The impact of taxation and signposting on diet: an online field study with breakfast cereals and soft drinks

DANIEL JOHN ZIZZO, MELANIE PARRAVANO, RYOTA NAKAMURA, SUZANNA FORWOOD AND MARC SUHRCKE*

We present a large scale study where a nationally representative sample of 1,000 participants were asked to make real purchases within an online supermarket platform. The study captured the effect of price changes, and of the signposting of such changes, for breakfast cereals and soft drinks. We find that such taxes are an effective means of altering food purchasing, with a 20% rate being sufficient to make a significant impact if (and only if) the tax is signposted. Signposting represents a complementary "nudge" policy that could enhance the impact of the tax, though its effectiveness depends on the product category.

Keywords: taxes, signposting, healthy diet, nudges, public health.
JEL classification numbers: C93, D12, H31, I12.

* Zizzo: University of Queensland; corresponding author. E-mail: d.zizzo@uq.edu.au. Parravano: BHRU and Newcastle University. Nakamura: Hitotsubashi University. Forwood: Anglia Ruskin University. Suhrcke: University of York and Luxembourg Institute of Socio-economic Research. This report is independent research commissioned and funded by the Department of Health Policy Research Programme (Policy Research Unit in Behaviour and Health (PR-UN-0409-10109)). The views expressed in this publication are those of the authors and not necessarily those of the Department of Health. We also gratefully acknowledge the support of the UKCRC Centre for Diet and Physical Activity Research (CEDAR) and the ESRC (Network for Integrated Behavioural Science, ES/K002201/1). We thank Susan Jebb, Dan Hamermesh, Melanie Lührmann, Bruce Lyons, Theresa Marteau, Peter Moffatt, Moritz Mosenhauer, Vicki Pyne, Martin Roland, Richard Smith, Robert Sugden, Lise Vesterlund, the Co-Editors handling the manuscript, two anonymous reviewers and participants to presentations at Cambridge, East Anglia, Essen, Glasgow, Malaga, two Newcastle seminars, Nottingham, Reading, Royal Holloway, Toulouse, Queensland, Washington DC (Southern Economic Association), Wageningen (SABE/IAREP) and York for relevant encouragement and advice. We also thank Kajsa Berg and Jia Liu for research assistance. The usual disclaimer applies. Data collection took place while Zizzo and Parravano were based at the University of East Anglia, and the study protocol received approval from the University of

East Anglia, School of Economics Ethics Committee (Application No: 0170). Online appendices and additional files are provided electronically as supplementary material.

1. Introduction

This paper presents a large scale online experiment with a nationally representative sample from the United Kingdom to look at the effect of price changes, and ways in which these price changes are signposted, on healthy diet choice in relation to soft drinks and breakfast cereals (cereals in the rest of the paper). Experimental participants used real budgets and had an opportunity to buy real groceries online which were then delivered to their door. The groceries were from a list of almost 1,000 products offered by Tesco, the largest U.K. supermarket retailer, and checks were made for any effect of online purchases on later purchases to better identify the net effect of our potential policy interventions. We use a purpose-built online supermarket that is carefully modelled on a major real world online supermarket (Tesco.com; for further details about the online supermarket platform, see section 3). The platform enables us to examine shoppers' purchases in a setting which is close to a real online supermarket, while fully preserving experimental control to help ensure interpretability of our empirical findings. We model taxes of 20% or 40%, either simply as a price change or minimally signposted with the amount of the tax, marked with T, next to the baseline price.

There has been considerable recent policy interest in the potential use of fiscal policy to influence dietary behavior (Mytton et al., 2012), for example in the form of a 'fat tax' or of a 'sugar tax' as already trialed in different forms in Australia, Chile, Denmark, France, Hungary, Mexico and Norway (Cornelsen and Carreido 2015; Public Health England, 2015) and since April 2018 also in the UK (UK Government, 2016). The WHO (2016) has recently recommended the introduction of a 20% tax on soft drinks worldwide – a conclusion that is (ex post) supported by the findings of a recent systematic review of taxation to promote health behaviors not only in terms of diet (Wright et al., 2017). There is in principle a defendable justification for public policy to correct what may be conceived as a market failure associated with the external costs arising from obesity (Cawley and Frisvold, 2017). A public policy response may also be justified on the basis of internalities, in the sense that at least a share of the health consequences incurred are not taken into account by the individual when making his or her food choices (e.g., Griffith et al, 2018).

Our study is useful and innovative in several respects against the existing literature. In terms of price interventions, as discussed in the literature review in the next section, the evidence from the small number of policy impact evaluations is surprisingly mixed, and the limitations of epidemiological modelling studies and econometric analysis make a complementary experimental approach useful. Unfortunately, as reviewed in the next section, the existing experimental evidence on price interventions is also mixed (e.g. Epstein et al., 2012; Shemilt et al., 2013) and has limitations in terms of very small number of studies properly incentivized, in terms of structuring the shopping task as an artificial assignment and in terms of lack of control for compensation effects between shopping in and outside the laboratory; almost all of it is also with students or other convenience samples. These limitations largely affect the only one previous and important study (Chetty et al., 2009) that combines price with signposting in a diet context; their signposting is also implemented in a different way than ours (see section 2).

In our study budgets and purchases are for real, i.e. participants are given budget and groceries (cereals and soft drinks) that will be delivered via a real supermarket to their doorstep. We also restrict the sample of participants to people that have a latent demand for the product categories being considered, in that they have consumed cereals and soft drinks at least once in the last month. This avoids any power reduction and estimation bias by including people who do not generally buy either, but who may spuriously feel they have to in an experimental setting. We consider a nationally representative, large sample of consumers (of our two product categories), the largest we are aware of in terms of tax studies on diet choices; and we have an intertemporal if admittedly partial check about the effect of the online purchase on other purchases. Specifically, we run a post-experiment shopper diary enabling at least partially to verify the trade-offs between experimental and post-experimental purchases (see sections 2.3, 5 and online appendix B). Looking at a representative sample enables us to verify whether taxes on less healthy products are more likely to be (in)effective with particular segments of the population, where evidence from existing research is mixed (see section 2).

We focus on cereals and soft drinks as they are reasonably popular among UK households, include sizeable proportions of both healthy and unhealthy products within each category and are priced within the experimental budget of UK £10. From a public policy perspective, taxes on sugary soft drinks have been implemented in countries such as Mexico and are being debated in many countries (Cornelsen and Carreido, 2015). In relation to the UK, an introduction of a price increase of a minimum of 10-20% via a tax or levy on high sugar products, such as on sugar-sweetened beverages, was one out of eight measures identified in Public Health England (2015) that could help facilitate a reduction in sugar intakes. Cereals are also interesting as they can make an important contribution to micronutrient intake. Many of them represent a good source of fiber (Williamson, 2010), though a considerable and growing proportion of cereals have high levels of sugar and salt (Action on Sugar, 2015).

We consider price changes of 20% and 40%. We do not try to model price promotions but rather taxes, as this has been the almost exclusive focus in the recent policy debate. A price change of (around) 20% is standard in the literature (Bonnet and Réquillart, 2013; Briggs et al., 2013; Epstein et al., 2007; Waterlander et al., 2012a, 2012b), it is at the upper bound of food taxes typically considered by policy makers, and the lower bound of what public health researchers consider as necessary for making a significant, positive impact on health (Mytton et al., 2012). A price change of 40% is a 'large' change and provides a stress test for whether any effect can be found, given the inconclusiveness of the results in the literature. In all cases price changes are in the form of taxes, which rules out the possibility that lab behavior may be driven by stockpiling (as there cannot be offers better than those found in the natural world, to which our baseline prices are benchmarked against), though it does not rule out other kinds of compensation effects, which we deal with using a post-experimental shopper diary.

Our experiment also considers whether signposting part of the price as a tax strengthens the effect of the tax in reducing demand. There are a number of mechanisms which explain why signposting could be effective as a nudge. First, it may increase the salience of a price change relative to a reference price. Second, it may inform consumers with limited attention that the price is different. We note that that the asymmetric results with taxes and subsidies by Feldman and Ruffle (2015), with signposting of taxes mattering but signposting of subsidies not mattering, cast some doubt on these specific interpretations. That said, because their implementation of signposting is different and they do not focus on diet, further research is needed to test these possible mechanisms.

Third, along the lines of Johnson et al. (2012), signposting may reduce the cognitive load of the decision maker by partitioning the products as taxed or untaxed. Fourth, it may inform the consumer that part of an overall price is made up of a tax. Both the third and fourth channels do not however explain why consumers react to a tax as they do. Tax aversion (Sussman and Olivola, 2011) explains this and is consistent with the evidence by Feldman and Ruffle (2015).

In brief, our key result is that taxation of less healthy cereals and soft drinks has a sizeable negative impact on their purchases, particularly, in the case of cereals, when the price change is salient (signposted), though more generally the effects of signposting are nuanced. A policy 'reasonable' tax value of 20% is sufficient to lead to large changes in behavior, which is robust across product categories if and only if signposting is used. We find no evidence of adverse distributional effects between socioeconomic groups to worry about as a result of the price changes. The rest of the paper is structured as follows. Section 2 provides a literature review. Section 3 describes the online supermarket platform, the product categories and healthfulness classification, and the experimental design, as well as providing details on participants, procedures, variables and the econometric model being employed. Section 4 describes the sample characteristics and experimental results and section 5 includes a discussion.

2. Literature Review

2.1 Price Interventions

With thus far few natural world policy impact evaluations and limited incentivized experimental evidence (see below), the bulk of the existing evidence on the effects of price (tax) of food and beverage on diet either relies (1) on epidemiological modelling studies (Briggs et al. 2013) or (2) on econometric analyses of the relationships between food prices and purchase, consumption or diet-related health (Powell and Chaloupka, 2009; Eyles et al., 2012). A limitation of the former approach is that its results critically depend on underlying assumptions and scenarios considered in the model. Also, a limitation of the latter approach is that the observed variation in prices may be endogenous, resulting in biased estimates of the impact of price changes. While it is widely recognized that the case for diet-related taxation is far more nuanced than that for other commodities such as tobacco and alcohol, the picture from the existing modeling and econometric evidence is mixed. Thow et al. (2014) and Cornelsen et al. (2015, p.18) express contrasting views on the effectiveness of the taxation on diet (positive and expressing doubt, respectively). Public Health England (2015) argue for price effects but note the limitations of existing research. Wright et al. (2017) express a similar qualified view but relying primarily on epidemiological modelling studies in their systematic review to argue for a positive finding – and not having any experimental study in their sample. A particularly useful, nuanced review of the empirical (and theoretical) evidence on SSB taxation is by Allcott et al (2018), if with a focus on the US evidence. Powell et al. (2013) provide estimates of price elasticities from their review, with an estimate of -1.2 for soft drinks; they do not study cereals and their sample is made only from US studies. As an example of a specific study, Nevo (2001) find that the elasticity is on average -1.53 in the US ready-to-eat cereals market. That is, assuming a linear relationship between the percentage change in price and the percentage change in demand, a 20% and 40% tax would lead to a reduction in purchases by ~30% and 60%, respectively. While valuable methodologies, the limitations of the epidemiological and econometric approaches justify the usefulness of an experimental approach to complement them like the one in this paper, as this can help identify causality more effectively.

