	-5.3940	1.0061	-2.4803	0.3118	-4.2290	0.6159	-3.6887	0.3947	-4.7742	1.0078
Ribena*single	-7.8957	1.0208	-4.3717	0.3551	-7.6408	0.7173	-4.9072	0.4073	-6.0474	1.0168
Other soda*single	-2.6475	0.0754	-1.6828	0.0962	-2.8996	0.1460	-2.3118	0.0708	-1.8817	0.0832
Fruit juice*single	-6.7158	1.0057	-3.9924	0.3087	-5.2681	0.6189	-5.3327	0.3917	-6.4837	1.0071
Flavoured milk*single	-6.9343	1.0197	-3.1554	0.3515	-6.9946	0.6420	-4.8586	0.4028	-5.6706	1.0226
Demographic specific size effects ($\delta^z_{d(i)}$)					Yes					
Time-demographic-brand effects ($\xi_{d(i)b(j)t}$)					Yes					
Retailer-demographic-brand effects ($\zeta_{d(i)b(j)r}$)					Yes					

Notes: We estimate demand on a sample of 3,314 households who we obseve on 302,383 at-home choice occasions. The table shows estimates are of the demographic specific preference parameters. Brand effects are shown for a baseline period. We allow these to vary across years and quarters.

B.1 Distributions of preference parameters

In Figure B.1 we plot contour plots of the bivariate preference distributions (based on the finite parts of the distribution). Figure B.2 shows how price and sugar preferences varies across the distribution of total equivalized grocery expenditure.

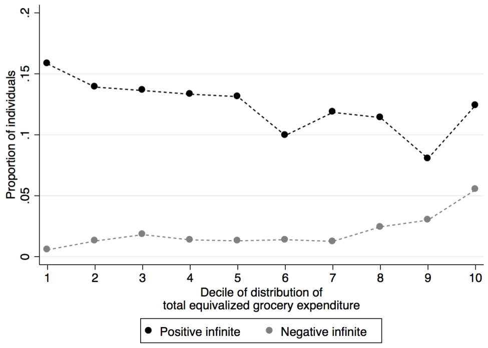
Figure B.1: *Bivariate distributions of consumer specific preference parameters*



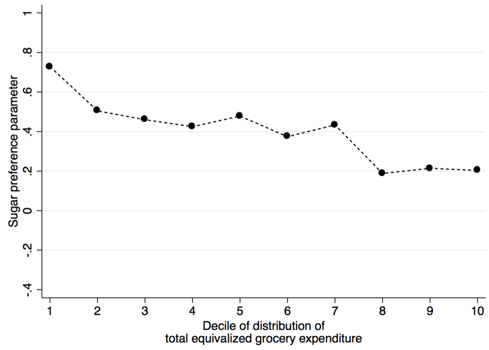
Notes: Distribution plots use consumers with finite preference parameters, those having infinite distaste for soft drinks or sugar are not included in this graph.

Figure B.2: *Preferences variation with equivalized expenditure*

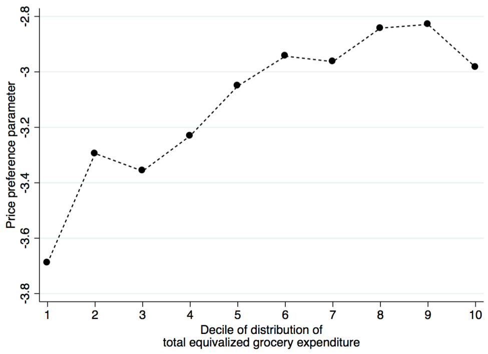
(a) infinite sugar preferences



(b) finite sugar preferences



(c) price preferences



(d) sugar-price correlation

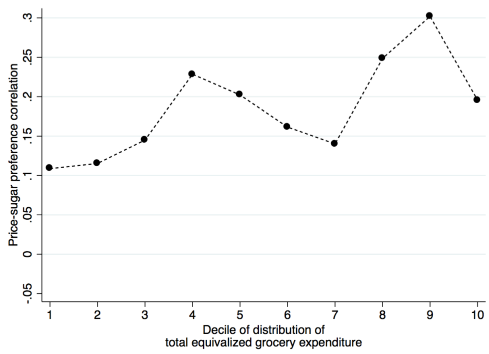


Figure shows how the share consumers with infinite sugar preferences, the mean of finite sugar preferences, the mean of price preferences and the correlation between sugar and price preferences vary deciles of the distribution of total annual equivalized grocery expenditure. 95% confidence intervals are shown by bars.

