Should We Tax Soda?
An Overview of Theory and Evidence

Hunt Allcott, Benjamin B. Lockwood, and Dmitry Taubinsky*

August 8, 2018

Abstract

Taxes on sugar-sweetened beverages (SSBs) are growing in popularity and have generated an active public debate. Are they a good idea? If so, how high should they be? Are such taxes regressive? Americans and some others around the world consume a remarkable amount of SSBs, and the evidence suggests that this generates significant health costs. Building on recent work by Allcott, Lockwood, and Taubinsky (2018) and others, we review the basic economic principles for an optimal sin tax on SSBs. The optimal tax depends on (1) *externalities*: uninternalized costs to the health system from SSB consumption; (2) *internalities*: costs consumers impose on themselves by overconsuming sweetened beverages due to poor nutrition knowledge or lack of self-control; and (3) *regressivity*: how much the financial burden and the internality benefits from the tax fall on the poor. We then summarize the empirical evidence on the key parameters that determine how large the tax should be, which suggests that SSB taxes can be welfare enhancing. We end with seven concrete suggestions for policymakers considering an SSB tax.
Introduction

Sin taxes are imposed to discourage activities that are deemed harmful, either for the actor or for others in society. Such taxes have a long history, but as the understanding of harm has evolved over time, so too has the class of goods to which sin taxes are commonly applied.

This paper focuses on a rapidly proliferating class of sin taxes: those on sugar sweetened beverages (SSBs), sometimes called “soda taxes,” although they typically target more than just carbonated soft drinks. As of mid-2018, seven U.S. cities and 34 countries around the world have implemented SSB taxes, mostly in the past few years (GFRP 2018). These policies have been spurred in part by the rise of sugar-related health conditions, including obesity, diabetes, and heart disease. Proponents point to a range of policy goals, including improving public health, reducing budget deficits, funding social programs, and raising revenues from SSB-producing corporations. There is also resistance to sin taxes, often on the grounds that SSBs are consumed most heavily by the poor, making the taxes regressive.

The goal of this paper is to provide an economic framework for understanding SSB taxes. Although SSB consumption can have many adverse consequences, which we summarize in the first part of the paper, standard economic frameworks do not justify a tax on soda in the absence of uninternalized externalities or “internalities”—harms that individuals impose on themselves due to behavioral biases. In the second part of the paper, we therefore draw on our recent work in Allcott, Lockwood, and Taubinsky (2018) to summarize the basic economic principles to determine the optimal size of SSB taxes. We discuss how internalities, externalities, the price elasticity of demand, distributional concerns, and the incidence on producers all shape the optimal SSB tax. In the third part, we summarize the growing empirical literature that measures these key parameters.

We end with seven concrete suggestions for policymakers. First, tax grams of sugar, not ounces of soda. Second, focus on counteracting externalities and internalities, not on minimizing SSB consumption. Third, design soda taxes to reduce consumption among people generating the largest externalities and internalities. Fourth, make taxes consistent across geographic boundaries. Fifth, use caution when pre-allocating tax revenues. Sixth, when judging regressivity, consider internality benefits, not just who pays the taxes. Finally, our read of the evidence is that taxing soda is probably a good idea.

Background: SSB Consumption and Health Harms

SSB Consumption Patterns

Americans consume a remarkable amount of calories from sugary drinks. A typical 12-ounce soft drink might contain 35-40 grams of sugar and about 140 calories, representing about seven percent of benchmark 2000-calorie diet. Using data from the National Health and Nutrition Examination Survey (NHANES) in 2013-2014, we calculate that the average American adult (aged 18 or older) consumes 157 calories per day from sugar-sweetened beverages, comprising 7.1 percent of calorie intake. (We define SSBs to include any beverages with caloric sweeteners, including carbonated soft drinks, sports drinks, energy drinks, fruit drinks, milk-based drinks, and coffee and tea with added sweeteners, but not 100% fruit juice or “diet” drinks with low-calorie or zero-calorie sweeteners.) Almost all of these calories are from added sugars. As a benchmark, the U.S. Dietary Guidelines recommend limiting added sugars from all food and drinks to no more than 10 percent of total calorie intake, or around 200 calories per day, while the World Health Organization is even more conservative. SSBs comprise 47 percent of the average American’s added sugar consumption (U.S.
Figure 1: Sugar-Sweetened Beverage Consumption by Income

![Graph showing sugar-sweetened beverage consumption by income for 2013-2014 using data from the National Health and Nutrition Examination Survey.]

Notes: This figure shows sugar-sweetened beverage consumption by household income for 2013-2014 using data from the National Health and Nutrition Examination Survey.

Many people consume SSBs: 50 percent of American adults consume at least one SSB on any given day. However, consumption amounts vary considerably across demographic groups. Figure 1 plots SSB consumption across the income distribution. People with household income below $25,000 per year consume 210 calories per day of SSBs, while people with household income above $75,000 per year consume only 106 calories per day. This underscores the concern that SSB taxes could be regressive.

Perhaps due to rising public awareness of the health effects of SSBs, consumption is falling over time in the U.S. and many other industrialized countries. In the NHANES data, the average American consumed 205 calories per day from sugar-sweetened beverages in 2003-2004, against the 157 calories in 2013-2014. Popkin and Hawkes (2016) find that SSB calorie consumption per capita declined from 2009-2014 in North America, Australasia, and Western Europe, but increased in the rest of the world. North Americans consume 3-4 times more calories from SSBs than the world average.

Harms from SSB Consumption

Figure 2 illustrates the pathways through which sugary drinks harm human health and impose private and social costs. SSB consumption harms health through three main channels: diabetes, cardiovascular disease, and weight gain, although there are also additional health effects such as tooth decay that we do not discuss here. For each of the main channels, we briefly discuss evidence on the magnitude of the effects, with the caveat that much of this evidence is from non-randomized epidemiological studies. Although these studies...
do their best to measure and control for confounding variables, unmeasured factors like eating and exercise habits and social conditions could mean that these conditional correlations are biased estimates of the causal relationship between SSB consumption and health outcomes.