There is a limited if growing number of studies looking at the implementation of significant fiscal policy measures (Wareham and Jebb, 2015; Cawley et al., 2019). Again there is a mixed picture, as apparent for example from what some see as a failure of the Danish fat tax (Snowdon, 2015), the limited population level effect in the case of the Chilean SSB tax (Caro et al., 2018, Nakamura et al., 2018) and conversely the seeming success of the Mexican excise tax on sugary soft drinks that has been found to have led to a reduction in purchasing of soft drinks by 6-9% over the first two years post-tax implementation (Colchero et al., 2016, 2017).[1] In the case of the city-level tax in Berkeley, studies have found a 9% decline in purchases based on sales data, while self-reported changes in consumption indicated a 21% reduction of taxed soft drinks (Falbe et al., 2016, Silver et al., 2017). To date, several of the recently implemented SSB taxes globally have not yet been subject to rigorous ex post-

[1] The previous studies also find that the tax policies might not necessarily be fully passed through to consumer prices. In Mexico there is evidence of full pass-through at least in urban areas (Colchero et al., 2015), and more recently evidence of a more complete pass-through of the tax onto consumer prices (Grogger, 2017). In Berkeley, there is evidence of a comparatively moderate pass-through of the tax to consumer prices (Cawley and Frisvold, 2017: 43%; Falbe et al., 2016: 46-69%). However, Silver et al. (2017) conclude that the degree of pass-through depends on store type: they find a complete pass-through in supermarket chains and gas station chains, whereas the pass-through was modest or even negative in pharmacies, independent corner stores and independent gas stations. Berardi et al. (2016) find evidence of pervasive if heterogeneous pass-through with French soda tax data, if incomplete for flavored waters.

evaluations, not least because the time period since implementation has been too short (e.g. in the United Kingdom case). While essential, a downside of real world tax policy evaluations remains in the limited degree of causal inference that can be achieved – due to the lack of a proper control group and the typically complex circumstances in which such policies are implemented. For example, many claims from observational findings are not seen when studied in randomized trials (Ioannidis, 2013) and may instead reflect complex and insufficiently controlled socio-economic patterns of behavior (Lawlor et al., 2004). Again, this implies that is an important complementary role for experimental studies such as this one.

There has been a small number of recent randomized controlled trials (RCTs) of subsidies to increase purchases of healthier foods, typically fruit and vegetables (Waterlander et al., 2013; Ball et al., 2015), as opposed to the potentially more interesting (from a policy viewpoint) case of taxes which we focus on in this paper. Epstein et al. (2012) and Shemilt et al. (2013) present reviews of experimental evidence on the effect of price changes (generally not labeled as taxes) and this presents mixed results: much of the existing evidence (e.g., Epstein et al., 2010; Giesen et al., 2011; Waterlander et al., 2012a, 2012b) is based on hypothetical choices by unincentivized participants (typically but not always students). Furthermore, it is common to structure the shopping task as an assignment (e.g., think of shopping for a dinner, or for a day of food), which may affect demand by inducing participants (especially those more sensitive to experimenter demand) to buy anyway. We now refer to four studies that go at least partially beyond this paradigm. Février and Visser (2004) used real budgets and a random population sample to test the GARP axiom with an orange juice purchasing task; their interest is not in estimating price effects. Yang and Chiou's study (2010) on beverages is partially incentivized and find evidences of price effects, more pronounced when they are exposed to health claims regarding the beverages. Theirs is an exception to a hypothetical setup in that the beverages were actually purchased; however, unspent budgets were lost, which may have biased the results. Darmon et al. (2014) and Muller et al. (2017) ran insightful incentivized shopping experiments across 180 products in a variety of product categories, with a convenience sample especially targeting low income women. Darmon et al. (2014) found a significant effect on the quantity of unhealthy products only with a combination of a 30% price increase and decrease on unhealthy and healthy food respectively, and recognized the potential of a regressive effect of the tax intervention. Muller et al. (2017) found that low income mothers consumed more unhealthy food than the benchmark group (i.e. medium and higher income mothers) – in line with theoretical predictions by Mytton et al. (2007) -, but they showed lower price sensitivity than the benchmark group. Limitations of these experiments are that the shopping tasks are structured as an assignment, which may affect demand schedules; and that there is not a one-to-one correspondence between shopping made and obtained (due to having access to a subset of grocery goods). A more general limitation of all experimental studies is the lack of control for compensation effects between shopping in the laboratory and later shopping.

Our experiment addresses these limitations by having a fully incentivized experiment with 'quasi-currency' (reward points convertible for money in a large number of popular websites: see section 5), with a one-to-one correspondence between shopping made and obtained, and with shopping tasks not structured as assignments other than in the minimal sense that subjects are invited to a website where they can shop.[2] We control for order effects by presenting our shopping tasks in random order. We also make a first (if undoubtedly partial) attempt of controlling for the trade-offs between shopping

[2] To some extent, this can still be interpreted as an assignment. However, by minimizing the context as much as possible and by recruiting subjects who consume these products on a regular basis (see section 3.4 for details), we can be reasonably confident that the participants spend the budget that is allocated to them on products they would usually allocate money to on a regular basis when they do go to either a physical or an online supermarket. Therefore, while we cannot rule out that the demand schedule is distorted, this is likely to be less so than in experiments with explicit assignments.

in the experiment and shopping outside (if afterwards) by verifying whether such compensation effects between the two exist using post-experimental diary data (see section 2.3, 5 and online appendix B).

Overall, from the existing literature there is not a clear conclusion on the health inequality impact resulting from tax interventions, neither from non-experimental evidence. Briggs et al. (2013) predict no significant difference in the impact of sugar sweetened drink tax in the UK, and Blakely et al. (2011) find no differential impact (by income and education group) of a subsidy on fruit and vegetables in their randomized control trial in New Zealand. Sharma et al. (2014) argue for a positive effect based on an econometric analysis of the Australian sugar tax. Smed et al. (2007) reach broadly a similar conclusion with Danish data estimated price elasticities. Based on theory-based simulations, Schroeter et al. (2008) warn however that a high calorie tax may be deemed equitable only if combined with income redistribution to low income households.

Some previous studies, irrespective of research designs, particularly investigate differential impacts of price/tax on consumption by socioeconomic and demographic group. Using consumer scanner data from the United Kingdom, Dubois et al. (2017) find that a hypothetical implementation of a soda tax would be more effective for younger and poor individuals than it would be for other types of individuals. They also find that the impact of the tax would not differ by baseline consumption level of sugar. Although their predictions are backed by previous systematic evidence review (Green et al., 2013), more recent impact evaluations of actual tax policies find otherwise. Cawley et al. (2018) evaluate the impact of the Philadelphia's soda tax on purchases of soft drinks, and they find no systematic differences in the magnitude of the impact by income. They did not find a significant tax impact on children, except for those whose pre-tax consumption level had been high. Nakamura et al. (2018) even find that the impact of soda tax for high socioeconomic group was greater than that for lower socioeconomic groups in Chile's tax reform on sugar sweetened beverages in 2014. Moreover, a modelling study from the United Kingdom shows that the impact of soda tax would be greater for younger than older generations, but there would be no differences by income group (Briggs et al., 2013). Hence, the distributional effects of a tax – e.g. by demographic characteristics, as we do in our experiment – remain very much an open empirical question that merits further investigation.

2.2 Signposting interventions

There has been growing policy interest in the use of "nudge" or "choice architecture" approaches to correct behavior (Thaler and Sunstein 2009; Marteau et al. 2011; Hollands et al. 2013, 2017; Guthrie et al. 2015). Nudges are "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives." (Thaler and Sunstein 2009, p. 8). One type of nudge that would be of policy interest to use alongside fiscal policies is recognized to be signposting of complementary information to raise awareness at the point-of-purchase (Hawkes et al 2015). Signposting information can affect behavior regardless of whether the consumer is aware of the reference price since the signposting can partition the products (taxed vs un-taxed) and reduce the cognitive load for the decision-maker (Johnson et al., 2012).

Guthrie et al. (2015) distinguishes nudge type interventions such as ours from education-based interventions. 'Significantly changing economic incentives' is normally interpreted with respect to the benchmark of the purely rational and self-interested agent. From the perspective of a rational consumer who only cares about the price and who has no attention constraints, and noting that only one shopping task counts towards payment, being informed that part of the price is made up of a tax should be irrelevant. Therefore, there is no significant change in economic incentives and signposting can be classified as a nudge.

Out of the wide range of potential nudge approaches, only very little research effort has gone into studying the effect of signposting information in the context of diet-related fiscal policies. With the exception of Zheng et al. (2013: theory and simulations), research has been conducted using an experimental approach.

In non-diet related settings, Chetty et al. (2009) may have been the first to examine the impact of making explicit a commodity tax in the price tag, finding that this did increase the consumers' sensitivity to the tax. Sussman and Olivola (2011) found a general tendency for tax aversion, namely consumers disliking taxes more than equivalent price changes. Goldin and Homonoff (2013) found that high income consumers were less responsive than low income consumers to a tobacco tax that was levied at the cashier but were similarly responsive to a tax that was presented in the price tag. Feldman and Ruffle (2015) found that, while tax-exclusive prices with the announcement of tax inclusion at point of purchase are effective in increasing consumption relative to a case where tax is explicitly included at point of purchase, subsidy-exclusive prices with the announcement of subsidy inclusion at point of purchase make no difference. Taubinsky and Rees-Jones (2018) found individual heterogeneity in the degree to which subjects under-react to a sales tax that is not made salient. They also show that this heterogeneity significantly increases the efficiency loss due to distortion of demand driven by such a non-salient tax.

In their systematic review of health focused tax interventions (not just food related), Wright et al. (2017) note the lack of studies combining the two even though they recognize it as a possible reason why US studies of tax interventions in their sample (where taxes are signposted) tend to show stronger effects than those elsewhere (where taxes are not sign-posted).

Together with Chen et al. (2015), ours is the first experiment looking at the effect of signposting in a diet context. Chen et al. (2015) conducted a laboratory experiment with non-student university-connected adults, who first made purchases with a $10 voucher with baseline food prices and then made purchases with another $10 voucher with prices of unhealthy products increased by 20%. They either have a treatment where on top of the screen they state "A 20% 'unhealthy food' excise tax has been added to the price of unhealthy food and beverages" or a treatment where on top of the screen they state that "A 20% 'unhealthy food' sales tax will be added to your purchase when you check out." They found effects on purchases of both signposting methods, though more in the first than the second case. There are a number of limitations to this study, however. First, as in Yang and Chiou (2010), unspent budgets are lost. Second, there is a wealth of information provided in terms of 'unhealthy food', making more psychological effects than just those entailed by a tax label at work, including potentially experimenter demand. Third, the effect of signposting is likely to be mediated by not controlled subjects' beliefs about what makes an unhealthy product, as well as by whether there is or is not recall of earlier prices in the first treatment; to clarify, both the subjects' beliefs and the extent of recall may have affected the results in terms of its stated aim of verifying the impact of signposting. Our experiment addresses the limitations in Chen et al.'s (2015). Only one shopping task, picked up at random at the end of the experiment, gets implemented in terms of outcomes, therefore making shopping tasks independent from a rational choice perspective. Also, in our experiments we do not label products as being healthy or unhealthy and our signposting procedure (described in section 3.3) does not depend on subjects' recall or beliefs about which products are unhealthy.