B.2 Price effects on demand

Table B.3: *Price Effects by product – on-the-go segment*

	Effect of 1% price increase on:				
	Own	cross demand for:			Total
	demand	sugary soft drinks	diet soft drinks	sugary alternatives	demand
Coca Cola 330	-2.91	0.16	0.08	0.07	0.02
Coca Cola 500	-2.27	0.41	0.18	0.25	-0.06
Coca Cola Diet 330	-2.90	0.06	0.29	0.02	0.02
Coca Cola Diet 500	-2.72	0.14	0.53	0.09	-0.05
Dr Pepper 330	-3.64	0.02	0.01	0.01	0.00
Dr Pepper 500	-2.73	0.07	0.03	0.04	-0.01
Dr Pepper Diet 500	-3.34	0.02	0.09	0.01	-0.01
Fanta 330	-3.58	0.03	0.01	0.01	0.00
Fanta 500	-2.62	0.07	0.03	0.05	-0.01
Fanta Diet 500	-3.23	0.03	0.10	0.02	-0.01
Cherry Coke 330	-3.60	0.02	0.01	0.01	0.00
Cherry Coke 500	-2.69	0.06	0.02	0.04	-0.01
Cherry Coke Diet 500	-3.25	0.02	0.07	0.01	-0.01
Oasis 500	-2.61	0.09	0.04	0.05	-0.01
Oasis Diet 500	-3.17	0.03	0.11	0.02	-0.01
Pepsi 330	-3.08	0.06	0.03	0.03	0.01
Pepsi 500	-2.76	0.17	0.08	0.10	-0.03
Pepsi Diet 330	-3.41	0.02	0.13	0.01	0.01
Pepsi Diet 500	-3.28	0.06	0.25	0.03	-0.03
Lucozade Energy 380	-2.84	0.10	0.05	0.06	0.01
Lucozade Energy 500	-2.57	0.07	0.03	0.05	-0.01
Ribena 288	-3.23	0.03	0.01	0.01	0.01
Ribena 500	-2.72	0.05	0.02	0.03	-0.01
Ribena Diet 500	-3.30	0.02	0.06	0.01	-0.01
Other	-2.27	0.51	0.22	0.28	-0.07
Other Diet	-2.76	0.09	0.30	0.05	-0.03
Fruit juice	-1.74	0.11	0.05	0.33	0.01
Flavoured milk	-2.37	0.03	0.01	0.07	-0.01
Fruit water	-2.42	0.02	0.01	0.04	-0.01

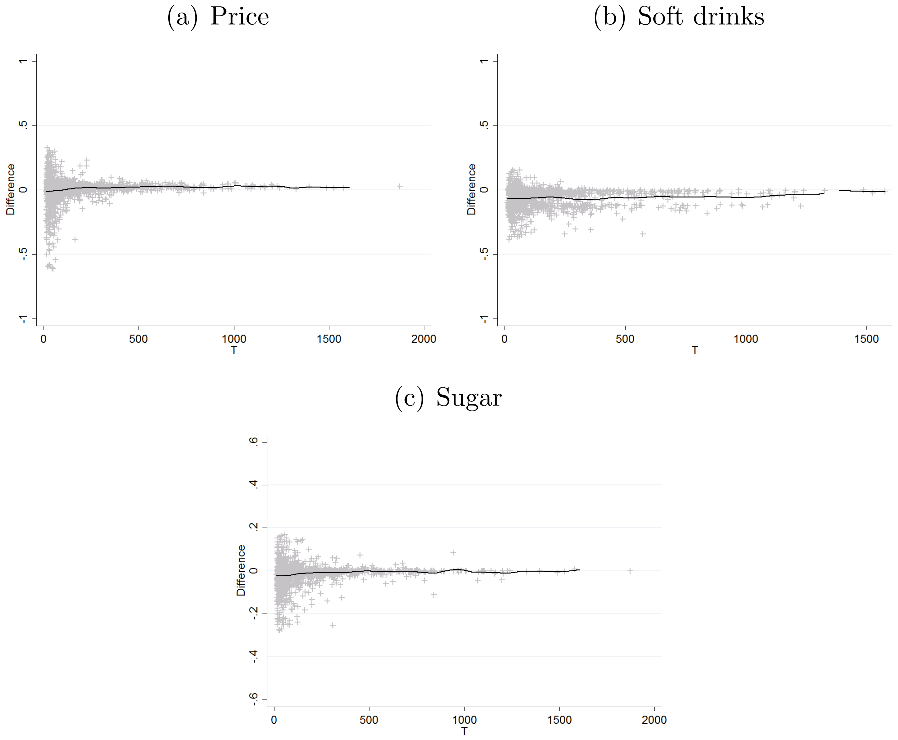
Notes: For each of the four products listed we compute the change in demand for that product, for alternative sugary and diet options and for total demand resulting from a 1% price increase. Numbers are means across time. 95% confidence bands in brackets.