The first main health harm is diabetes. SSBs have high “glycemic loads,” meaning that they contain large amounts of rapidly digestible sugars. (Sugars are digested more quickly when they come from drinks than when they come from foods.) When high glycemic load foods are digested, they prompt a quick release of glucose into the blood stream and the secretion of a corresponding amount of insulin in response. Over time, these states of elevated blood glucose (hyperglycemia) and insulin (hyperinsulinemia) can cause insulin resistance, often a precursor to diabetes (Ludwig, 2002; Janssens, Shapira, Debeuf, Michiels, Putman, Bruckers, Renard, and Molenberghs, 1999; Raben, Møller, Flint, Vasilaras, Christina Møller, Juul Holst, and Astrup, 2011). A meta-analysis of 11 cohort studies found that people who drink one or more SSB servings per day have a 26 percent higher risk of developing diabetes than those who drink less than one per month (Malik, Popkin, Bray, Després, Willett, and Hu, 2010).

The second main health harm is cardiovascular disease. Diets high in refined carbohydrates, including added sugars, can increase one’s risk of coronary heart disease (CHD) by increasing blood pressure and causing abnormal cholesterol and triglyceride levels, a state known as dyslipidemia. High-carbohydrate diets have been found to be associated with higher systolic and diastolic blood pressure, higher triglyceride levels, lower “good cholesterol” (HDL), and higher “bad cholesterol” (LDL), even when weight is held constant (Appel, Sacks, Carey, Obarzanek, Swain, Miller, Conlin, Erlinger, Rosner, and Laranjo, 2005; Siri-Tarino, Sun, Hu, and Krauss, 2010; Santos, Esteves, da Costa Pereira, Yancy Jr, and Nunes, 2012; DiNicolantonio, Lucan, and O’Keefe, 2016; Welsh, Sharma, Abramson, Vaccarino, Gillespie, and Vos, 2010; Te Morenga, Howatson, Jones, and Mann, 2014). A meta-analysis of four studies found that consuming one additional SSB per day is associated with 17 percent higher risk of CHD (Xi, Huang, Reilly, Li, Zheng, Barrio-Lopez, Martinez-Gonzalez, and Zhou, 2015).1

The third main health harm is weight gain. SSBs contain calories, but because these calories come in liquid form, they do not make people feel satiated. Experimental studies show that when people consume calories from solid foods instead of liquids (e.g. jelly beans instead of soda, or cheese instead of milk), they eat less later in the day, resulting in significantly lower overall calorie intake (DiMeglio and Mattes, 2000; Mourao, Bressan, Campbell, and Mattes, 2007). Other experiments have found that when participants are provided with the same foods and either caloric or non-caloric beverages, they consume the same amount of

---

1Some additional studies have focused on the association between SSB consumption and hypertension, a key risk factor for CHD. A meta-analysis of six cohort studies found that heavy SSB consumers (people who drink one or more servings per day have a 12 percent higher risk of developing hypertension compared to those who never consumed SSBs (risk ratio=1.12; 95% CI: 1.06, 1.17) (Jayalath, de Souza, Ha, Mirrahimi, Blanco- Mejia, Di Buono, Jenkins, Leiter, Wolever, and Beyene, 2015).
food regardless of the beverage provided and report no difference in feelings of fullness between conditions (DellaValle, Roe, and Rolls, 2005; Flood, Roe, and Rolls, 2006). This excess calorie consumption causes weight gain, which has a variety of costs. Weight gain is thought to independently affect diabetes and cardiovascular disease, and it mediates the statistical relationships reported above between SSB consumption and those conditions (e.g. Schulze, Manson, Ludwig, Colditz, Stampfer, Willett, and Hu, 2004; Fung, Malik, Rexrode, Manson, Willett, and Hu, 2009).

Both experimental and observational evidence suggests that SSB consumption causes weight gain. For example, Ebbeling, Feldman, Chomitz, Antonelli, Gortmaker, Osganian, and Ludwig (2012) randomized 224 overweight adolescents in Boston who regularly consumed SSBs to receive either home deliveries of non-caloric drinks for one year with instructions not to buy additional SSBs, or to receive supermarket gift cards and no instructions. The diet drink group consumed less calories and after one year weighed 1.9 kilograms less. de Ruyter, Olthof, Seidell, and Katan (2012) randomized 641 children aged 6-12 in Amsterdam who regularly consumed SSBs to receive either SSBs or diet drinks every day for 18 months. While there was no difference in the number of drinks consumed by each group, the diet drink group weighed about one kilogram less. In observational analysis of three cohort studies, Mozaffarian, Hao, Rimm, Willett, and Hu (2011) find that one serving per day additional SSB consumption is conditionally associated with weight gain of one pound per four-year follow-up period, controlling for a variety of biological and lifestyle factors.2

The public health community has focused on taxing sugary drinks instead of taxing sugar in foods because, for the reasons described above, sugar consumed through beverages is thought to be more harmful than the same amount of sugar consumed through foods. There is less evidence that added sugars are significantly more harmful than natural sugars in drinks, although 100% fruit juice has typically been excluded from sugary drink taxes.

Quantifying Costs

As illustrated in Figure 2, the diseases caused by SSBs impose several kinds of costs. First, they can reduce income, through missed work hours, lower productivity, and/or labor market discrimination (e.g. Cawley, 2004). Second, they can impose non-financial costs—in the extreme, early death (Bhattacharya and Sood, 2011). Third, they can increase health care costs.