2.3 Compensatory behavior and price ceiling effects

As we have already mentioned in this section, one of the key limitations of the existing experimental/intervention studies is the lack of control for compensatory consumption of unhealthy diet which may occur off the experiment. Even if a tax on dietary product leads to healthier outcomes in terms of direct study outcomes (e.g. reduction in purchases of sugary drinks within the experiment), this effect may be undermined by later unhealthy purchases (e.g. increase in purchases of sugary drinks in real supermarket).

Wisdom et al. (2010) emphasize the importance of controlling for the compensatory behavior in their interventional study to induce healthier choices of sandwiches in lunch time. They find that although their intervention led the participants to choose healthier sandwiches, the participants compensated by increasing calorie consumption from side menus and drinks (which were not subject to their intervention). Furthermore, in the literature of impact evaluations of tax policies using observational data, studies tend to focus on purchases of targeted items (e.g. soft drinks) only and do not necessarily investigate the substitution or compensation (Colchero et al. 2016; Nakamura et al. 2018; Cawley et al. 2018).

A related effect is the so-called price ceiling effect, namely a subject should not agree to obtain the same commodity in an experiment at a price that could be beaten outside the experiment (Harrison et al., 2004). This effect is identical to compensation effects with positive taxes, but exists even in the lack of taxes.

As a key methodological contribution to the existing literature, we conduct a post-experimental survey of purchasing of targeted items (cereals and soft drinks) and an additional item (chocolate) to investigate if and how post experimental shopping patterns differ based on experimental purchases. Because of our experimental design, we can also test for price ceiling effects other than compensation effects, by verifying how purchases reported after the experiment relate to what participants purchased when there were no taxes. We discuss this in section 5 and in more detail in online appendix B.

3. Study Platform and Design

3.1 Online supermarket platform

A website built to emulate an online supermarket platform was used. The website was developed by Cauldron Inc. for the BHRU. The website mimics the appearance of a regular online grocery store, including browsing, search, unique product pages, trolley and checkout. Image, price, full description, and a table of macronutrients were available for each product. The platform automatically collects a range of data, i.e. product name, price, number of units purchased, price information, nutrient content, browsing history and time spent within the site. The platform was built for assessing the effectiveness of a series of food purchase interventions (Forwood et al., 2015) and further modified for this study. In this study the range of products was restricted to two target categories – cereals (189 products) and soft drinks (709 products). The food database was a copy of the range of products for the corresponding categories from the largest supermarket chain in the United Kingdom, Tesco (scraped from Tesco.com in August 2015), supplemented with nutrient composition per 100 g (100 ml) from food (drinks) labels available at Tesco.com or from databases for common foods supplied by MRC Human Nutrition Research (Fitt et al., 2015). Tesco's market share at the time of the experiment was 28.3% (McKevitt, 2015). Grocery deliveries were organized by the team using Tesco's home delivery service.

In our experiment we only included breakfast cereals and soft drinks (only one category per task). The number of products per page was customized to be up to 50, and price manipulations and signposting were implemented according to the treatment. A demo (not restricted to breakfast cereals and soft drinks, and with baseline prices) can be found at http://woodssupermarket.co.uk/login?demo=1. Additional demo videos showing relevant price manipulations and signposting, as implemented in the experiment, are available upon request.

3.2 Product categories and healthfulness classification

As noted, we focused on the following product categories: (i) cereals, which included ready-to-eat breakfast cereals, muesli and granola and (ii) soft drinks, which included sports and energy drinks, waters, juices and fizzy drinks. We exclude fresh milk, fresh fruit juices and similar products on the grounds of their perishability and the risk of problems associated with delivery. Products in both the cereal and the soft drinks categories were divided into two groups: healthier and less healthy. We classified products into those groups based on the same criteria used by the UK broadcasting regulator Ofcom to enforce restrictions of television advertising to children. Those criteria rely on the UK Food Standards

Agency's nutrient profiling (NP) model (Rayner et al. 2005). An advantage of this method is that it provides a unified measure of healthiness across all available food and drink products (Nakamura et al., 2015). The model uses a simple scoring system based on the nutrient content of 100g of a food or drink. The score for each product is based on the energy density, saturated fat, total sugar and sodium and protein contents together with an estimate of the fruit, vegetable, and nut contents. An increase in the NP score reflects a reduction in the healthiness of the product. In order to classify products, we followed the same thresholds used by Ofcom to regulate food and drinks advertisement to children, that is, foods scoring 4 or more and drinks scoring 1 or more are defined as less healthy, and those scoring less are defined as healthier.[3] Lobstein and Davies (2009) offer a comprehensive discussion on nutrient profiling models for public health policy purposes.

3.3 Design

Participants had a budget (U.K. £10) and undertook a series of online shopping tasks.[4] In each shopping task, items from only one product category (either cereals or soft drinks) were available for purchase. In the baseline tasks, one for each product category, item prices were matched with Tesco.com prices (excluding promotions). In the intervention tasks, the prices of either less healthy (healthier) products were taxed by either 20% or 40%, while the prices of the healthier (less healthy) products remained at the baseline level. This results in a total of ten shopping tasks, five for each product category, namely: (i) baseline prices, (ii) 20% tax on less healthy products, (iii) 40% tax on less

---

[3] Some examples of less healthy cereals are: Kellogg's Rice Krispies (score = 6), Nestle Cheerios Cereal (score = 8) and Kellogg's Frosties Cereal (score = 13) and examples of healthier cereals are: Nestle Shredded Wheat Cereal (score = -6), Alpen No Added Sugar Muesli (score = -3) Dorset Cereals Honey Granola (score = 0). Examples of less healthy soft drinks are: regular tonic water (score = 1), regular Coke (score = 2) and Lucozade Energy (score = 3), while water (score = 0), Diet Coke (score = 0) and orange juice from concentrate (score = -3) are examples of healthier drinks. Notice that a healthier food or drink is a relative concept and depends on the threshold applied; hence it is not necessarily equal to what would be generally accepted as a healthy food or drink. A good example is diet sodas, which are healthier than sugar added soft drinks, but would not be considered healthy per se.

[4] A potential concern is that £10 could induce an income effect that might distort our results. While we do not believe this credible in a UK context (£10 is just a little over 1 hour of minimum wage pay), we test for whether household income correlates to purchases made and find no evidence for this (see section 4). Also, a pilot experiment controlled for whether the fact that the £10 were not earned made a difference, which it did not (see online appendix C).

healthy products, (iv) 20% tax on healthier products, and (v) 40% tax on healthier products. These price interventions allowed us to estimate the effect of our price interventions within each product category. While only the case of taxes on the less healthy products is of direct policy interest and of interest in this paper, we decided to be symmetrical between the two cases in order to avoid making overly transparent what the experiment is about, and therefore to reduce the likelihood of experimenter demand effects (Zizzo, 2010).

In addition, participants were randomly allocated to one of the following treatments: *Signposted* tax and *Not Signposted* tax. In the Signposted tax treatment the amount of the price increase is presented separately from the original price and flagged as a tax, whereas in the Not Signposted treatment the price increase is kept implicit (only the final price is presented). Figure 1 shows an example of a 20% tax on a less healthy cereal. An important distinction between this treatment and previous studies such as Chetty et al. (2009) is that in our study the participants always face tax, signposted or not, at the point of choice, whereas in other studies consumers only face tax at the cashier in a non-salient condition (but they see tax in the price tag in a salient condition).

The price intervention was a within-subject manipulation while the signposting intervention was done between-subjects, with 500 subjects in each category (with the number of observations within each product category being equal to 500 subjects x 5 tasks).[5] We used a standard Random Incentive Mechanism, i.e. once participants complete all shopping tasks one shopping task was randomly selected and played out for real, i.e. participants received the groceries delivered to their homes and the remainder of the budget was credited to them by the survey company (Research Now). The instructions never referred to taxes as either ‘fat taxes’ or ‘sugar taxes’. This follows standard experimental methodology, and again helps control for experimenter demand effects particularly in an initial study.

[5] We discuss potential concerns with our within-subject design in the discussion section.

The order in which products appeared on the Woods online supermarket platform was exactly the same within subjects, price manipulations and signposting treatments. This order was pre-determined by the order that Tesco.com had set at the time of the experiment to organize products online (all the information including the order was retrieved from their application-programming interface – API).[6] To minimize order effects, we set up each product page to show up to 50 products without having to click "Next page", and so in one page most of the products belonging to the same sub-category were shown together. In addition, as in real online retail shopping sites, in our experiment the participants had the possibility of using a search tool located at the top of the webpage to find products or group of products within the range available for that task.

Participants were free to spend as much as desired from their £10 shopping budget, including the option of checking out with an empty basket. We decided to implement this design feature against the option of forcing participants to spend all or at least part of their shopping budget, because the latter would inflate the purchasing data unrealistically and potentially bias the estimation of tax effects. One of the potential negative consequences of this was having a large number of participants checking out with an empty basket. We reduced the risk of this happening by limiting subject eligibility to consumers who bought a product from each of the two target product categories (cereals and soft drinks) at least once in the last month.

3.4 Participants

A total of 1,000 participants (54.10% female; mean age 46.95, SD=15.9 years; BMI= 26.49, SD=5.28) completed the experiment. All participants were based in the UK and the sample was representative in terms of regions, gender, age

[6] Products organization follows an internal hierarchy grouping which is similar to the brick and mortar (B&M) supermarket departments, aisles and shelves. For instance, all drinks are grouped in the same product category (as they would be in the same department in a supermarket), and within this category products that share similar characteristics are grouped by sub-categories (as they will be within a supermarket aisles and shelves). As an illustration, "Drinks/ Soft Drinks/Colas" was a sub-category that a consumer in a Drinks task could see with a couple of clicks to get there.

ranges and social classification. Participants were recruited via Research Now, a market research online panel company (http://www.researchnow.com). Quota targets were based on the Office for National Statistics annual mid-year estimates 2014. We used three screening questions in order to restrict our sample to consumers who purchased the target products – cereals and soft drinks – at least once a month and were responsible for at least half of the shopping of their household. We also excluded respondents in a diet for medical reasons and those answering incorrectly one or more of the data quality questions. See Q1.2-Q1.5 and Q1.7 in the appendix E for the questions' exact phrasing.