B.3 Incidental parameters problem

Figures B.3, B.4 and B.5 show, for the price, soft drinks and sugar preference parameters, how the jackknife ($\widetilde{\theta}_{split}$) and the maximum likelihood estimate ($\widehat{\theta}$) relate to a) the time individuals are in the sample, b) age and c) total dietary sugar. They show no systematic relationship in the mean of ($\widetilde{\theta}_{split} - \widehat{\theta}$) with any of these variables, with the dispersion of ($\widetilde{\theta}_{split} - \widehat{\theta}$) falling in T.

Figures B.6 plots the distributions of price, soft drinks and sugar preference parameter estimates for both the estimators $\widehat{\theta}$ and $\widetilde{\theta}_{split}$, showing there is little difference in the distributions.

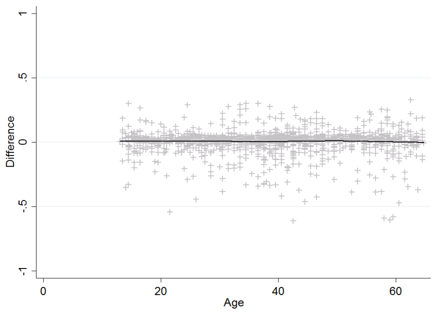
Figure B.3: *Relationship between bias and time in sample*



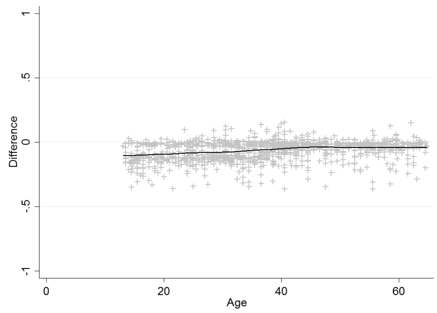
Notes: Marks represent consumer level differences. Lines are local polynomial regressions.

Figure B.4: *Relationship between bias and age*

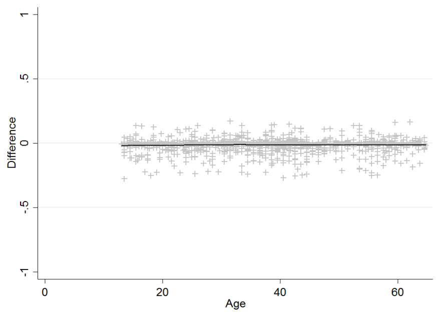
(a) Price



(b) Soft drinks



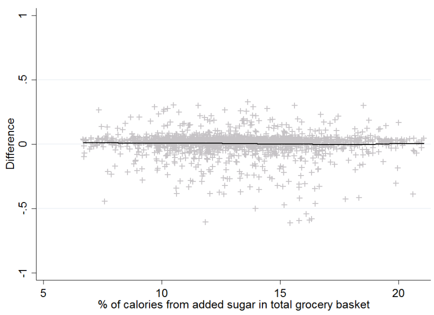
(c) Sugar



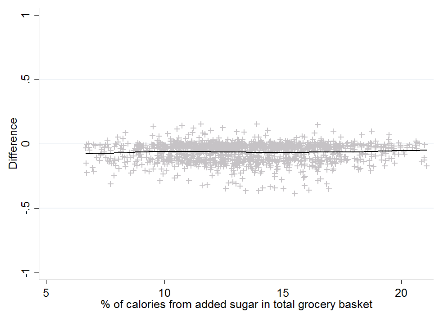
Notes: Marks represent consumer level differences. Lines are local polynomial regressions.