By combining estimates of the price elasticity of SSB demand, the effect of SSBs on diabetes, cardiovascular disease, and obesity, and the costs of treating these diseases, it is possible to estimate the effects of an SSB tax on health care costs. Wang, Coxson, Shen, Goldman, and Bibbins-Domingo (2012) estimate that over 10 years, a one cent per ounce tax would save $17.1 billion in health care costs, while Long, Gortmaker, Ward, Resch, Moodie, Sacks, Swinburn, Carter, and Wang (2015) estimate the 10-year savings to be $23.6 billion. Of course, these estimates are incomplete measures of social welfare. We will want to separate these costs into private costs versus externalities, as indicated in Figure 2, and trade them off against the benefits of SSB consumption. The next section provides a framework for doing this.

---

2 See Mattes, Shikany, Kaiser, and Allison (2011) for a summary of additional randomized experiments, and see Malik, Schulze, and Hu (2006); Vartanian, Schwartz, and Brownell (2007); Malik, Willett, and Hu (2009); Malik, Pan, Willett, and Hu (2013) for meta-analyses and systematic review of other cohort studies on weight gain.
Some Simple Economics of Soda Taxes

The economic logic behind an SSB tax begins builds from the classic principles of externality-correcting taxes (Pigou, 1920): if a good has harmful effects that aren’t considered by its consumers, then in an unregulated market people will consume too much of it. Thus, a tax can raise welfare by reducing consumption toward the efficient level at which marginal social cost equals marginal social benefit.

Externalities are an important part of the rationale for SSB taxes, since some costs from the adverse health effects are externalized through health insurance, whether public or private. Additionally, a growing body of research in behavioral economics indicates that people sometimes ignore harmful or beneficial effects to themselves—for example, because they are misinformed, or because those consequences lie in the distant future. These costs are sometimes called “internalities,” and we view their presence as a key distinction of “sin taxes” on goods like soda, cigarettes, and alcohol.

Although internality and externality costs operate somewhat similarly, there are important differences between the two, and we consider each in turn. We will use Figure 3, which illustrates the effect of an SSB tax on demand from a single consumer, to discuss both concepts.

Welfare Effects Due to Externalities

Focusing first on the case of pure externalities, $D_1$ plots the individual’s demand curve for SSBs at various prices (or, equivalently, the consumer’s marginal private benefit from soda at each quantity). The vertical distance $b$ represents the per-unit externality cost, so that $D_2$ plots the marginal social benefit, net of externalities, as a function of soda quantity consumed. A tax that raises the price from $p_0$ to $p_t$ then has three distinct effects on welfare. The area $A = t \times q_t$ is transferred from the consumer to the government, in the form of tax revenue. The area $C = \Delta q \times t/2$ represents a further decrease in the consumer’s welfare from foregone soda consumption due to the tax. (Here $\Delta q \approx \frac{dD_1}{dp} \times t$ is the reduction in soda consumption due to the tax.) The area $B + C = \Delta q \times b$, represents an increase in welfare for the externality bearer. In the context of soda, a natural benchmark assumption is that the externality reduction accrues to the government’s budget (in present value terms), for example due to reduced Medicare expenditures on treatments for conditions such as heart disease and diabetes. Therefore the net effect of the tax is twofold: a transfer of $A + C$ from the consumer to the government, and a further increase in government funds of $B$.

The total welfare effects of a soda tax depend on aggregating these components across individuals. Since the tax involves transfers between parties, something must be assumed about the social value of resources in the hands of the government relative to consumers, and across consumers of different types. A common assumption is that the marginal utility from consumption is decreasing with consumers’ incomes—a common justification for progressive income tax schedules. A simple way to capture such distributional implications is to assign “social marginal welfare weights” (e.g., Saez and Stantcheva, 2016) to different households depending on their income (or, possibly, other attributes), so that a weight of, say, 1.5 on household $x$ implies that society places the same value on $\$1$ in the hands of household $x$ as on $\$1.50$ in the hands of the government. Then the transfer $A + C$ from the consumer to the government generates a net a social gain if the weight of on the consumer in question is less than 1, and a social loss otherwise.

Putting these pieces together, we must aggregate these effects by summing the externality benefit $B$ and the transfer $A + C$ across consumers, weighted appropriately. The area $B$ scales with its width (proportional to the individual elasticity of soda demand) multiplied by its height (the externalized health costs from soda consumption). Therefore this benefit depends positively on the aggregate demand elasticity for soda, the
average externality across all consumers, and the covariance between demand elasticity and externality cost across individuals. Intuitively, if a tax causes a larger drop in SSB demand among individuals who have larger health costs from SSB consumption, then the welfare benefit from the tax is greater.

The transfer $A + C$ has the same height for all consumers ($p_t - p_0$), but its width depends on the quantity of soda consumed by each consumer. Moreover, this summation across consumers is weighted by the difference between their welfare weight and the value of public funds. As such, the sign of this welfare effect can be either positive or negative, but will tend to be negative if poorer consumers (with high welfare weights) tend to purchase more of the externality-producing good, as is the case for soda.\(^3\) This effect formalizes the common concern a downside of soda taxes is their regressive incidence.

### Welfare Effects Due to Internalities

Now suppose SSBs generate only internality costs. We can reinterpret Figure 3, with $D_1$ representing the consumer’s observed demand curve, and $D_2$ representing the latent demand curve that would arise if consumers did not suffer from internalities. Then the vertical distance $b$ represents an ignored internality cost, measured in money units.

Internalities operate similarly to externalities, with a crucial difference: the area $B + C$ accrues to the consumer, rather than to the government (or to the externality bearer, more generally). This does not change the interpretation of the transfer $A$, from the individual to the government, which will again be negative if poorer consumers purchase more SSBs. However, it does change the interpretation of $B$, which (unlike in

\(^3\)Note that the welfare effect of this transfer depends on the level of soda consumption across the income distribution, and not on soda consumption as a share of consumers’ income.
the case of externalities) is multiplied by the individual’s welfare weight.\(^4\) As a result, the welfare benefits from the tax are larger to the extent that poorer consumers generate larger internalities, and therefore larger areas of \(B\). Similarly, if poor consumers reduce their consumption relatively more in response to the tax (higher \(\Delta q\)) then that also leads poorer consumers to have larger areas \(B\).\(^5\) These relationships illustrate an important conceptual difference between internalities and externalities, and appear to be quantitatively important in the context of SSB consumption. The relationship between internalities, elasticities, and welfare weights is derived formally (and empirically measured) in Allcott, Lockwood, and Taubinsky (2018).