3.5 Procedure

Participants took part in the study over the internet between August and October 2015, and both participant and experimenter were blind to treatment allocation. The experimental instructions can be found in online appendix E. The study was conducted using specialized online survey software (www.qualtrics.com) and the Woods online supermarket discussed in section 3.1. At initial assessment, participants completed a consent form, questions about their shopping habits and quality control questions. Participants not meeting the inclusion criteria and/or failing to answer correctly one or more of the quality control questions were excluded at this point. Eligible participants were randomly allocated to one of the two experimental treatments. Participants each completed ten shopping tasks, presented in random order. Both randomizations were performed using a built-in feature within the survey software. Product name, number of units purchased, price information, and total spending were recorded for analysis. Participants provided their address, relevant contact details and delivery preferences. Following completion of the shopping tasks, participants were informed which task (and hence products) was (randomly) selected to be delivered to them. Personally identifiable information was used solely for the purpose of organizing the delivery of the groceries to the participants, who knew this. Subsequently, participants answered a post-experiment questionnaire. Within the next eleven days participants received the purchased groceries via Tesco.com and any unspent budget was transferred by the survey company via panel points rewards of equivalent value in a large set of popular websites as detailed in section 5 (within 28 days after completion of the experiment).

3.6 Variables

We used the following *outcome variables*: the total quantity of less healthy cereals (weight in grams) purchased in each task by each participant; the total quantity of healthier cereals (weight in grams) purchased in each task by each participant; the total quantity of less healthy soft drinks (volume in ml) purchased in each task by each participant; and the total quantity of less healthy soft drinks (volume in ml) purchased in each task by each participant. For simplicity, from now on we will call these variables "volume purchased of …" as appropriate. We now list the other variables.

Socioeconomic status: As a measure of individual level socioeconomic status, participants provided their highest level of educational qualification attained, coded onto a six point scale ranging from 0 'No qualifications' to '6 Degrees or higher'. As a measure of income, participants provided both their personal yearly income and their total household income before tax (participants selected from 4 income bands in each case).

Participant characteristics: Data on gender, age, self-reported height and weight (used to calculate BMI), the number of adults and children living in their household, and being on a weight loss diet, was collected.

Other control variables: We elicited participants' time preferences using a ten choices discount rate price list for a six months' time horizon, similar to Harrison et al. (2002). Participants were asked to provide an estimation of the number of days their stocks of soft drinks and their stocks of cereals would last until they would run out and need to buy some more. Participants were also asked to provide a measure of the frequency they purchased groceries online, and the frequency they shopped at Tesco, both variables coded on a 7 point Likert scale ranging from never to more than once a week. Finally, as a measure of price recollection, participants were asked to indicate their best guess of the regular retail price of six products, i.e. three cereals and three soft drinks.

3.7 Econometric model

We exploit the panel nature of the data in the analysis. Since the data on the volume of products purchased are censored at 0, we have the following Random Effects Tobit regression model (Wooldridge, 2010). For subject j in task *t*:

(1)
$$\begin{aligned}\ln\mathrm{Y}^{*}_{jt} = {} & \alpha_j + \beta_1 \mathrm{tax20}^{H}_{jt} + \beta_2\left[\mathrm{tax20}^{H}_{jt} \times \mathrm{signpost}_j\right] + \gamma_1 \mathrm{tax40}^{H}_{jt} \\ & + \gamma_2\left[\mathrm{tax40}^{H}_{jt} \times \mathrm{signpost}_j\right] + \theta_1 \mathrm{tax20}^{LH}_{jt} \\ & + \theta_2\left[\mathrm{tax20}^{LH}_{jt} \times \mathrm{signpost}_j\right] + \mu_1 \mathrm{tax40}^{LH}_{jt} \\ & + \mu_2\left[\mathrm{tax40}^{LH}_{jt} \times \mathrm{signpost}_j\right] + \delta \mathrm{signpost}_j + \mathbf{z}_j^{'} \boldsymbol{\varphi} \\ & + \varepsilon_{jt};\end{aligned}$$

$\ln\mathrm{Y}_{jt} = \ln\mathrm{Y}^{*}_{jt}$ if $\ln\mathrm{Y}^{*}_{jt} > 0$ and $\ln\mathrm{Y}_{jt} = 0$ otherwise.

$\ln Y_{jt}$ gives the volume of healthier and less healthy products purchased in each product category. We use a log-scaled variable because the distribution of Y_{jt} is highly skewed (see the online appendix A for distributions). The log-scaled variable closely follows a normal distributions with no outlier values, except zero, which is addressed by using Tobit regression models.

The key independent variables $\mathrm{tax20}^{LH}_{jt}$ and $\mathrm{tax40}^{LH}_{jt}$ are indicators of the price manipulations, representing the 20% or 40% tax on less healthy (LH) products, respectively. These variables are interacted with the indicator of the signpost condition. Note that the coefficients alone represent the impacts of the interventions on the latent outcome ($\ln\mathrm{Y}^{*}_{jt}$), rather than the impacts on the actual outcome ($\ln\mathrm{Y}_{jt}$). We are interested in the latter, and hence we re-evaluate the estimated coefficients to represent the partial effects on the actual outcome; this involves weighting of the estimated coefficients by the probability of not being censored. The vector $\mathbf{z}_j$ includes various other variables as discussed. In order to take into account the within-subject correlation in purchasing, we estimate the standard errors of the parameters via block bootstrap, which gives subject-level cluster-robust standard errors (Bertrand et al. 2004; Cameron et al. 2008).

4. Results

Demographic characteristics of participants are presented in Table 1. The total number of participants was 1,000, with a total of 506 in the Not Signposted treatment and 494 in the Signposted treatment. There were no significant between-treatment differences in these characteristics. Tests employed are Pearson's χ^2 (for gender) and the two-sample Mann-Whitney U test for the ratio variables. The distribution of the ratio variables, which departs considerably from normality justifies using a non-parametric approach.

The mean quantity purchased across all tasks was 1.5 kg of cereals and 5.6 liters of soft drinks. Only about 5.6% purchases were zero (5.8% for cereals and 5.4% for soft drinks). When only considering less healthy products within each task, the amount of zero purchases was 13.5% for cereals and 13.2% for soft drinks. The weighted average price of purchased less healthy breakfast cereals (across all the tasks and weighted based on the quantity purchased) was £0.52, (SD: 0.20) per 100g and for soft drinks it was £0.16 (SD: 0.13) per 100ml.

Figure 2 shows mean purchases of less healthy (healthier) products when prices of less healthy (healthier) were manipulated. Figure 3 relies on the econometric model to show the estimated effect and confidence intervals for each price manipulation on the volume of purchases of the less healthy taxed products for cereals and soft drinks, respectively. The model estimates are in Table 2, while Table 3 extracts the key information by presenting price and cross-price effect estimates for less healthy cereals and soft drinks corresponding to each price manipulation and treatment. While we report the results of regression models with the control variables, the general results remain the same with different combinations of controls or no controls at all. The corresponding results for healthier products are in the online appendix D. Mean values and distributions of volumes purchased at baseline prices and each price manipulation are also available in the online appendix (Figures A.1-A.4 and Table A.1; this appears mainly driven by fewer units bought).

Let us first focus on the Not Signposted treatment. Both the 20% and the 40% taxes on less healthy cereals significantly reduced the volume purchased of the taxed products ($p<0.01$: see Table 3), though the effect of the 40% tax was not significantly larger than that of a 20% tax. A 20% tax was sufficient to induce a 48% demand decrease for cereals. For soft drinks, in the Not Signposted treatment, a 40% tax was instead required to significantly reduce the purchasing of less healthy soft drinks.

Result 1. With no signposting, while a 40% tax always works, a 20% tax is effective in reducing the demand for less healthy products in relation to cereals but not in relation to soft drinks.

Let us now consider the Signposted treatment. Again the 20% and the 40% taxes on less healthy cereals significantly reduced the volume purchased of the taxed products ($p<0.01$), with no difference between the two. A 20% tax was now sufficient to induce a 54% demand decrease for cereals. With signposting, a 20% tax was sufficient to induce a similar (53%) and significant demand decrease for soft drinks ($p<0.01$).

Result 2. With signposting, a 20% tax is sufficient to significantly reduce the demand for less healthy products in relation to both cereals and soft drinks.

We now consider cross-price effects.

Result 3. For both a 20% and a 40% tax rate, there is a sizeable cross-price effect increase in the purchasing of healthier cereals when a tax on less healthy cereals is introduced and there is no signposting. There is no such cross-price effect for soft drinks.

Table 3 provides evidence for Result 3 (e.g. $p<0.05$ with a 20% for cereals with no signposting).

As supplementary findings, and focusing only on ones consistent between cereals and soft drinks, Table 2 show that men generally bought less healthy cereals and soft drinks than women, and consumers with greater stocks at home bought less unhealthily. Interestingly, we found no robust evidence of greater income and a better education as a predictor for better diet choices, though there is some partial evidence to this effect in relation to soft drinks only. Stocks of food were negatively correlated with purchases of less healthy food (in supplementary work, we found this effect robust to whether there is signposting).

We also checked for potential differences in tax and signposting sensitivity by socioeconomic status, in particular participants' household income level and educational level. Online appendices Tables A.2-A.3 show that the estimated effect of our price manipulations are not statistically different between the lower and the higher income group, while Tables A.4-A.5 show that there were no significant differences in price manipulation sensitivity between the less educated and the more educated group (95% confidence intervals for estimated effects overlap). Using the same criteria, we found no significant differences when comparing the groups with lower and higher body mass index (Tables A.6-A.7) and between patient and impatient individuals, classified using our time preferences measure (Tables A.8-A.9).

5. Discussion

Do taxes on less healthy products lead to healthier diet choices in relation to breakfast cereals and soft drinks? The broad answer from our experiment is a clear yes, and quite independently of signposting. However, in the case of cereals, our findings suggest that signposting does seem to affect whether such taxes increase the purchasing of more healthy cereals (non-signposted case) or just reduce that of the less healthy cereals (signposted case).

Does signposting matter? Yes in making the effectiveness of the tax interventions more robust, though the specific answer depends both on the product category and on what we are trying to achieve. Based on our results, the 20% tax recommended by the WHO (2016) would only be effective in reducing the consumption of less healthy soft drinks if signposting is used. In the context of cereals, if the goal is to reduce the demand of less healthy cereals, there is clear evidence of a potentially substantial effect. The magnitude of this effect is potentially substantial, and would lend itself to a fairly straightforward, inexpensive policy implication – that food-related taxation could be more effective when it is combined with signposting on the price tag. The implementation of such a complementary ‘nudge’ policy would enhance the impact of the tax without imposing severe welfare loss, though we should recognize that its effectiveness may depend on the product category. The further qualification to this is that the consumption of healthier cereals would not be increased in case of signposting – this may or may not be problematic depending on a more comprehensive evaluation of optimal dietary choices.

Our results are noteworthy as they are based on real purchases of a nationally representative sample of 1,000 consumers in the UK, the largest sample we are aware of for tax field studies on diet. The large sample size and representativeness of the data, combined with experimental control over price and signposting variations as well as controls for compensation effects and a range of other factors, provide a valuable contribution on the impact of the tax on food and beverages, and should thus credibly inform the current policy debate on the subject. Our study also goes beyond Chen et al. (2015; see sub-section 2.2) in innovatively providing an evaluation of the effectiveness of signposting in diet context.