Figure B.5: *Relationship between bias and dietary sugar*

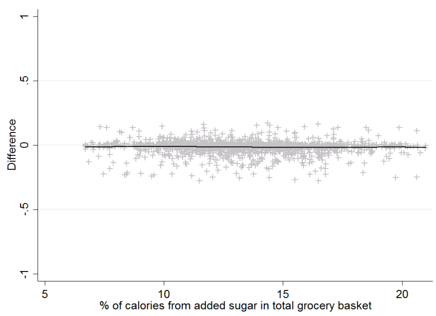
(a) Price



(b) Soft drinks



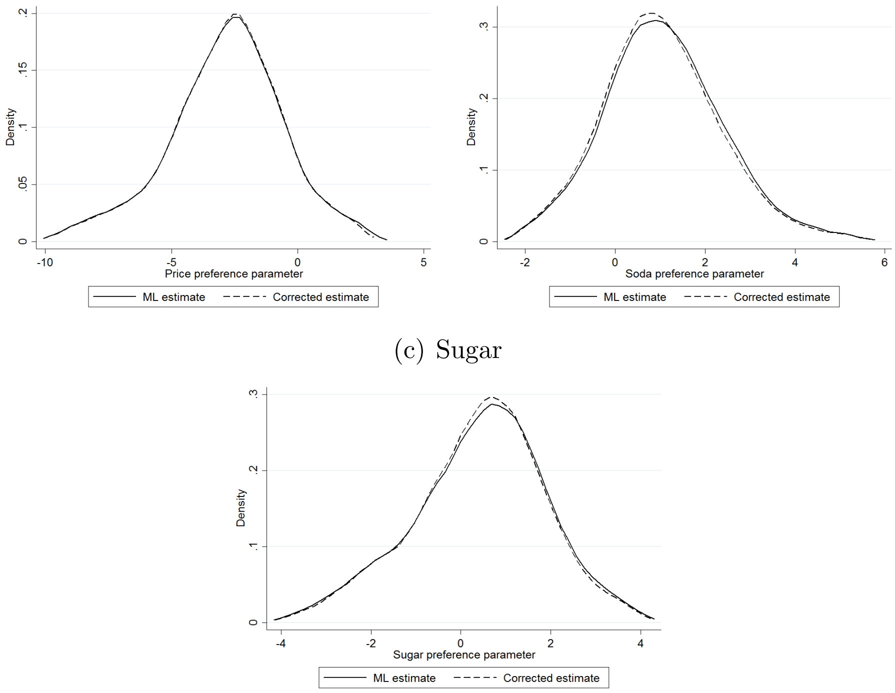
(c) Sugar



Notes: Marks represent consumer level differences. Lines are local polynomial regressions.

Figure B.6: *Preference parameter distribution*

(a) Price (b) Soft drinks (c) Sugar



Notes: Lines are kernel density estimates.

C An alternative soda tax

The paper focuses on the impact of a soda tax levied only on sugary soft drinks. We also simulate the impact of a soda tax incidence on all soft drinks products (both regular and diet); this tax takes the form

$$p_{jm} = \begin{cases} \tilde{p}_{jm} + \pi l_j & \forall j \in \Omega_{ws} \bigcup \Omega_{wn} \\ \tilde{p}_{jm} & \forall j \in \Omega_{as} \bigcup \Omega_{an}. \end{cases}$$

Here we refer to this as a broad soda tax and the tax we focus on in the main paper as a sugary soda tax. We simulate the same rate for the broad soda tax as for the sugary soda tax (25 pence per liter) using the same supply side model estimates in the first step and conducting the counterfactual simulation of pass-through of this tax to consumer prices.

