REGRESSIVE SIN TAXES, WITH AN APPLICATION TO THE OPTIMAL SODA TAX[*]

Hunt Allcott
Benjamin B. Lockwood
Dmitry Taubinsky

A common objection to "sin taxes"—corrective taxes on goods that are thought to be overconsumed, such as cigarettes, alcohol, and sugary drinks—is that they often fall disproportionately on low-income consumers. This paper studies the interaction between corrective and redistributive motives in a general optimal taxation framework and delivers empirically implementable formulas for sufficient statistics for the optimal commodity tax. The optimal sin tax is increasing in the price elasticity of demand, increasing in the degree to which lower-income consumers are more biased or more elastic to the tax, decreasing in the extent to which consumption is concentrated among the poor, and decreasing in income effects, because income effects imply that commodity taxes create labor supply distortions. Contrary to common intuitions, stronger preferences for redistribution can increase the optimal sin tax, if lower-income consumers are more responsive to taxes or are more biased. As an application, we estimate the optimal nationwide tax on

sugar-sweetened beverages, using Nielsen Homescan data and a specially designed survey measuring nutrition knowledge and self-control. Holding federal income tax rates constant, our estimates imply an optimal federal sugar-sweetened beverage tax of 1 to 2.1 cents per ounce, although optimal city-level taxes could be as much as 60% lower due to cross-border shopping. *JEL* Codes: D04, D61, D62, H21, H23.

> "The only way to protect all of us, including the poor, from further harm is through a sugary drink tax."
>
> —Huehnergarth (2016)

> "A tax on soda and juice drinks would disproportionately increase taxes on low-income families in Philadelphia."
>
> —U.S. Senator Bernie Sanders (2016)

> "They've [big soda] made their money off the backs of poor people, but this money [soda tax revenue] will stay in poor neighborhoods."
>
> —Philadelphia Mayor Jim Kenney
> (quoted in Blumgart 2016)

I. Introduction

A large literature in behavioral economics suggests that biases, such as self-control problems, inattention, and incorrect beliefs, can lead to overconsumption of "sin goods," such as cigarettes, alcohol, unhealthy foods, and energy-inefficient durable goods. Consumption of these goods can also generate externalities in the form of health care costs or pollution. Consequently, "sin taxes" that discourage consumption of such goods could increase social welfare. This argument has led to widespread taxation of cigarettes and alcohol, as well as newer taxes on sugar-sweetened beverages in seven U.S. cities and 34 countries around the world (GFRP 2019).

What is the optimal level of a sin tax? The existing literature frequently invokes a corrective logic dating to Pigou (1920) and Diamond (1973): the optimal corrective tax equals the sum of the externality and the average mistake (or "internality") of marginal consumers.[1] This principle, however, assumes that consumers do

---

1. For Pigouvian taxation of internalities, see O'Donoghue and Rabin (2006), Gruber and Kőszegi (2001), Allcott, Mullainathan, and Taubinsky (2014), Allcott and Taubinsky (2015), Heutel (2015), and Mullainathan, Schwartzstein, and Congdon (2012).

not vary in their marginal utility of money, and thus that policy makers care equally about the poor versus the rich. This assumption is starkly out of sync with public debates about sin taxes. As highlighted by the quote from Senator Bernie Sanders, a common objection to sin taxes is that they are regressive.[2]

In response to such objections, others argue that the harms caused by overconsumption from behavioral biases are themselves regressive, so a corrective tax might confer greater benefits on the poor than on the rich. For example, smoking and sugary drink consumption cause lung cancer, diabetes, and other health problems that disproportionately affect the poor. Furthermore, as emphasized in the quote from Philadelphia Mayor Jim Kenney, regressivity can be reduced by "recycling" sin tax revenues to programs that benefit the poor.

This article presents a general theoretical model delivering the first explicit formulas for an optimal commodity tax that accounts for the three central considerations in such policy debates: correction of externalities and/or consumer bias, regressivity, and revenue recycling. We use the theoretical formulas to estimate the optimal nationwide tax on sugar-sweetened beverages using new data and empirical techniques to estimate elasticities and biases.