Repetition and experimenter demand effects. Having subjects play 10 periods had advantages and disadvantages. The key potential disadvantage is that repetition enhances the salience of the changes, in terms of prices and in terms of signposting, leading to a potential experimenter demand effect confound. Specifically, by drawing attention to the price changes – particularly in the presence of signposting -, this might lead to an overestimation of the effect both of price changes and of signposting. Alternatively, one might think that the within-subject changes in prices may reduce the impact of signposting. However, since signposting is always combined with the prices being changed in our experiment, it is more plausible that the effect of signposting is enhanced rather than reduced by the repetition.

Natural world food shopping is a repeated task, and so a within-subjects design is ecologically valid in reproducing real-world conditions, including how changes in policy get implemented – consumers see them as a change relative to the previous time period –. It is also in line with the experimental research reviewed in section 2.

That said, we accept that the focus on two product categories in repeated tasks may potentially distort the effectiveness of signposted taxes, and it is therefore useful to test for this. If a potential experimenter demand effect due to task repetition were to exist, in period 1 there should be greater effectiveness of taxation under signposting. In the online appendix Table A.10 tests this via 3-way interaction effects and finds no evidence that this is the case.

There is of course a broader disadvantage from repetition in that there could be potential order effects. We control for this by presenting tasks in random order, and our data analysis checks did not find evidence for any time sequence effects.

Having a control group of participants who never face any tax would only be possible if we had opted for a full between-subjects design, i.e. each individual is exposed to only one treatment (price manipulation). If we had followed this approach, we would have required a much larger sample to have similar statistical power, which would have been unfeasible. Note that each group would need to be sufficiently large to contain a representative sample of the UK population. A further advantage of having a within-subject design is that its internal validity does not depend on random assignment and they lend themselves to more powerful econometric techniques (Charness et al, 2012).

Signposting. Natural world implementation of signposting of a price would be framed as a tax and therefore we retained this feature. Our tax signposting was deliberately minimal, and set up for an initial experiment such as ours (together with the taxation on healthy products) to minimize the transparency of the objective of the experiment and therefore avoid experimenter demand effects (Zizzo, 2010). Future research could helpfully determine the effect of an entirely neutral label as this could help test for tax aversion as an explanation. It could also determine the effect of a stronger frame, e.g. labeling the tax as a 'health tax'. This could plausibly increase the impact of the health tax, although there have been enough examples of public backlash against soda tax and other health interventions that this is not a foregone conclusion (Wright et al., 2017; Just, 2017).

Own-price effects. We recognize that our estimates of the own-price effects are larger than what has been found in previous research (Briggs et al., 2013; Andreyeva et al., 2010). The fact that we have pre-filtered the sample for participants with latent demand does not explain why there is a higher estimate than in econometric studies estimates, as these are based on empirical demand functions that, by definition, rely on latent demand (e.g., if you never buy cereals, you will not affect the slope of the demand for cereals function). In comparison to our around 50% drop in demand for a 20% tax, Nevo (2001) found that the elasticity is on average -1.53 in the US ready-to-eat cereals market, i.e. assuming linearity (that is, a linear relationship between the percentage change in price and the percentage change in demand), a 20% and 40% tax would lead to a reduction in purchases by about 30% and 60%, respectively. Our results suggest that the marginal proportionate effect of taxation diminishes as the tax rate is raised from 20% to 40%, for both cereals and soft drinks. When clearly signposted, even a small tax could make a substantial difference, which is consistent with the other signposting research as discussed in sub-section 2.2 (though comparing magnitudes is unfeasible given differences in experimental designs and how signposting is implemented).

A potential issue with our results, as well as other research in this area, is that there may be compensation effects such that the study manipulation, while effective in terms of direct study outcomes, would be made pointless by later unhealthy purchases. We need to recognize that the current non-experimental evidence also does not necessarily address the compensation effect problem (Wisdom et al., 2010). We do not know the extent to which, for example, Mexican consumers have replaced sugar consumption with equally harmful sugar surrogates (Colchero et al., 2016, 2017).

Compensation and price ceiling effects. In order to examine any potential compensation effect between our online shop and alternative shopping opportunities, upon completion of the online shopping, participants were invited to record their purchases of the two relevant product categories, namely cereals and soft drinks, as well as the additional category “chocolate” for one week after they received the shopping basket delivered to their homes. There is some evidence of substitution between different sources of sugar, including, specifically, between demand for high sugar soft drinks and demand for chocolate (see Smith et al., 2018, based on UK Kantar data). Below, we refer to this part of our study as the shopper diary.

Online appendix B describes the results from the shopper diary trying to provide a first and partial control for compensation effects. We do this in two ways and exploit the fact that only one shopping task was implemented for real, and subjects did know which one after the experiment. We look at how the volume of less healthy/healthier products received in the experiment affects post-experimental purchases, both in general and in relation to the baseline and to each of the different tax conditions. We also look at how taxes in the implemented task directly affected post experimental purchases. We do not find evidence for compensation effects using either method.

The shopper diary also enables us to test price ceiling effects (Harrison et al., 2004) other than compensation effects, by focusing on the cases where the implemented task had baseline prices. Specifically, we can see whether the volume of less healthy / healthier products purchased at baseline prices correlates with the volume of less healthy / healthier products purchased after the experiment (see Table B.5.2 in online appendix B). We find no effects, other than a higher volume of cereals being bought in the experiment being strongly correlated with a lower volume of healthier cereals bought after the experiment. However, the marginal effect of the purchase of less healthy cereals on healthier cereals (-0.475) is virtually identical to that of the purchase of healthier cereals on healthier cereals (-0.461), suggesting that this effect does not interact with our findings on the effect of taxation on less healthy products in a way that may affect our results.

That said, this is *some* evidence of subjects not treating the experimental task in isolation; another piece of evidence is the strong negative correlation between prior stocks of soft drinks/cereals and experimental purchases. The volume of quantity purchased in the experiment clearly decreases with a 40% tax, though the effect is muted or non-existent with a 20% tax (see Table 2, and Table A.1 for mean volumes by treatment). One may query why there is not a bigger substitution between healthy and unhealthy products of the same category. This is however not surprising as preferences for soft drinks and perhaps especially cereals is likely to be habitual to some degree, and so at least some consumers may be disinclined to replace (for example) frosty corn flakes with regular corn flakes. The lack of compensatory effects is particularly reassuring in this context.

Why do we observe at least partial isolation of experimental choices? One explanation could be similar to the one of online as well as physical supermarket shopping: there is a transaction cost (in time, if nothing else) from looking for other suppliers and risking for a better offer not to be there (Sugden et al., 2019). This possibility is perhaps made more credible by the market competitive nature of the Tesco's online prices which we rely on. Shopping at high end supermarkets or local shops can be considerably more expensive, and we were more competitive even than Tesco's online supermarket in that we did not charge subjects for delivery fees (£4 at the time of the experiment, which, given the £10 budget, implied that, for a comparable expenditure, shopping with us was never more expensive, even with a 40% tax). Another and possibly complementary explanation is experimental narrow framing, as found at least partially in other settings such as risk taking (e.g., Cox and Sadiraj, 2006; Fafchamps et al., 2015; Andersen et al., 2018).

We acknowledge that our shopping diary data is not conclusive regarding the absence of any compensation effects, which of course may take place over a period longer than one week after receipt of the groceries, and may involve other product categories. We cannot therefore entirely rule out that our price effects may be overestimated because of potential compensation effects, or indeed from unobserved price ceiling effects. One piece of additional suggestive evidence of compensation effects or price ceiling effects not being a major issue is provided by the fact that our results are robust to the introduction of our Tesco prices recollection measure among the control variables used in the regression analysis. Overall, our results should be considered as evidence for the effectiveness of taxation, particularly when signposted, but not clearly as evidence for the specific quantitative size of the effect, which obviously needs to be seen in complement with other experimental and non-experimental evidence. Equally clearly, more experimental research with more comprehensive shopper diary follow-ups, and perhaps with experimental questions specifically probing price ceiling effects, would be useful.

Other limitations. Another potential limitation of our analysis is that participants may be subject to house money effects, namely they may behave differently if the money is just given to them – as it was in our setup – as opposed to being earned. However, in a pilot we ran with university students (see online appendix C), we tested for house money effects and found no evidence of them in our setting. We therefore opted for simplicity in the online study design.

A different limitation of our analysis is that we were unable to directly provide money because of constraints from the market survey company. However, we note that the market survey company sees rewards points as effective in recruiting and retaining participants in their subject pool for a wide range of studies as 'quasi-currency' can be converted into vouchers/gift cards of a wide variety of retailers including Tesco.com, another major UK supermarket chain

(Sainsbury's), the largest UK department store (M&S), the largest UK catalogue retailer (Argos), major coffee and cinema chains, other major retailers and even Amazon (see https://www.valuedopinions.co.uk/rewards). This supports their status as 'quasi-currency'. Clearly, it would be useful in future research if actual currency could be used.

A possible limitation of our analysis is that, since the purchased goods are not immediately available, the choices do not have immediate consequences and may therefore be less subject to behavioral biases such as signposting. In that sense, the study looks at medium term consumption. We note however that the same holds for any online shopping; that the goods purchased in supermarkets, particularly breakfast cereals, are not normally for immediate consumption in any case; and that, if anything, this potential confound should work in the direction of reducing the effect of signposting, whereas we find that this is in any case effective in making a 20% tax on less healthy products robustly reduce consumer demand. In other words, our key finding would be strengthened if this confound were to matter.

A further limitation of our analysis might be considered that we focus on just two product categories. Clearly, more research is needed. Focusing on one category of products at a time helps simplify the decision task considerably for our online shopper (particularly those less familiar with an online shopping environment), and given the strict constraints on time for an online experiment of this kind, we saw this as important to preserve data quality. One possible concern is that this may enhance compensation effects with shopping outside the laboratory. We recognize that further research on compensation effects with more comprehensive post experiment shopping diaries would be useful. Another possible concern is that, if consumers only see products from one product category, they are less likely to be distracted and more likely to be influenced by attributes of these products (price in particular) than in a natural setting. As subjects did the experiment online – whether from their homes, their offices or elsewhere –, it is not obvious that they would be distracted less than if, for example, they were to focus on buying cereals in the breakfast cereals aisle of a supermarket. It is also not clear that any resulting bias would be in the direction of paying more attention to the product attributes as opposed to introducing more noise in the data. Nevertheless, it would be interesting to introduce distractor tasks (e.g., Sitzia et al., 2015) in future research and see to what extent inattention affects shopping behavior.

An alternative simplification procedure would have been to have a ‘sparse set’, with only a few products available in each product category. However, and bearing in mind the habitual nature of consumption of many product categories (e.g., in relation to regular consumption of cereals), rather than solving the problem of unobserved compensation effects, it would have led to preference-based biases in demand patterns due to the specific products on sale (e.g., if Rice Krispies cereals are taxed but a healthier alternative which I would be comfortable buying is not on offer).