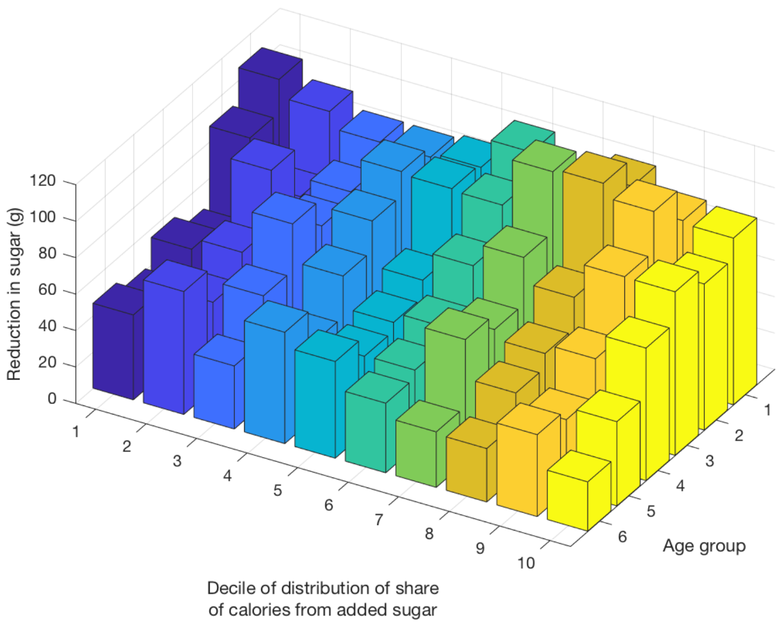
Table C.1 summarizes the impact of the broad soda tax on equilibrium prices and market shares (it contains analogous information to Table 4.1 in the main paper). The main difference between a tax levied on only sugary soft drinks and one levied on all soft drinks is that the latter leads to prices increases for diet products (that on average are similar to those for sugary products). The result is that the broad soda tax leads to a much smaller reduction in demand for sugary soft drinks and a fall (rather than increase) in demand for diet soft drinks (relative to the sugary soda tax). Figure C.1 shows that a broad soda tax does achieve larger reductions in sugar among the young than the old, but fails to achieve relatively large reductions among those with high total dietary sugar.

Table C.1: *Effects of "broad" soda tax*

	Tax (pence)	Δ price (pence)	Δ share (p.p.)
Sugary soda	10.65	11.44	-2.43
Diet soda	11.65	12.40	-1.35
Sugary alternatives	0.00	0.00	1.28
Outside option	0.00	0.00	2.49

Notes: Numbers are means across products.

Figure C.1: *Reductions in sugar by age and total dietary sugar*



Notes: Numbers are for the on-the-go segment. Figure shows variation in the reduction in sugar conditional on being a soft drinks purchaser. Age groups are 1=<22, 2=22-30, 3=31-40, 4=41-50, 5=51-60, 6=60+.

D Substitution to food

The choice model we outline in Section 3 captures consumer choice between drink products $j = \{0, 1, ..., J\} = \Omega_{\mathcal{D}}$. The drink products comprise water $j = 0$, soft drinks, $j = \{1, ..., j'\} = \Omega_a$ and juice $j = \{j' + 1, ..., J\} = \Omega_n$ The expected utility to the consumer of purchasing a drink is:

$$\begin{aligned}
E_{\epsilon_{ijt}}\left[\max_{j\in\Omega_{\mathcal{D}}} U_{ijt}\right] =& \ln\Bigg(\exp\left(\xi_{d(i)0t} + \zeta_{d(i)0t}\right) + \sum_{j\in\Omega_a\cup\Omega_n} \exp\left(\alpha_i p_{jrt} + \beta_i s_j + \right.\\
&\left.\gamma_i w_j + \delta^z_{d(i)} z_j + \delta^h_{d(i)} h_{c(i)t} + \xi_{d(i)b(j)t} + \zeta_{d(i)b(j)r}\right)\Bigg)\\
\equiv& W_{i\mathcal{D}t}.
\end{aligned}$$

Consider a first stage decision in which the consumer chooses between options $k = \{\varnothing, 1, ..., K, \mathcal{D}\}$, where $k = \varnothing$ denotes the outside option of a non-sugar snack, $k = \{1, ..., K\} = \Omega_c$ indexes chocolate products and $k = \mathcal{D}$ indexes choosing a drink. Suppose utility from these options takes the form:

$$\begin{aligned}
V_{i\varnothing t} &= \varepsilon_{i\varnothing t}\\
V_{ikt} &= \mu_c + W_{ikt} + \varepsilon_{ikt} \quad \text{for all } k \in \Omega_c\\
V_{i\mathcal{D}t} &= \mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t} + \varepsilon_{i\mathcal{D}t},
\end{aligned}$$

where

$$W_{ikt} = \alpha_i p_{krt} + \beta_i s_k + \vartheta_{b(k)}$$

and $(\varepsilon_{i0t}, \varepsilon_{i1t}, ..., \varepsilon_{iKt}, \varepsilon_{i\mathcal{D}t})$ are distributed i.i.d. extreme value. Note the nesting of the errors terms – consumers get a draw of first stage error terms $\boldsymbol{\varepsilon}$ and if they choose $k = \mathcal{D}$, they get a draw of second stage errors, $\boldsymbol{\epsilon}$, when selecting what drink product to choose. These idiosyncratic shocks are sequentially observed.