**Putting it Together**

In a context with both externalities and internalities, one must add the externality and (welfare-weighted) internality benefits, netted against any welfare effects due to the transfer of resources from consumers to the government. Externality benefits depend (positively) on the aggregate elasticity of SSB demand, the average externalized health cost from consumption, and their covariance. Internality benefits similarly depend on the aggregate elasticity and average uninternalized health costs, and their covariance, as well as the extent to which uninternalized health costs and soda demand responses are higher among poor consumers. Finally, the welfare cost of the resource transfer is larger to the extent that soda consumption is higher among poor households.

**Are Soda Taxes “Regressive”?**

A common concern about soda taxes is that they may hurt poor households, since low earners tend to purchase more soda. We formally study the theory of regressive sin taxes, and estimate the welfare costs of a soda across the income distribution, in Allcott, Lockwood, and Taubinsky (2018). But the concepts of externalities, internalities, and transfers from Figure 3 illustrate the basic forces at work.

To understand who is helped and hurt by a soda tax, we need to draw a distinction between who pays the most in taxes and who is benefitted or harmed, all things considered. While it is true that poorer consumers will pay more in soda taxes on average (due to their disproportionate soda consumption), if there are internality costs from drinking soda, the beneficial reductions of health conditions like heart disease and diabetes will also accrue to low income households, as highlighted by Gruber and Köszegi (2004). In terms of Figure 3, although poorer consumers incur more costs due to area \(A\) on average, those may be offset (partially, or more-than-fully) by the gained area \(B\). As a result, the fact that poorer consumers purchase more soda does not imply they are made worse off by the tax. Indeed, the findings in Allcott, Lockwood, and Taubinsky (2018) suggest that the net benefits of a soda tax are reasonably flat across the income distribution, and possibly highest for the poorest consumers.

**Welfare Effects on SSB Suppliers**

The exposition so far accounts only for the consumer side of the market, and therefore leaves out two key issues: the question of tax pass-through (what portion of the tax is borne by consumers in the form of a price increase), and producer surplus. To illustrate these forces, Figure 4 depicts a simple supply and demand model of the SSB market. \(D^m\) represents observed market demand for SSBs, while \(b^m\) represents

\(^4\)In the context of internalities, the area \(C\) can be regarded as a transfer from the consumer to herself, and can be ignored.

\(^5\)If the elasticity of demand is constant across consumers, the demand response \(\Delta q\) will be higher for those who consume more.
the average marginal externality (weighted by elasticities of demand) plus average marginal internality, (weighted by elasticities of demand and welfare weights), so that $D^m_2$ represents market demand less the uninternalized social cost of consumption (normalized by the marginal value of public funds) at each quantity. (The pictured tax is a little lower than the optimal level, $b^m$.)

In a simple model like that illustrated in Figure 4, the conventional explanation for incomplete tax pass-through is that some of the tax incidence falls on producers, rather than consumers. To account for this possibility, we allow that market supply $S$ may slope upward. The share of the tax that is passed through to consumers is $\frac{p^m - p_0}{1}$—a quantity which rises with the elasticity of soda supply and falls with the (absolute) elasticity of soda demand. The tax then has three distinct effects on welfare—a transfer from producer surplus to the government, represented by the vertically-hatched area $X$, a transfer from consumers to the government, represented by the horizontally-hatched area $Y$, and a beneficial reduction in externalities and internalities (now combined), represented by the angle-hatched area $Z$. Relative to a model with infinitely elastic soda supply (corresponding to full pass-through to consumers), the key difference is that some of the costs of the tax are borne by producers, rather than consumers. If marginal resources are valued equally in the hands of SSB producers and (welfare-weighted) SSB consumers, the issue of pass-through is irrelevant—the tax should be adjusted maximize the welfare gain from the internality and externality benefit $Z$, and the weighted transfer of resources $X + Y$. On the other hand, if resources are valued more in the hands of SSB consumers than SSB producers—for example, if marginal resources accrue to firm shareholders who have a lower average welfare weight then SSB consumers—then a lower pass-through will imply a larger net welfare benefit from the tax, and a higher tax at the optimum. Conversely, if a higher welfare weight is placed on producers, then partial pass-through calls for a lower soda optimal tax.

Other explanations for partial pass-through—such as discrete pricing policies by grocers or an inability to separately price regular and diet soda fountain sales at fast-food restaurants—might generate different implications. In particular, if a portion of the tax is absorbed by producers with no reduction in quantity supplied, then the optimal tax may need to be larger than $b^m$ in order to achieve the efficient reduction in soda consumption. However, this possibility depends on understanding the reason for partial pass-through, in addition to quantifying the pass-through rate itself.

Further Considerations

In addition to the mechanics discussed thus far, several additional issues may affect the welfare impacts of an SSB tax.

First, when consumers reduce their SSB consumption due to the tax, they may also raise or lower consumption of other (untaxed) sugary goods. To the extent that they do, the resulting change in externalities and internalities from those goods should be considered when setting the tax on SSBs. The sign of this effect is ambiguous. For example, consumers may view sugary snacks as a substitute for SSBs—an alternative way to get a desired “sugar kick”—in which case some of the internality and externality reductions from an SSB tax may be offset by increased internalities and externalities from substitution to other sugary goods. As discussed earlier, however, SSBs may have particularly large health damages per gram of sugar; the relevant statistic depends on the marginal internality and externality on the goods that people substitute to, not amount of sugar in these substitutes. On the other hand, the SSBs and unhealthy foods may be complements if consumers tend to purchase or consume such snacks together—then the analysis above will underestimate the benefits of a soda tax.