Focusing on our two product categories also helps with the interpretability of the results in terms of definition and therefore identification of healthy and unhealthy products. We designed the intervention to target nutrient profile and not individual macro-nutrients. Specifically, for both cereals and soft drinks we used the UK Food Standards Agency’s NP model to score and classify products into two types: healthier and less healthy, and tax them accordingly. In line with this, our main outcome variable focuses on the quantity of foods/drinks as classified using the aforementioned criteria. Though NP does not address all aspects of nutrition, it is more comprehensive than looking at a single macro-nutrient or at calories. Within composite foods healthier foods with lower nutrient profile scores may have lower content of undesirable nutrients (sugar, fat or calories), and they may have higher content of desirable nutrients (fiber, protein), and benefits offered by an intervention based on nutrient profile are likely spread across all macronutrients and not any one in particular. This is in line with a recognition that no single macronutrient is responsible for poor dietary health. That said, within both cereals and soft drinks, our nutrient profile does follow sugar content quite closely, making the interpretation of our results more straightforward also in this respect.

Finally, focusing on only two product categories helps us have enough statistical power to detect the heterogeneity of effects between product categories. This is important as it teaches the need of caution in generalizing the effectiveness of moderate taxes (such as 20%) across product categories, particularly where signposting is not implemented. In further data analysis (see Table A.11), we tried to better understand the source of the two key differences between breakfast cereals and soft drinks: namely, the greater price sensitivity of breakfast cereals and the fact that with soft drinks a 20% tax is only effective where signposted. Specifically, we separately considered the decision whether to buy (the extensive margin decision) from the decision of how much to buy conditional on buying (the intensive margin decision). We find that the greater price elasticity of cereals is evident at the extensive margin. In the intensive margin, with signposting, the reduction in consumption of soft drinks is comparable to that of cereals. The effect of signposting with a 20% tax on less healthy soft drinks is therefore largely driven by the intensive margin: the way signposting operates does not appear as one of drawing attention away from soft drinks. Intuitively, highlighting something is likely to draw attention to something rather than vice versa. However, if consumers want soft drinks, signposting *will* factor in the higher price more in reducing the amount they buy. Overall, the results are more consistent with a tax aversion mechanism than with an attentional mechanism for the effect of signposting.

Socioeconomic differences. Our results show no significant difference in the impacts of taxes between socioeconomic groups. This finding contrasts with a recent empirical estimate using observational, home scanner data in the United Kingdom, which finds that soda tax would be more effective for poorer individuals (Dubois et al., 2017). The fact that, as noted above, subjects appear partially to be treating the experimental task in isolation might be a reason for the discrepancy in findings. That said, we note that our finding is consistent with a modelling study to predict the effect of soda tax in the United Kingdom (Briggs et al., 2013) and with recent policy impact evaluations such as Nakamura et al. (2018; Chile) and Cawley et al. (2018; Philadelphia). Further research would obviously be helpful.

In contrast to Goldin and Homonoff's (2013) tobacco study, we also do not find significant differences in the effect of the signposting intervention across different socioeconomic groups. Therefore, our results do not support the claim that diet-related taxation (with or without signposting) would reduce socioeconomic inequalities in diet and related health outcomes. Neither though would our results give rise to the concern that such inequalities will increase. Taken literally, this would imply that taxation could improve overall population diet without having to incur the cost of widening socioeconomic inequalities in diet (and related health outcomes).

Policy implementation. A policy implementation of a tax on less healthy food and drinks would obviously need to make choices on which products should be taxed (whether that be based on the FSA Nutrition score or some other criterion). If the tax is based on a nutrient attribute, industry may modify the attribute to avoid it, i.e. pass-through should not be considered as automatic. Ito and Sallee (2018) found that this has happened in the market of cars, where a fuel economy tax is levied based on the size/weight of the car. Grogger (2017) looked at the tax pass-through of the Mexican soda tax, and found that, in the short run, the price of sodas actually increased by more than the amount of the tax.

A related policy question is the value of the optimal tax rate. The maximum level of the tax that has been considered by policy makers is 20%, and obviously there are good reasons not to tax more than is needed. Our finding of a nonlinear impact of taxation provides some support to a tax rate of 20% rather than a higher value, though this depends on the rate of tax pass-through as well as the use of signposting. Further research could look in a more nuanced way at the impact of tax rates up to 20% to identify where the greatest marginal gains in terms of tax rate increases are obtained.

REFERENCES

Action on Sugar, 2015. Cereals still stuffed with sugar. (Accessed August 21, 2018).

Allcott, H., Lockwood, B.B., Taubinsky, D., 2018. Should We Tax Soda? An Overview of Theory and Evidence. Working Paper. (Accessed December 11, 2018)

Andersen, S., Cox, J.C., Harrison, G.W., Lau, M.I., Rutstrom, E.E., Sadiraj, V., 2018. Asset integration and attitudes toward risk: Theory and evidence. Review of Economic Studies. 100(5), 816-830.

Andreyeva, T., Long, M.W., Brownell, K.D., 2010. The Impact of Food Prices on Consumption: A Systematic Review of Research on the Price Elasticity of Demand for Food. American Journal of Public Health. 100(2), 216.

Ball, K., McNaughton, S.A., Le, H.N., Gold, L., Mhurchu, C.N., Abbott, G., Pollard, C., Crawford, D., 2015. Influence of Price Discounts and Skill-Building Strategies on Purchase and Consumption of Healthy Food and Beverages: Outcomes of the Supermarket Healthy Eating for Life Randomized Controlled Trial. American Journal of Clinical Nutrition. 101(5), 1055-1064.

Berardi, N., Sevestre, P., Tépaut, M., Vigneron, A., 2016. The impact of a 'soda tax' on prices: evidence from French micro data. Applied Economics. 48(41): 3976-3994.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How Much Should We Trust Differences-in-Differences Estimates? The Quarterly Journal of Economics. 119(1), 249-275.

Blakely, T., Mhurchu, C.N., Jiang, Y., Matoe, L., Funaki-Tahifote, M., Eyles, H.C., Foster, R.H., McKenzie, S., Rodgers, A., 2011. Do Effects of Price Discounts and Nutrition Education on Food Purchases Vary by Ethnicity, Income and Education? Results from a Randomised, Controlled Trial. Journal of Epidemiology & Community Health. 65(10), 902-908.

Bonnet, C., Réquillart, V., 2013. Tax Incidence with Strategic Firms in the Soft Drink Market. Journal of Public Economics. 106(0), 77-88.

Briggs, A.D.M., Mytton, O.T., Kehlbacher, A., Tiffin, R., Rayner, M., Scarborough, P., 2013. Overall and income specific effect on prevalence of overweight and obesity of 20% sugar sweetened drink tax in UK: Econometric and Comparative Risk Assessment Modelling Study. British Medical Journal. 347, f6189.

Cameron, A.C., Gelbach, J.B., Miller, D.L., 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. The Review of Economics and Statistics. 90(3), 414-427.

Caro, J.C., Corvalán, C., Reyes, M., Silva, A ., Popkin, B., Taillie, L.S., 2018. Chile's 2014 sugar-sweetened beverage tax and changes in prices and purchases of sugar-sweetened beverages: An observational study in an urban environment. PLOS medicine. 15(7), e1002597.

Cawley, J., Thow, A.M., Wen, K., Frisvold, D., 2019. The Economics of Taxes on Sugar-Sweetened Beverages: A Review of the Effects on Prices, Sales, Cross-Border Shopping, and Consumption. Annual Review of Nutrition. doi: 10.1146/annurev-nutr-082018-124603Cawley, J., Frisvold, D., 2017. 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.

Cawley, J., Frisvold, D., Hill, A., Jones, D. 2018. The impact of the Philadelphia beverage tax on purchases and consumption by adults and children. NBER Working Paper 25052. National Bureau of Economic Research.

Charness, G., Gneezy, U., Kuhn, M., 2012. Experimental methods: Between-subject and within-subject design. Journal of Economic Behavior and Organization. 81: 1-8.

Chen, X., Kaiser, H.M., Rickard, B.J., 2015. The impacts of inclusive and exclusive taxes on healthy eating: An experimental study. Food Policy. 56, 13–24.

Chetty, R., Looney, A., Kroft, K., 2009. Salience and Taxation: Theory and Evidence. American Economic Review. 99(4), 1145-1177.

Colchero, M.A., Popkin, B.M., Rivera, J.A., Ng, S.W., 2016. Beverage Purchases from Stores in Mexico under the Excise Tax on Sugar Sweetened Beverages: Observational Study. British Medical Journal. 352, h6704.

Colchero, M.A., Rivera-Dommarco, J., Popkin, B.M., Ng, S.W., 2017. In Mexico, evidence of sustained consumer respose two years after implemeiting a sugar-sweetened beverage tax. Health Affairs. 36 (3), 564-571.

Colchero, M.A., Salgado, J.C., Unar-Munguia, M., Molina, M., Ng, S., Rivera-Donmarco, J.A., 2015. Changes in prices after an excise tax to sweetened sugar beverages was implemented in Mexico: Evidence from urban areas. PLOS One. 10(12): e0144408.

Cornelsen, L., Carreido, A., 2015. Health-Related Taxes on Foods and Beverages. Food Research Collaboration Policy Brief. (Accessed August 21, 2018).

Cornelsen, L., Green, R., Dangour, A., Smith, R., 2015. Why fat taxes won't make us thin. Journal of Public Health. 37(1), 18-23.

Cox, J.C., Sadiraj, V., 2006. Small- and large-stakes risk aversion: implications of concavity calibration for decision theory. Games and Economic Behavior. 56 (1), 45–60.

Darmon, N., Lacroix, A., Muller, L., Ruffieux, B., 2014. Food price policies improve diet quality while increasing socioeconomic inequalities in nutrition. International Journal of Behavioral Nutrition and Physical Activity. 11(66).

Dubois, P., Griffith, R., O'Connell, M., 2017. How well targeted are soda taxes? TSE Working Paper. 17-868 Toulouse School of Economics.

Epstein, L.H., Dearing, K.K., Paluch, R.A., Roemmich, J.N., Cho, D., 2007. Price and Maternal Obesity Influence Purchasing of Low- and High-Energy-Dense Foods. American Journal of Clinical Nutrition. 86(4), 914-922.

Epstein, L.H., Dearing, K.K., Roba, L.G., Finkelstein, E., 2010. The Influence of Taxes and Subsidies on Energy Purchased in an Experimental Purchasing Study. Psychological Science. 21, 406-414.

Epstein, L.H., Jankowiak. N., Nederkoorn, C., Raynor, H.A., French, S.A., Finkelstein, E., 2012. Experimental research on the relation between food price changes and food-purchasing patterns: a targeted review. American Journal of Clinical Nutrition. 95, 789-809

Eyles, H., Mhurchu, C.N., Nghiem, N., Blakely, T., 2012. Food Pricing Strategies, Population Diets, and Non-Communicable Disease: A Systematic Review of Simulation Studies. PLOS Medicine. 9(12), e1001353.