Our theoretical model in Section II builds on Saez's (2002a) extension of Atkinson and Stiglitz (1976) by considering an economy of consumers with heterogeneous earning abilities and tastes who choose labor supply and a consumption bundle that exhausts their after-tax income. The policy maker chooses a set of linear commodity taxes and a nonlinear income tax, which can be used to provide transfers to poor consumers, raise money for commodity subsidies, or distribute commodity tax revenue (in a progressive way, if desired). Unlike those models, we allow for a corrective motive of taxation, driven by externalities or internalities from consumer mistakes.

Our theoretical results fill two gaps in the literature. First, even in the absence of internalities or externalities, there was no known general formula for optimal commodity taxes in the presence of nonlinear income taxation and preference heterogeneity.[3]

---

2. The poor disproportionately consume cigarettes and sugary drinks, while the rich disproportionately take up energy efficiency subsidies (Gruber and Kőszegi 2004; Goldin and Homonoff 2013; Allcott, Knittel, and Taubinsky 2015; Davis and Borenstein 2016; Davis and Knittel 2016).

3. Saez (2002a) answered the qualitative question of when a "small" commodity tax can increase welfare in the presence of preference heterogeneity, but left

As a result, tax economists' policy recommendations have relied on the canonical result of Atkinson and Stiglitz (1976) that commodity tax rates should be uniform. This extreme result, however, requires a homogeneous preferences assumption that is likely unrealistic in many settings—and that we show is strongly rejected in the case of sugary drinks. Second, there has been no general account of how redistributive motives and behavioral biases jointly shape optimal commodity taxes. Although this is not the first article to study internality-correcting commodity taxes, it is the first to embed internalities in the dominant optimal taxation framework of public economics, allowing for redistributive motives and a nonlinear income tax, as well as commodity taxes. Most previous papers studying internality taxes abstract from redistributive motives. Those that do not (such as Gruber and Kőszegi 2004; Bernheim and Rangel 2004; Farhi and Gabaix 2015) focus on specialized models of bias and/or consider simplified tax environments that do not permit redistribution through nonlinear income taxation, means-tested transfer programs, or progressive revenue recycling.[4]

Our optimal commodity tax formula decomposes into two terms that address these gaps. The first term represents the "redistributive motive": the desire to use the commodity tax to transfer money from the rich to the poor by subsidizing goods consumed by low earners. This motive depends on the extent of between-income preference heterogeneity, which is assumed away in the special case studied by Atkinson and Stiglitz (1976). We derive a novel sufficient statistic that quantifies this preference heterogeneity: the difference between the cross-sectional variation in consumption of the sin good across incomes and the (causal) income effect.[5]

---

to future work the task of deriving an expression for the optimal tax, writing, "It would of course be extremely useful to obtain optimal commodity tax formulas" in such a general framework. See also Diamond and Spinnewijn (2011) and Gauthier and Henriet (2018) for work on discrete type models.

4. Gruber and Kőszegi (2004) study the incidence of cigarette taxes on low- and high-income consumers with self-control problems, but they do not characterize the optimal tax implications. Both Farhi and Gabaix (2015) and Bernheim and Rangel (2004) assume that commodity taxes are the sole source of redistribution: the commodity tax revenue has to be distributed as a lump sum and cannot, for example, be spent on transfers to the poor or programs that benefit the poor.

5. Our results also generalize Jacobs and Boadway (2014), Jacobs and de Mooij (2015), and Kaplow (2012), who study optimal commodity taxes under strong homogeneity assumptions. Our work extends the "double dividend" literature

The second term in our optimal commodity tax formula represents the "corrective motive": the desire to reduce overconsumption arising from internalities and externalities by imposing taxes on harmful goods. To incorporate internalities, we follow the sufficient statistics approach to behavioral public finance (Mullainathan, Schwartzstein, and Congdon 2012; Farhi and Gabaix 2015; Chetty 2015) by adopting a money-metric definition of bias, which can transparently accommodate many specific behavioral biases and lends itself to direct empirical quantification.