Second, the benefits of a soda tax may be further affected by behavioral adjustments that affect tax
revenues in other domains. For example, economic theory predicts that taxing “normal goods” (i.e., goods whose consumption increases when consumers get additional income) will reduce the appeal of earning income generally, creating a distortion that lowers income tax revenue. This highlights the importance of distinguishing between causal income effects (which imply that commodity taxes distort labor supply decisions) and between-income preference heterogeneity (which do not)—an issue we formalize and quantify in Allcott, Lockwood, and Taubinsky (2018). More generally, to the extent that SSB consumption (and the resulting expected health consequences) affect labor supply patterns such as retirement age or disability insurance take-up, the resulting changes in income tax revenues should be incorporated when quantifying the welfare effects of a soda tax.

### Empirical Estimates of Key Parameters

In this section, we review the empirical estimates of the key parameters identified in the theory, with an eye to the strengths and weaknesses of different estimation strategies.

#### Demand Elasticities

Perhaps the simplest form of demand estimation is a regression of the natural log of quantity consumed on the natural log of price, often controlling for some additional variables. The coefficient on price is the demand elasticity. For any product—not just SSBs—demand estimation requires a number of considerations. For example, to what extent should we aggregate similar products into groups? How should we address stockpiling: when the price is low in period \( t \), people buy more and store it for use in future periods,
affecting future demand? If there are some periods with zero purchases, how should we model this censored demand? How should researchers parameterize substitution patterns across goods?

Two empirical challenges are particularly important when estimating SSB demand. First, the standard datasets either do not provide complete measures of SSB consumption or do not have plausibly exogenous price variation. One common type of dataset is household-level scanner data, such as the U.S. National Consumer Panel (also known as Homescan) or Kantar Worldpanel. Households in these panels are asked to scan the bar codes of all groceries that they bring home, but they do not record “away from home” consumption, such as purchases at restaurants, vending machines, and ballparks. If soda taxes are imposed on away from home consumption, then the parameter of interest is the elasticity of demand for all SSBs, including away from home consumption. The demand elasticity for “at-home” consumption may not generalize if away-from-home purchases are more or less elastic, and there may also be bias due to substitution, if households respond to higher grocery prices by making more away from home purchases.

A second common type of dataset is self-reported consumption from dietary recall studies such as NHANES, in which people record food and drink consumed over the past 24 hours or some other period. Self-reports may have more measurement error and do not track the same individuals over time, making it difficult to use the type of strategies detailed below to exploit exogenous price variation. Various papers (e.g. Silver, Ng, Ryan-Ibarra, Taillie, Induni, Miles, Poti, and Popkin 2017; Allcott, Lockwood, and Taubinsky 2018) use combinations of these two types of datasets. Dubois, Griffith, and O’Connell 2017 have an unusual dataset that includes smartphone-enabled scanner data measuring “on-the-go” purchases of packaged SSBs, although this does not cover grocery or restaurant purchases.

The second particularly important empirical challenge is isolating quasi-random variation in prices. Simultaneity bias is a standard problem in demand estimation, in which unobserved demand shifters affect prices. Furthermore, measurement error in prices can attenuate the estimated response of purchases to prices, as Einav, Leibtag, and Nevo (2010) demonstrate in the Homescan data. Both of these forms of omitted variables bias can make demand appear more inelastic than it actually is.

The literature has used two main identification strategies to isolate quasi-random variation. The first is to include a large set of fixed effects in an attempt to control for demand shifters and thereby isolate quasi-random price variation. For example, Dubois, Griffith, and O’Connell (2017) include brand, time, and other fixed effects, thereby identifying the demand elasticity only off of variation in prices of the same product across retailers and variation in the slope of non-linear pricing (the relative prices of small vs. large containers) across brands. This approach does not address possible measurement error, and in different situations it may not be clear whether the fixed effects fully control for demand shifters that affect prices.

The second approach to isolating quasi-random price variation is to instrument for price. For example, a standard set of instruments for the price of SSBs (or other goods) in a given city at a particular time is the average price of the good at that time in all other cities around the country (Hausman, 1996; Nevo, 2001). There is active debate about these instruments (e.g. Bresnahan, 1996), as they require that demand shocks are uncorrelated across markets, which rules out the possibility of national-level advertising campaigns, public health information, or other types of national-level taste variation. Allcott, Lockwood, and Taubinsky (2018) instrument for price of each household’s UPCs with the deviation from national average price at the retailers where the household shops, outside the household’s county. This affords more power than the Hausman (1996) instruments, because much of the variation in prices is from retailer-specific promotions. Finkelstein, Zhen, Bilger, Nonnemaker, Farooqui, and Todd (2013) instrument for a household’s price paid with prices paid by other households in the same city and quarter, excluding households living in the household’s Census
tract. While such IV strategies can address measurement error, in different situations it may not be clear whether the instruments isolate price variation that is not driven by responses to demand shifters.

Andreyeva, Long, and Brownell (2010) and Powell, Chriqui, Khan, Wada, and Chaloupka (2013) review the literature estimating the aggregate price elasticity of demand for SSBs. Andreyeva, Long, and Brownell (2010) report that across 14 studies, the mean price elasticity is -0.79, with a range from -0.13 to -3.18. Powell, Chriqui, Khan, Wada, and Chaloupka (2013) review 12 studies, finding a mean price elasticity of -1.21, with range from -0.71 to -3.87. These wide ranges suggest that the literature is unsettled. Furthermore, some of the studies summarized in these reviews use empirical strategies that do not address simultaneity bias or measurement error. Allcott, Lockwood, and Taubinsky (2018) estimate an aggregate elasticity of -1.40.