This first stage choice probabilities are:

$$\begin{aligned}
P_{it}(k = 0) =& \frac{1}{1 + \sum_{k'\in\Omega_c} \exp\left(\mu_c + W_{ik't}\right) + \exp\left(\mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t}\right)}\\
P_{it}(k = \tilde{k}) =& \frac{\exp\left(\mu_c + W_{i\tilde{k}t}\right)}{1 + \sum_{k'\in\Omega_c} \exp\left(\mu_c + W_{ik't}\right) + \exp\left(\mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t}\right)} \qquad \text{for all } \tilde{k} \in \Omega_c\\
P_{it}(k = \mathcal{D}) =& \frac{\exp\left(\mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t}\right)}{1 + \sum_{k'\in\Omega_c} \exp\left(\mu_c + W_{ik't}\right) + \exp\left(\mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t}\right)}.
\end{aligned}$$

The second stage drinks choice model allows us to identify the drinks inclusive value, $W_{i\mathcal{D}t}$, and the preference parameters (α_i, β_i) among all other drinks demand parameters. Let Ω_c^B denote the set of chocolate brands and ω_b be the set of chocolate products that belong to brand b. The second stage model also enables us to identify the chocolate brand indices as:

$$z_{ibt} = \ln \sum_{k \in \omega_b} \exp\left[\alpha_i p_{krt} + \beta_i s_k\right].$$

Note that

$$\begin{aligned}
\sum_{k \in \Omega_c} \exp\left(\mu_c + W_{ikt}\right) &= \sum_{b \in \Omega_c^B} \sum_{k \in \omega_b} \exp\left(\mu_c + W_{ikt}\right) \\
&= \sum_{b \in \Omega_c^B} \sum_{k \in \omega_b} \exp\left(\mu_c + \left[\alpha_i p_{krt} + \beta_i s_k + \vartheta_{b(k)}\right]\right) \\
&= \sum_{b \in \Omega_c^B} \exp\left(\tilde{\vartheta}_b + z_{ibt}\right),
\end{aligned}$$

where $\tilde{\vartheta}_b = \mu_c + \vartheta_b$ so that the first stage purchase probabilities can be written:

$$\begin{aligned}
P_{it}(k=0) =& \frac{1}{1 + \sum_{b' \in \Omega_c^B} \exp\left(\tilde{\vartheta}_{b'} + z_{ib't}\right) + \exp\left(\mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t}\right)} \\
P_{it}(k \in \omega_b) =& \frac{\exp\left(\tilde{\vartheta}_b + z_{ibt}\right)}{1 + \sum_{b' \in \Omega_c^B} \exp\left(\tilde{\vartheta}_{b'} + z_{ib't}\right) + \exp\left(\mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t}\right)} \quad \text{for all } b \in \Omega_c^b \\
P_{it}(k = \mathcal{D}) =& \frac{\exp\left(\mu_{i\mathcal{D}} + \psi_{\mathcal{D}} W_{i\mathcal{D}t}\right)}{1 + \sum_{b' \in \Omega_c^B} \exp\left(\tilde{\vartheta}_{b'} + z_{ib't}\right) + \exp\left(\mu_{i\mathcal{D}} + \psi_{i\mathcal{D}} W_{i\mathcal{D}t}\right)}.
\end{aligned}$$

Given identified parameters from the second stage and data on decisions consumers make over purchases of chocolate products, drinks or other snacks, the first stage choice model allows us to identify the remaining parameters $\tilde{\boldsymbol{\vartheta}} = (\tilde{\vartheta}_1, ..., \tilde{\vartheta}_B)'$, $\mu_{i\mathcal{D}}$ and $\psi_{i\mathcal{D}}$.

We allow for heterogeneity in the parameters $\mu_{i\mathcal{D}}$ and $\psi_{i\mathcal{D}}$ across age groups. Table D.1 shows estimates of these parameters.

Table D.1: *Upper stage model estimates*

	$\hat{\mu}_{iD}$		$\hat{\psi}_{iD}$	
	Estimate	Standard error	Estimate	Standard error
<22	1.3132	0.0164	0.4498	0.0043
22-30	1.5677	0.0116	0.4024	0.0034
31-40	1.4522	0.0093	0.3726	0.0025
41-50	1.1361	0.0098	0.4805	0.0027
51-60	1.2328	0.0112	0.5070	0.0028
60+	1.5641	0.0182	0.4347	0.0056

Notes: Estimates based on sample of 324,818 choice occasions. Chocolate brand effects were also estimated.