The price elasticity of demand determines the relative importance of the corrective versus redistributive motives. High demand elasticity implies a large change in sin good consumption for a given degree of redistribution, and thus that the effects of the tax are primarily corrective rather than redistributive. Conversely, low demand elasticity implies that the effects of the tax are mostly redistributive rather than corrective.

In contrast to the conventional wisdom that inequality aversion unambiguously reduces optimal taxes on sin goods heavily consumed by the poor, we show that it can either decrease or increase the optimal sin tax. Although inequality aversion does magnify the redistributive motive, pushing toward a lower sin tax, it also amplifies the corrective motive when poor consumers are relatively more biased or more elastic, which pushes toward a higher sin tax.

A key contribution of our theoretical work is that it delivers formulas for optimal commodity taxes as a function of sufficient statistics that can be estimated in a wide variety of empirical applications. In Section III, we apply the theory by estimating the necessary statistics for the optimal nationwide tax on sugar-sweetened beverages (SSBs). We use Nielsen Homescan, a 60,000-household, nationally representative panel data set of grocery purchases, and Nielsen Retail Measurement Services (RMS), a panel data set from 37,000 stores covering about 40 percent of all U.S. grocery purchases. A plot of the Homescan data in Figure I illustrates how SSB taxes could be regressive: households with annual income below $10,000 purchase about

---

analyzing the interaction between carbon taxes and income tax distortions (e.g., Goulder 1995, 2013; Bovenberg and Goulder 1996; Goulder and Williams 2003); this literature assumes linear income taxes and does not consider redistributive motives.
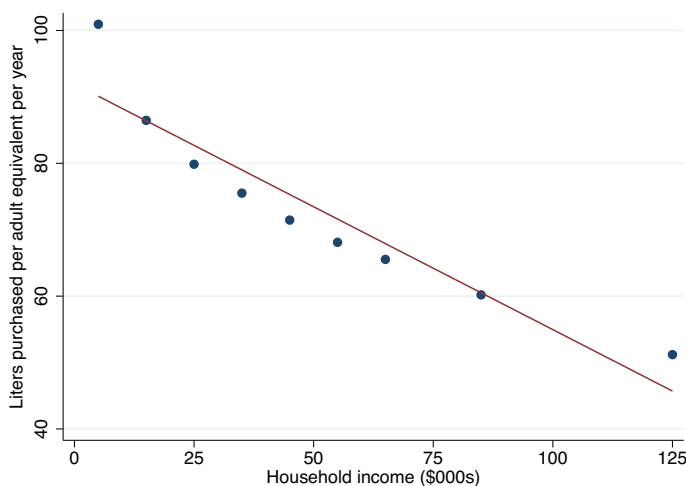
FIGURE I

Homescan Sugar-Sweetened Beverage Purchases by Income

This figure presents the average purchases of sugar-sweetened beverages by household income, using Nielsen Homescan data for 2006–2016. Purchases are measured in liters per "adult equivalent," where household members other than the household heads are rescaled into adult equivalents using the recommended average daily consumption for their age and gender group. Observations are weighted for national representativeness.

101 liters of SSBs per adult each year, whereas households with income above $100,000 purchase only half that amount.

To identify the price elasticity of demand, we develop an instrument that exploits retail chains' idiosyncratic pricing decisions for the UPCs that a household usually buys at the retailers where the household usually buys them. For example, if Safeway puts Gatorade on sale, people who often buy Gatorade at Safeway face a lower price for their SSBs compared with people who don't buy Gatorade or don't shop at Safeway. We ensure that the instrument is not contaminated by local or national demand shocks by using the retailer's average price charged outside of a household's county and by using only deviations from each UPC's national average price. Because retailers regularly vary prices independently of each other while keeping prices fairly rigid across their stores (DellaVigna and Gentzkow, forthcoming), the instrument delivers higher power while relaxing stronger assumptions required

for traditional instruments such as those proposed by Hausman (1996) and Nevo (2001).

We estimate the income elasticity of demand using within-county income variation over time. We find a small positive income elasticity, which means that the downward-sloping consumption-income profile illustrated in Figure I is driven by strong between-income preference heterogeneity, not by causal income effects. This strong preference heterogeneity—but not the declining consumption-income profile per se—reduces the socially optimal SSB tax.