As discussed earlier, the welfare effects of SSB taxes also depend on whether they affect consumption of other untaxed goods that generate externalities or internalities. Various papers estimate demand systems that capture these substitution patterns between SSBs and other foods and beverages. Possibly due to the challenges in data quality and variation in identification strategies, there is very little agreement in this literature. For example, Duffey, Gordon-Larsen, Shikany, Guilkey, Jacobs, and Popkin (2010) find that pizza is a strong substitute for SSBs. On the other hand, Finkelstein, Zhen, Bilger, Nomemaker, Farooqu, and Todd (2013) find that only one food category is a statistically significant substitute for SSBs: canned soup. By contrast, Zhen, Finkelstein, Nomemaker, Kuns, and Todd (2014) find that canned soup is a complement to carbonated soft drinks but a substitute for sports drinks, energy drinks, and juice drinks. Smith, Lin, and Lee (2010) and Dharmasena and Capps (2012) find that SSBs and diet drinks are complements, while other studies disagree. These conflicting and sometimes counterintuitive results highlight the difficulty of estimating substitution patterns. Using the instrumental variable strategy described above, Allcott, Lockwood, and Taubinsky (2018) find that only one type of beverage is a statistically significant substitute or complement: diet soft drinks are a substitute for regular SSBs.

A separate but related parameter is the elasticity of SSB consumption with respect to an SSB tax. As illustrated in Figure 4, this is different than the elasticity with respect to the SSB price, and it depends on both the supply and demand elasticity. This parameter is of interest because it determines the public health effect of a tax. It also captures how interest groups’ advertising campaigns and public debates about sin taxes, and the resulting public awareness, could affect demand over and above the effect of a price increase, as highlighted by Rees-Jones and Rozema (2018). Unfortunately, the tax elasticity is also difficult to identify, and is likely to be highly context-dependent because both the political climate and the political process by which the tax is implemented should affect the extent of public debates and interest groups’ advertising. Fletcher, Frisvold, and Teft (2010) study how SSB consumption responds to changes in how SSBs are treated in state sales and excise taxes, but this variation is very limited: among states with a non-zero tax during their sample period, the average tax rate was no more than about five percent. Silver, Ng, Ryan-Ibarra, Taillie, Induni, Miles, Poti, and Popkin (2017), Bollinger and Sexton (2017), and others study responses to the Berkeley SSB tax. While this tax rate is higher than the taxes studied by Fletcher, Frisvold, and Teft (2010), having only one city limits the sample size and requires the strong assumption that no factors other than the tax change affected Berkeley SSB demand. As more cities implement soda taxes, the sample size will grow. Economically, the knowing the tax elasticity is not sufficient for calculating welfare effects of SSB taxes. It is important to measure how it decomposes into price versus non-price effects on demand.
Externalities

In the context of the simple framework presented above, the health effects of SSB consumption generate two main types of externalities: health cost externalities and other fiscal externalities. The key statistic is the average marginal externality caused by SSB consumption, weighted by both the elasticity of demand and the welfare weight of the person or entity whom the externality affects.\(^6\)

Health cost externalities result because most Americans have health insurance, typically through their employers, Medicare, or Medicaid, and thus most of the health costs caused by SSB consumption are paid for by others.\(^7\) Wang, Coxson, Shen, Goldman, and Bibbins-Domingo (2012) and Long, Gortmaker, Ward, Resch, Moodie, Sacks, Swinburn, Carter, and Wang (2015) both estimate the health system costs of SSB consumption, reporting estimates of approximately one cent per ounce of SSB consumed. The U.S. Department of Health and Human Services (Yong, Bertko, and Kronick, 2011) estimates that on average, about 15 percent of health costs are borne by the individual, while 85 percent are covered by insurance. Cawley and Meyerhoefer (2012) estimate that 88 percent of the total medical costs of obesity are borne by third parties. This suggests that the average health cost externality from SSB consumption might be 0.8 to 0.9 cents per ounce.

The results of Bhattacharya and Bundorf (2009), however, suggest that obese people in jobs with employer-sponsored health insurance face the full health costs of obesity through lower wages. The wage gap between obese and non-obese people exists only in jobs that provide health insurance and is robust to a wide array of controls. As discussed above, however, physiological and statistical evidence suggests that SSB consumption causes diabetes and cardiovascular disease independently of effects through body weight. It may be more difficult for employers to discriminate against people who weigh the same but are more likely to suffer from diabetes or cardiovascular disease, because these diseases are not observable. Furthermore, the Bhattacharya and Bundorf (2009) results do not apply to people with government-sponsored health insurance (Medicare and Medicaid).

In addition to health cost externalities, SSB consumption imposes other fiscal externalities, i.e. positive or negative effects on the government’s budget. For example, obesity appears reduce people’s life spans, reducing the amount of social security benefits that obese people will claim (Bhattacharya and Sood, 2011; Fontaine, Redden, Wang, Westfall, and Allison, 2003).

One important but unknown statistic is the covariance between people’s elasticity of demand and the marginal health damages of SSB consumption. For example, low-income people are thought to be more price elastic, and their health cost externalities may be higher (if their health costs are not offset by wage reductions because they are on Medicaid) or lower (if they are more likely to be uninsured). Dubois, Griffith, and O’Connell (2017) argue that SSB consumption by young people might generate larger health harms, and they show that young people are more price elastic. SSB consumption by people who are pre-diabetic—that is, just below the threshold for receiving diabetes treatment—likely involves large health cost externalities.

---

\(^6\) Because obesity is one of the main channels of health harms from SSBs, the article “Who Pays for Obesity” (Bhattacharya and Sood, 2011) in this Journal is also highly relevant.