Fafchamps, M., Kebede, B., Zizzo, D.J., 2015. Keep up with the winners: Experimental evidence on risk taking, asset integration, and peer effects. European Economic Review. 79, 59-79.

Falbe, J., Thompson, H.R., Becker, C.M., Rojas, N., McCulloch, C.E., Madsen, K.A., 2016. Impact of the Berkeley excise tax on sugar-sweetened beverage consumption. American Journal of Public Health. 106(10), 1865-1871.

Feldman, N., Ruffle, B., 2015. The impact of including, adding, and subtracting a tax on demand. American Economic Journal: Economic Policy 7 (1), 95–118.

Février, P., Visser, M., 2004. A Study of Consumer Behavior Using Laboratory Data. Experimental Economics. 7(1), 93-114.

Fitt, E., Cole, D., Ziauddeen, N., Pell, D., Stickley, E., Harvey, A., Stephen, A.M., 2015. DINO (Diet In Nutrients Out) - an integrated dietary assessment system. Public Health Nutrition. 18, 234-241.

Forwood, S.E., Ahern, A.L., Marteau, T.M., Jebb, S.A., 2015. Offering within-Category Food Swaps to Reduce Energy Density of Food Purchases: A Study Using an Experimental Online Supermarket. International Journal of Behavioral Nutrition and Physical Activity. 12, 85.

Giesen, J.C., Payne, C.R., Havermans, R.C., Jansen.A., 2011. Exploring How Calorie Information and Taxes on High-Calorie Foods Influence Lunch Decisions. American Journal of Clinical Nutrition. 93(4), 689-694.

Goldin, J., Homonoff, T., 2013. Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity. American Economic Journal: Economic Policy 5(1), 302-336.

Green, R., Cornelsen, L., Dangour, A.D., Turner, R., Shankar, B., Mazzocchi, M., Smith, R.D., 2013. The effect of rising food prices on food consumption: systematic review with mata-regression. British Medical Journal. 346:f3703.

Griffith, R., O'Connell, M. and Smith, K., 2018. Corrective Taxation and Internalities from Food Consumption. CESifo Economic Studies, 64(1),1-14.

Grogger, J., 2017. Soda Taxes and the Prices of Sodas and Other Drinks: Evidence from Mexico. American Journal of Agricultural Economics 99 (2), 481-498.

Guthrie, J., Mancino, L., Lin, C.T., 2015. Nudging consumers towards better food choices: Policy approaches to changing food consumption behaviors. Psychology and Marketing. 32(5), 501–511.

Harrison, G.W., Lau, M.I., Williams, M.B., 2002. Estimating individual discount rates in Denmark: A field experiment. American Economic Review. 92(5), 1606-1617.

Hawkes, C., Smith, T.G., Jewell, J., Wardle, J., Hammond, R.A., Friel, S., Thow, A.M., Kain, J., 2015. Smart food policies for obesity prevention. The Lancet. 385, 2410–2421.

Hollands, G.J., Bignardi, G., Johnston, M., Kelly, M.P., Ogilvie, D., Petticrew, M., Prestwich, A., Shemilt, I., Sutton, S., Marteau, T.M., 2017. The TIPPME intervention typology for changing environments to change behaviour. Nature Human Behaviour. 1, 0140.

Hollands, G.J., Shemilt, I., Marteau, T.M., Jebb, S.A., Kelly, M.P., Nakamura, R., Suhrcke, M., Ogilvie, D., 2013. Altering Micro-Environments to Change Population Health Behaviour: Towards an Evidence Base for Choice Architecture Interventions. BMC Public Health. 13(1), 1218.

Ioannidis, J.P.A., 2013. Implausible Results in Human Nutrition Research. British Medical Journal. 347, f6698.

Ito, K., Sallee, J.M., 2018. The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards. Review of Economics and Statistics. 100 (2), 319-336.

Johnson, E.J., Shu, S.B., Dellaert, B.G.C., Fox, C., Goldstein, D.G., Häubl, G., Larrick, R.P., Payne, J.W., Peters, E., Schkade, D., Wansink, B., Weber, E.U., 2012. Beyond nudges: Tools of a choice architecture. Marketing Letters. 23, 487–504.

Just, D.R., 2017. The Behavioral welfare paradox: Practical, ethical and welfare implications of nudging. Agricultural and Resource Economics Review. 46(1), 1-20.

Lawlor, D.A., Davey Smith, G., Bruckdorfer, K.R., Kundu, D., Ebrahim,S., 2004. Those Confounded Vitamins: What Can We Learn from the Differences between Observational versus Randomised Trial Evidence? Lancet. 363(9422), 1724-1727.

Lobstein, T., Davies, S., 2009. Defining and Labelling 'Healthy' and 'Unhealthy' food. Public Health Nutrition. 12(3), 331-340.

Marteau, T.M., Ogilvie, D., Roland, M., Suhrcke, M., Kelly, M.P., 2011. Judging Nudging: Can Nudging Improve Population Health? BMJ. 342, d228.

McKevitt, F., 2015. Sainsbury's the Only One of the 'Big Four' Retailers to Increase Sales. (Accessed August 21, 2018).

Muller, L., Lacroix, A., Lusk, J.L., Ruffiex, B., 2017. Distributional Impacts of Fat Taxes and Thin Subsidies. Economic Journal. 127 2066-2092.

Mytton, O.T., Clarke, D., Rayner, M., 2012. Taxing Unhealthy Food and Drinks to Improve Health. British Medical Journal. 344, e2931.

Mytton, O., Gray, A., Rayner, M., Rutter, H., 2007. Could targeted food taxes improve health? Journal of Epidemiology & Community Health. 61(8), 689-694.

Nakamura, R., Mirelman, A.J., Cuadrado, C., Silva-Illanes, Dunstan, J., Suhrcke, M.E., 2018. Evaluating the 2014 sugar-sweetened beverage tax in Chile: An observational study in urban areas. PLOS Medicine. 15 (7): e1002596.

Nakamura, R., Suhrcke, M., Jebb, S.A., Pechey, R., Almiron-Roig, E., Marteau, T.M., 2015. Price promotions on healthier compared with less healthy foods: A hierarchical regression analysis of the impact on sales and social patterning of responses to promotions in Great Britain. American Journal of Clinical Nutrition. 101(4), 808-816.

Nevo, A., 2001. Measuring Market Power in the Ready-to-Eat Cereal Industry. Econometrica. 69(2), 307-342.

Powell, L.M., Chaloupka, F.J., 2009. Food Prices and Obesity: Evidence and Policy Implications for Taxes and Subsidies. Milbank Quarterly. 87(1), 229-257.

Powell, L.M., Chriqui, J.F., Khan, T., Wada, R., Chaloupka, F.J., 2013. Assessing the Potential Effectiveness of Food and Beverage Taxes and Subsidies for Improving Public Health: A Systematic Review of Prices, Demand and Body Weight Outcomes. Obesity Reviews. 14, 110-128. doi:10.1111/obr.12002.

Public Health England, 2015. Sugar Reduction: The Evidence for Action.

Rayner, M., Scarborough, P., Boxer, A., Stockley, L., 2005. Nutrient Profiles: Development of Final Model, London: Food Standards Agency.

Sharma, A., Hauck, K., Hollingsworth, B., Siciliani, L., 2014. The effects of taxing sugar-sweetened beverages across different income groups. Health Economics. 23(9), 1159-1184.

Schroeter, C., Lusk, J., Tyner, W., 2008. Determining the impact of food price and income changes on body weight. Journal of Health Economics. 27(1) 45-68.

Shemilt, I., Hollands, G.J., Marteau, T.M., Nakamura, R., Jebb, S.A., Kelly, M.P., Suhrcke, M., Ogilvie, D., 2013. Economic instruments for population diet and physical activity behaviour change: a systematic scoping review. PLOS One. 8(9), e75070.

Silver, L.D., Ng, S.W., Ryan-Ibarra, S., Taillie, L.S., Induni, M., Miles, D.R., Poti, J.M., Popkin, B.M., 2017. Changes in prices, sales, consumer spending, and beverage consumption one year after a tax on sugar-sweetened beverages in Berkeley, California, US: A before-and-after study. PLOS Medicine. 14(4), e1002283.

Sitzia, S., Zheng, J., Zizzo, D.J., 2015. Inattentive consumers in markets for services. Theory and Decision, 79(2), 307–332.

Smed, S., Jensen, J.D., Denver, S., 2007. Socio-economic characteristics and the effect of taxation as a health policy instrument. Food Policy. 32(5-6): 624-639.

Smith, R.D., Cornelsen, L., Quirmbach, D., Jebb, S.A., Marteau, T.M. (2018). Are sweet snacks more sensitive to price increases than sugar-sweetened beverages: analysis of British food purchase data. BMJ Open. 8(4), e019788.

Snowdon, C., 2015. A tax on sugar won't work, as the shipwreck of the Danish 'fat tax' shows. *Spectator*. (Accessed August 21, 2018).

Sugden, R., Wang, M., Zizzo, D.J., 2019. Take it or leave it: Experimental evidence on the effect of time-limited offers on consumer behaviour. Journal of Economic Behavior and Organization. 168, 1-23.

Sussman, A.B., Olivola, C.Y., 2011. Axe the Tax: Taxes Are Disliked More than Equivalent Costs. Journal of Marketing Research. 48(Special Issue), S91-S101.

Taubinsky, D., Rees-Jones, A., 2018. Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment. Review of Economic Studies. 85, 2462-2496.

Thaler, R.H., Sunstein, C.R., 2009. Nudge: Improving Decisions About Health, Wealth and Happiness. Penguin UK.

Thow, A.M., Downs, S., Jan, S., 2014. A Systematic Review of the Effectiveness of Food Taxes and Subsidies to Improve Diets: Understanding the Recent Evidence. Nutrition Reviews. 72(9), 551-565.

UK Government, 2016. Soft drinks industry levy. (Accessed August 21, 2018).

Wareham, N., Jebb, S., 2015. What is the evidence base for various fiscal measures? Cambridge: 2015. (Accessed August 21, 2018).

Waterlander, W.E., Steenhuis, I.H.M., de Boer, M.R., Schuit, A.J., Seidell, J.C., 2012a. The Effects of a 25% Discount on Fruits and Vegetables: Results of a Randomized Trial in a Three-Dimensional Web-Based Supermarket. International Journal of Behavioral Nutrition and Physical Activity. 9, 11.

Waterlander, W.E., Steenhuis, I.H.M., de Boer, M.R., Schuit, A.J., Seidell, J.C., 2012b. Introducing Taxes, Subsidies or Both: The Effects of Various Food Pricing Strategies in a Web-Based Supermarket Randomized Trial. Preventive Medicine. 54(5), 323-330.

Waterlander, W.E., Steenhuis, I.H.M., de Boer, M.R., Schuit, A.J., Seidell, J.C., 2013. Effects of Different Discount Levels on Healthy Products Coupled with a Healthy Choice Label, Special Offer Label or Both: Results from a Web-Based

Supermarket Experiment. International Journal of Behavioral Nutrition and Physical Activity. 10, 59.