To quantify consumer bias, we designed a survey of 18,000 Homescan households measuring nutrition knowledge and self-control. We find that both bias proxies are strongly associated with SSB purchases. For example, households in the lowest decile of nutrition knowledge purchase more than twice as many SSBs as households in the highest decile. Furthermore, the distribution of these proxies suggests that bias may be regressive. People with household incomes below $10,000 have 0.82 standard deviations lower nutrition knowledge and report that they have 0.40 standard deviations lower self-control than do people with household incomes above $100,000.

We formally quantify consumer bias using what we call the "counterfactual normative consumer" strategy, which builds on Bronnenberg et al. (2015), Handel and Kolstad (2015), and other work. We estimate the relationship between SSB consumption and bias proxies after conditioning on a rich set of preference measures and demographics and correcting for measurement error. We then predict "normative" consumption—that is, consumption if people had the nutrition knowledge of dietitians and nutritionists as well as perfect self-control. To interpret this prediction, we assume that any unobserved preferences are conditionally independent of bias. This unconfoundedness assumption is the key weakness of our approach.

We predict that American households would consume 31%–37% fewer SSBs if they had the nutrition knowledge of dietitians and nutritionists and perfect self-control. This estimated over-consumption is higher among the poor, accounting for 37%–48% of consumption for households with incomes below $10,000, compared with 27%–32% of consumption for households with incomes above $100,000. This regressive bias implies a higher optimal soda tax.

In Section IV, we implement our optimal tax formulas using our empirical results. In our baseline specification, the optimal federal-level SSB tax is 1.42 cents per ounce, or 39% of the quantity-weighted average price of SSBs recorded in Homescan. For a broad range of specifications, the optimal federal SSB tax lies in the range of 1 to 2.1 cents per ounce, or 28%–59% of the quantity-weighted average price. Our preferred estimates imply that the welfare benefits from implementing the optimal tax are between $2.4 billion and $6.8 billion a year. Although SSB consumption is highly concentrated among low earners, the overall welfare effects are distributed much more evenly across incomes, since our estimates imply that the internality corrections are also greatest at low incomes. The welfare gains are about $100 million a year higher than what would be realized by imposing a 1 cent per ounce federal tax—currently the modal policy among U.S. cities that have implemented SSB taxes. After adjusting for the estimated cross-border shopping induced by recent city-level taxes, however, we estimate that the optimal city-level tax could be as low as 0.53 cents per ounce, which is lower than the current modal policy. Finally, we emphasize the importance of accounting for behavioral biases when designing policy: a tax designed without accounting for behavioral biases forgoes nearly $1 billion a year in potential welfare gains.

In addition to contributing to optimal tax theory and behavioral public economics, our work connects to a large and growing empirical literature on SSB taxes. One set of papers estimates the price elasticity of SSB demand and/or the effect of SSB taxes on consumption.[6] Our work contributes transparent estimates in a large nationwide sample, whereas most previous papers require more restrictive identifying assumptions or deliver less precise estimates, for example, because they exploit only one specific SSB tax change. A second set of papers additionally estimates how SSB taxes would affect consumer surplus, including Dubois, Griffith, and O'Connell (2017), Harding and Lovenheim (2015), Wang (2015), and Zhen et al. (2014). These papers do not attempt to quantify consumer bias, making it difficult to use the estimates

---

6. This includes Bollinger and Sexton (2019), Duffey et al. (2010), Finkelstein et al. (2013), Fletcher, Frisvold, and Teft (2010), Rojas and Wang (2017), Silver et al. (2017), Smith, Lin, and Lee (2010), Tiffin, Kehlbacher, and Salois (2015), Zhen et al. (2011), and others; see Andreyeva, Long, and Brownell (2010), Powell et al. (2013), and Thow, Downs, and Jan (2014) for reviews.

to evaluate a policy motivated by consumer bias. The most important difference between our article and the existing soda tax literature is that to our knowledge, ours is the only one that attempts to answer the following basic question: what is the optimal soda tax? More broadly, this article provides a theoretical and empirical framework for calculating optimal commodity taxes that can be applied in a wide variety of contexts.