\(^7\) This is different from moral hazard, which is the elasticity of health care utilization with respect to insurance coverage. The term “health cost externalities” reflects that some costs of SSB consumption are borne by others, regardless of whether there is any elasticity of SSB consumption with respect to health insurance coverage.
There are three main challenges to measuring consumer bias in a way that can be integrated in the behavioral public economics framework sketched in Figures 3 and 4. First, there is a mechanical tension in evaluating paternalistic policies, which are predicated on the idea that consumers do not act in their own best interest, using revealed preference techniques, which are predicated on the idea that consumers do act in their own best interest. Following Bernheim and Rangel (2009), behavioral welfare analyses must somehow establish a “welfare-relevant domain”—that is, a subset of consumer choices that are assumed to be unbiased—versus another subset of “suspect” choices that may be affected by bias. This exercise is controversial and seemingly not solvable with choice data alone. Second, the researcher must identify the causal impact of bias, which involves the same type of identification challenges as those present in estimating price elasticities and other parameters. Third, this causal impact must be translated into dollar units, as highlighted by the fact that the internality and/or externality $b$ is a vertical distance separating the demand curves on Figures 3 and 4. While much of the behavioral economics literature has focused on simply testing whether some behavioral bias exists, behavioral welfare analysis requires that the effects of a bias be quantified specifically in dollar terms, just as environmental policy analysis requires researchers to quantify the social cost of carbon or other pollutants in units of dollars.

Our read of the SSB tax debate is that there are two reasons why consumers might not act in their own best interest when buying SSBs. The first is imperfect information: consumers might not know how bad SSBs are for their health. The second is self-control: time-inconsistent consumers might underweight the future health costs of SSB consumption relative to how they would like to weight those costs in the future. Different empirical strategies can be used to identify each source of behavioral bias. For imperfect information, researchers can estimate the effects of information provision, as in Allcott and Taubinsky (2015) and others. For self-control, researchers can compare choices made for consumption now versus in the future, as in Read and van Leeuwen (1998), Augenblick, Niederle, and Sprenger (2015), and others. For example, Sadoff, Samek, and Sprenger (2015) take advance orders for grocery delivery and allow people to re-optimize their choices at the time that the groceries are delivered, finding that people tend to re-optimize toward less-healthy options and that one-third of people would like to restrict their own future ability to re-optimize. However, standard “preference reversal” experiments cannot provide a quantification of the effects limited self-control in dollar terms.

Alternatively, a “counterfactual normative consumer” approach can be used to simultaneously measure multiple biases, and quantify their effects in dollar terms. As an example of this approach, Bronnenberg, Dubé, Gentzkow, and Shapiro (2013), show that sophisticated shoppers—in their application, doctors and pharmacists—are more likely to buy generic instead of branded drugs, and they conduct welfare analysis assuming that only sophisticates’ choices are welfare-relevant. Bartels (1996), Handel and Kolstad (2015),
Johnson and Rehavi (2016), and Levitt and Syverson (2008) similarly compare informed to uninformed agents to identify the effects of imperfect information. Allcott, Lockwood, and Taubinsky (2018) use this approach to measure both imperfect information and self-control. Specifically, they survey Nielsen Homescan panelists to measure nutrition knowledge and perceived overconsumption of SSBs, and they relate these survey bias proxies to SSB consumption, controlling for demographic variables and survey-based controls for SSB tastes and health preferences, and correcting for measurement error in the survey measures of bias. The key weakness of this approach is that it requires the assumption that preferences are conditionally independent of measures of consumer bias.

Under this unconfoundedness assumption, Allcott, Lockwood, and Taubinsky (2018) predict that the average American household would consume 38 to 48 percent fewer SSBs if they had the nutrition knowledge of dietitians and nutritionists and perfect self-control. Translated into dollar terms for use as part of the $b$ term in Figure 2, this means that the average marginal consumer overweights the marginal utility of consuming SSBs by an average marginal bias of 1.14 to 2.78 cents per ounce, or 34 to 84 percent of the average purchase price. This average marginal internality is about 30 percent larger at household incomes below $10,000 per year compared to at household incomes above $100,000 per year. Since low-income consumers appear to be more biased, the above theoretical framework implies that the bias correction benefits from soda taxes are progressive, even though the poor pay more in soda taxes.

Pass-through

Bollinger and Sexton (2017), Cawley and Frisvold (2017), Falbe, Rojas, Grummon, and Madsen (2015), Rojas and Wang (2017), and Silver, Ng, Ryan-Ibarra, Taillie, Induni, Miles, Poti, and Popkin (2017) estimate the pass-through of the Berkeley SSB tax. Between them, these papers use two complementary data sources. The first is scanner data, which in practice is limited to a small number of large chain retailers. The second is audit studies of prices, which allow coverage of smaller retailers such as independent convenience stores and gas stations, which may have very different pass-through rates, but are expensive to gather. Although the point estimates differ substantially across the papers and across types of retailers and SSBs, all four of these papers find less than full pass-through. This implies that at least some of the incidence of the tax is on suppliers.

Consumers’ ability to substitute to stores outside of Berkeley makes demand more elastic, which Figure 4 shows will reduce pass-through rates. Bollinger and Sexton (2017) find that approximately half of SSB purchase reductions appear to be substituted to retailers just outside of Berkeley. This “leakage” reduces the possible benefits from Berkeley’s tax and may generate deadweight loss in the form of higher travel costs incurred by consumers.

Bollinger and Sexton (2017) also document how retailers’ overall pricing strategies interact in important ways with a local tax on a small subset of products. First, as documented by DellaVigna and Gentzkow (2017) and Hitsch, Hortacsu, and Lin (2017), large retail chains often set uniform prices across many stores in many cities. This limits the extent to which a local cost increase from a local tax is passed through to retail prices in that area. Second, retailers often use category pricing, where, for example, all two-liter bottles of regular and diet soda might have the same price. If retailers maintain equal prices for regular and diet soda and diet soda consumption involves lower internalities or externalities because it does not contain sugar, this reduces the tax’s welfare gains.
Guiding Principles for Policy Makers

As the above review of the empirical evidence makes clear, there is still substantial uncertainty about the value of some relevant parameters. Nevertheless, economic theory does provide some guiding principles which may prove useful to policy makers interested in designing soda taxes or other corrective policies. We emphasize seven.