WHO, 2016. Fiscal Policies for Diet and Prevention of Noncommunicable Diseases: Technical Meeting Report, 5-6 May 2015. (Accessed August 21, 2018).

Williamson, C.S., 2010. Breakfast Cereals – Why All the Bad Press? Facts behind the headlines. Nutrition Bulletin. 35(2010), 30-33

Wisdom, J., Downs, J.S., Loewenstein G., 2010. Promoting Healthy Choices: Information versus Convenience. American Economic Journal: Applied Economics, 2 (2): 164-78.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Wright, A., Smith, K.E., Hellowell, M., 2017. Policy lessons from health taxes: a systematic review of empirical studies. BMC Public Health. 17(1), 583.

Yang, C.C., Chiou, W.B., 2010. Substitution of Healthy for Unhealthy Beverages among College Students. A Health-Concerns and Behavioral-Economics Perspective. Appetite. 54(3), 512-516.

Zheng, Y., McLaughlin, E., Laiser, H.M., 2013. Taxing food and beverages: Theory, evidence, and policy. American Journal of Agricultural Economics 95, 705-723.

Zizzo, D.J., 2010. Experimenter Demand Effects in Economic Experiments. Experimental Economics. 13(1), 75-98.

FIGURE 1. PRICE MANIPULATION: NOT SIGNPOSTED VS. SIGNPOSTED

Notes: The signposting manipulation was between-subjects, and therefore participants would either see the tax-inclusive prices as in the first panel, or the tax identified separately and flagged as a tax as in the second panel.

FIGURE 2. MEAN QUANTITY PURCHASED OF LESS HEALTHY (HEALTHIER) PRODUCTS WHEN PRICES OF LESS HEALTHY (HEALTHIER) WERE MANIPULATED

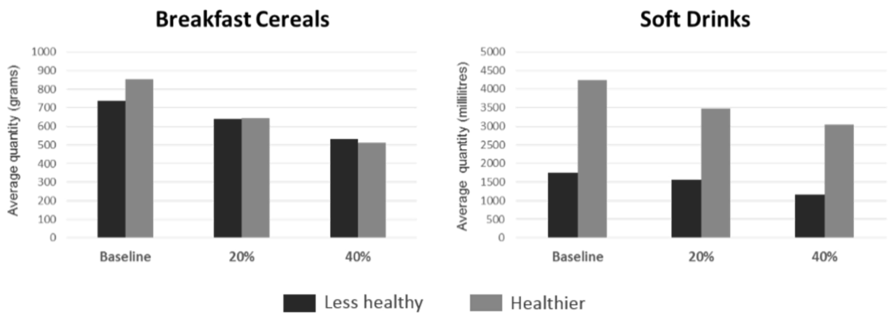
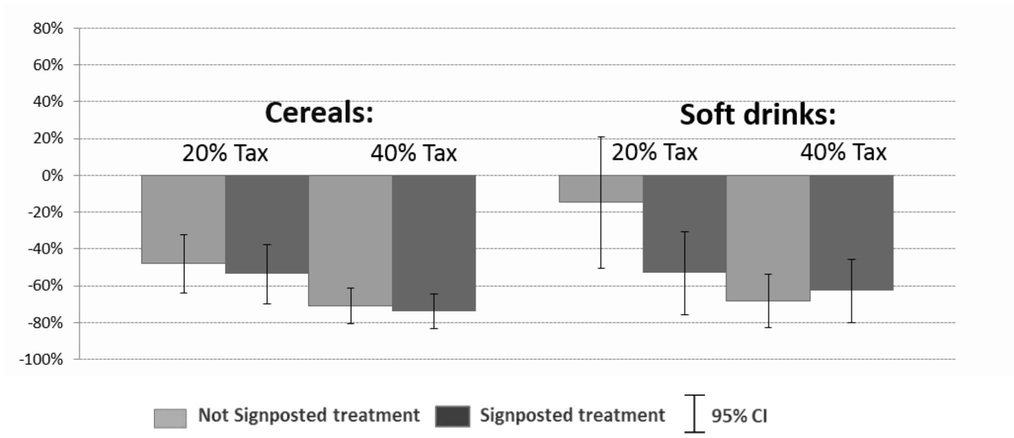


FIGURE 3. ESTIMATED TAX IMPACT ON THE VOLUME PURCHASED ON TAXED LESS HEALTHY PRODUCTS



Notes: Own-price effects that are estimated from the random effects Tobit regression models (see the regression models in Table 3). The bars show the average percentage change in the volume of purchases in each tax condition, as compared to the baseline condition without any price manipulations. Dark grey bars show the effects when tax was combined with the signposted treatment. 95% confidence intervals are presented in the figure and are based on block-bootstrap cluster robust standard errors at participant level.

TABLE 1. BASELINE CHARACTERISTICS OF PARTICIPANTS

			Treatment			
	All	n	Not Signposted	n	Signposted	n
Age, years	46.95	1000	46.72	506	47.18	494
Gender (F), %	54.10		53.75	506	54.45	494
BMI, kg/m^2	26.49	983	26.81	487	26.16	494
Education level		999		506		493
No qualifications, %	6.91		5.34		8.52	
<5 GCSEs/NVQ Level 1, %	15.42		15.81		15.01	
5 or more GCSEs/NVQ Level 2/1 A- level, %	19.62		20.75		18.46	
2 or more A-levels/NVQ Level 3, %	21.22		19.37		23.12	
Bachelor's degree, %	25.23		27.27		23.12	
Post-Graduate degree or qualification, %	11.61		11.46		11.76	

Household income		908		455		453
Up to £15,499 per year, %	19.71		18.46		20.97	
£15,500–£24,999 per year, %	23.46		25.05		21.85	
25,000–£39,999 per year, %	29.3		28.13		30.46	
£40,000 or more per year, %	27.53		28.35		26.71	
Household size, n	2.64	1000	2.68	506	2.59	494
Total participants, n		1000		506		494

TABLE 2. IMPACT OF TAXATION AND SIGNPOSTING ON THE LOG VOLUME OF LESS HEALTHY BREAKFAST CEREALS AND SOFT DRINKS PURCHASED (RANDOM EFFECT TOBIT MODEL).

	Breakfast cereals		**Soft drinks**	
	(1)	**(2)**	**(3)**	**(4)**
Tax Less-healthy 20%	-0.442***	-0.479***	-0.125	-0.147
	(0.080)	(0.081)	(0.170)	(0.183)
Interaction with signposted	-0.141	-0.109	-0.459***	-0.450**
	(0.192)	(0.214)	(0.159)	(0.179)
Tax Less-healthy 40%	-0.657***	-0.709***	-0.670***	-0.683***
	(0.055)	(0.050)	(0.072)	(0.075)
Interaction with signposted	-0.322*	-0.0945	0.0793	0.173
	(0.165)	(0.226)	(0.320)	(0.393)
Tax Healthier 20%	-0.117	-0.184*	0.127	0.240
	(0.111)	(0.112)	(0.224)	(0.265)
Interaction with signposted	0.368	0.601*	0.245	0.113
	(0.258)	(0.334)	(0.347)	(0.335)
Tax Healthier 40%	-0.0195	-0.0514	0.061	0.0476
	(0.148)	(0.147)	(0.21)	(0.223)
Interaction with signposted	0.412	0.609*	0.665	0.497
	(0.301)	(0.362)	(0.453)	(0.457)
Signposted	-0.423***	-0.444***	-0.207	-0.196
	(0.127)	(0.126)	(0.220)	(0.243)
Gender (male=1)		-0.368***		-0.291*
		(0.122)		(0.171)
Age		-0.0232***		-0.0151*
		(0.00681)		(0.00849)
Log BMI		-0.177		-0.653
		(0.579)		(0.704)
In a weight loss diet (yes=1)		0.0529		-0.464***
		(0.221)		(0.148)
Online grocery shopping frequency		0.0256		0.0693
		(0.0572)		(0.0669)
Tesco shopping frequency		0.0678		-0.0170
		(0.0588)		(0.0692)

Time preferences (discount)		0.0305 (0.0271)		-0.0118 (0.0331)
Education level		-0.0411 (0.0725)		-0.153* (0.0873)
Household Income		0.135 (0.0918)		-0.201* (0.117)
Stocks of cereals (days)		-0.0264** (0.0106)		
Stocks of soft drinks (days)				-0.0310** (0.0144)
Accuracy guessing Cereal prices		0.187 (0.253)		
Accuracy guessing soft drink prices				-2.322** (1.100)
Observations	5000	4265	5000	4265
Number of subjects	1000	853	1000	853

Notes: The analysis is based on a random effect Tobit model where each choice made by a subject provides an observation. Partial effects on the observed volume of log purchases are presented. The estimated partial effects are also transformed to represent the proportionate effects. Subject-level cluster-robust standard errors in parentheses (block-bootstrapped, 1000 replications). The number of observations corresponds to five shopping tasks and is smaller in the regressions with controls because of omitted answers to control questions. *** p<0.01, ** p<0.05, * p<0.1.

TABLE 3. ESTIMATED PRICE EFFECTS ON THE LOG VOLUME OF LESS HEALTHY AND HEALTHIER BREAKFAST CEREALS AND SOFT DRINKS BY TREATMENT

	Breakfast cereals		**Soft drinks**	
	Less healthy	**Healthier**	**Less healthy**	**Healthier**
Tax Less-healthy 20%	-0.479*** (0.081)	1.060** (0.427)	-0.147 (0.183)	-0.242* (0.141)
Tax Less-healthy 20% in signposted intervention	-0.536*** (0.083)	0.0821 (0.191)	-0.531*** (0.114)	0.142 (0.235)
Tax Less-healthy 40%	-0.709*** (0.050)	1.543*** (0.534)	-0.683*** (0.075)	-0.064 (0.213)
Tax Less-healthy 40% in signposted intervention	-0.736*** (0.048)	0.221 (0.230)	-0.628*** (0.087)	0.365 (0.247)
Tax Healthier 20%	-0.184* (0.112)	0.103 (0.220)	0.240 (0.265)	-0.515*** (0.103)
Tax Healthier 20% in signposted intervention	0.306 (0.201)	-0.560*** (0.084)	0.38 (0.314)	-0.365** (0.146)
Tax Healthier 40%	-0.0514 (0.147)	-0.226 (0.163)	0.048 (0.223)	-0.503*** (0.102)
Tax Healthier 40% in in signposted intervention	0.526** (0.244)	-0.741*** (0.054)	0.568* (0.335)	-0.675*** (0.078)
Control variables	Yes	Yes	Yes	Yes
Observations	4265	4265	4265	4265

Number of subjects	853	853	853	853

Notes: The results are based on models (2) and (4), presented in Table 2. Partial effects on the observed volume of log purchases are presented. The estimated partial effects are also transformed to represent the proportionate effects. *** p<0.01, ** p<0.05, * p<0.1.