1. Tax grams of sugar, not ounces of soda.

Most SSB taxes to date have been structured as a per-ounce-of-drink tax on all beverages with added sugar (or, in some cases, other added sweeteners). That means drinks with high and low amounts of added sugar are both taxed the same way. From the perspective of the theoretical rationale for soda taxes, this structure makes little sense. Adverse health consequences are driven by sugar consumption, and therefore drinks containing more sugar generate greater externality and internality harms. Scaling the tax with sugar content, rather than drink volume, encourages consumers to switch to lower-sugar drinks, and encourages producers to reduce the sugar content itself.\(^\text{12}\)

2. Focus on counteracting externalities and internalities, not on minimizing SSB consumption.

Many papers in the public health literature focus on how much a tax would reduce the frequency of harmful health conditions reviewed earlier. Such analyses are incomplete measures of the social welfare effects of SSB taxes, because they do not account for the countervailing consumption benefits from SSBs. Put differently, the policy that maximizes health (i.e. minimizes diseases caused by SSBs) would be an infinite tax (i.e., an outright ban) on soda, whereas the economic logic from Figures 3 and 4 shows that the costs of soda consumption must be weighed against their consumption benefits. Taxation is optimal only to the extent that consumers underweight those costs due to externalities or internalities. We should maximize social welfare, not narrowly maximize health or minimize consumption of unhealthy goods.

3. Design soda taxes to reduce consumption among people generating the largest externalities and internalities.

A key message from Figure 3 is that the welfare effect of a soda tax depends on the size of externality and internality costs among consumers most responsive to the tax. Yet the responsiveness of different consumers to a tax often depends on features of implementation—for example, whether it is masked or heavily advertised. Therefore, policy makers should seek out tax design features which elicit large responses from consumers with the largest externality and internality costs. For example, if externalities and internalities are largest among children—for example, due to limited self control, or because their decisions lead to lifelong consumption patterns—then a higher tax (or an outright ban) may be justified in settings catering primarily to children, such as schools.

\(^\text{12}\)See Francis, Marron, and Reuben (2016) and Zhen, Brissette, and Ruff (2014) for estimates of the gains from SSB taxes that scale in sugar or calorie content. The UK’s tax on sugary drinks, which features a lower per-volume tax rate on drinks below a specific sugar density, has already led some producers to lower the sugar content in some beverages.
4. Make soda taxes consistent across geographic boundaries.

Many recent SSB tax policies have been implemented in specific (and sometimes quite small) geographic areas. While such local policies can provide a fertile testing ground for policy proposals, they raise the possibility that consumers may travel across borders to avoid the tax. Such cross-border shopping imposes costs on consumers, while underlining the externality and internality reduction benefits of the tax.

This relates to the broader principle about taking cross-price effects into account. SSBs in bordering geographic areas are a sugary substitute good for SSBs in one’s own geographic area. Thus, just as an SSB tax is less effective when consumers substitute to other sugary foods that generate high internalities and externalities, it is also less effective when it is easy for consumers to substitute to SSBs in bordering geographic areas. This highlights the importance of unifying soda tax policies across geographies.

5. Use caution when pre-allocating soda tax revenues.

Some recent SSB tax policies have explicitly preallocated anticipated revenues for specific uses. For example, a portion of the funds generated by the beverage tax in Philadelphia are preallocated early childhood education programs. Although such precommitments may help bolster political support for soda taxes, we urge caution in relying on them heavily, for a simple reason: if the SSB tax is effective at reducing consumption, the tax base will shrink, and the projected funds may not materialize. In addition to risking a backlash among constituencies who were promised soda tax funds, precommitments can create a perverse incentive for policy makers to design SSB taxes to raise revenue (e.g., by applying them to less harmful beverages) while undermining the corrective objectives that originally motivated the tax.

6. When judging regressivity, consider internality benefits, not just who pays the taxes.

As we have emphasized, an SSB tax may not make poor consumers worse off—even if they pay more in taxes than other consumers—because heavy SSB consumers also benefit from reduced internalities. As shown in Figure 3, although poorer consumers incur more costs due to paying more in taxes (area A), that may be offset (partially, or more-than-fully) by the gains they incur from having the tax reduces the harms that they impose on themselves from overconsumption (area B). Therefore, when making judgments about whether SSB taxes are regressive, policymakers should look both at the financial burden of the taxes themselves, and at the benefits conferred by reduced internalities, across the income distribution. This tradeoff is intimately related the elasticity of demand for soda—if consumers are highly elastic, and substantially reduce consumption when a tax is imposed, then corrective benefits are large relative to the financial burden, making the tax less regressive.

7. Taxing soda is probably a good idea

Our read of the evidence summarized above is that there is a strong case to be made that SSBs impose costs on the health system and on consumers themselves, that people do not internalize into their consumption decisions. Allcott, Lockwood, and Taubinsky (2018) formally integrate the theoretical and empirical considerations discussed in this overview, finding that the socially optimal SSB tax lies in the range of 1.5 to 2.8

---

13 One threat to such experimentation is a rising trend of state-level policies prohibiting local soda taxes; see Dana and Nadler (2017).
cents per ounce. While there is considerable uncertainty in these optimal tax estimates, they are not zero, and in fact somewhat higher than the levels set in most U.S. cities to date. The social welfare benefits from implementing the optimal tax nationwide are estimated to lie between $3.6 billion and $12.1 billion per year. These benefits would be considerably larger for a tax that followed our first principle above by taxing grams of sugar instead of ounces of beverage.