II. Deriving the Optimal Sin Tax

II.A. Model

We begin with a conventional static optimal taxation setting: consumers have multidimensional heterogeneous types $\theta \in \Theta \subset \mathbb{R}_+^n$, distributed with measure $\mu(\theta)$. They supply labor to generate pretax income z, which is subject to a nonlinear tax $T(z)$. Net income is spent on two goods: a numeraire consumption good c and a "sin good" s, with pretax price p, which is subject to a linear commodity tax t. Therefore the consumer's budget constraint is $c + (p + t)s \leqslant z - T(z)$.

Each consumer chooses a bundle (c, s, z), subject to her budget constraint, to maximize "decision utility" $U(c, s, z; \theta)$. U is assumed to be increasing and weakly concave in its first two arguments and decreasing and strictly concave in the third. Decision utility may differ from "normative utility" $V(c, s, z; \theta)$, which the consumers would choose to maximize if they were fully informed and free from behavioral biases. Sin good consumption also generates a fiscal cost to the government of e per unit of s consumed. Pecuniary fiscal externalities are a natural case for sin goods such as sugar-sweetened beverages and cigarettes that raise the health care costs for public programs like Medicare.

The policy maker selects taxes $T(\cdot)$ and t to maximize normative utility, aggregated across all consumers using type-specific Pareto weights $\alpha(\theta)$,

$$(1) \qquad \max_{T, t} \left[\int_\Theta \alpha(\theta)[V(c(\theta), s(\theta), z(\theta); \theta)]d\mu(\theta) \right],$$

subject to a government budget constraint, which includes the externality costs of sin good consumption,

$$(2) \qquad \int_\Theta (ts(\theta) + T(z(\theta) - es(\theta)))d\mu(\theta) \geqslant R$$

and to consumer optimization

$$\{c(\theta), s(\theta), z(\theta)\} = \arg\max_{\{c,s,z\}} U(c, s, z; \theta)$$

(3)   s.t. $c + (p + t)s \leqslant z - T(z)$ for all θ.

The difference between U and V can capture a variety of different psychological biases. For example, consumers may have incorrect beliefs about certain attributes of s, such as calorie content, future health costs, or energy efficiency (Attari et al. 2010; Bollinger, Leslie, and Sorensen 2011; Allcott 2013). Alternatively, consumers may have limited attention or salience bias with respect to certain attributes of s (Allcott and Taubinsky 2015). Finally, present focus may lead consumers to underweight the future health costs of some goods (e.g., potato chips or cigarettes) as in Gruber and Kőszegi (2004) and O'Donoghue and Rabin (2006). Our framework allows us to treat present focus as a bias. However, our framework also allows us to study other welfare criteria that may be applied to the model—for example, the policy might place some normative weight both on the "future-oriented self" and on the "in-the-moment self."

A key goal of our theoretical analysis is to derive optimal tax formulas that can accommodate a variety of possible consumer biases while remaining empirically implementable. We do this by constructing a price metric for consumer bias.

Formally, let $s(p + t, y, z, \theta)$ be the sin good consumption chosen at total price $p + t$ by a type θ consumer who earns z and has disposable income y.[7] Analogously, define $s^V(p + t, y, z, \theta)$ to be the amount of s that would be chosen if the consumer were maximizing V instead. We define the bias, denoted $\gamma(p + t, y, z, \theta)$, as the value for which $s(p + t, y, z, \theta) = s^V(p + t - \gamma, y - s\gamma, z, \theta)$. In words, γ is equal to the compensated price reduction that produces the same change in demand as the bias does. In terms of primitives, $\gamma = \frac{U_s'}{U_c'} - \frac{V_s'}{V_c'}$.[8] (Throughout, we use the notation f_x'

---

7. Recall that because our model allows utility not to be weakly separable in leisure and consumption of s, s depends not only on disposable income but also on earned income z.

8. The first-order condition for consumer choice is $\frac{U_s'(c,s,z)}{U_c'(c,s,z)} = p + t$, with $s \cdot (p + t) + c = z - T(z)$. By definition, $\frac{V_s'(c,s,z)}{V_c'(c,s,z)} = p + t - \gamma$ with $s \cdot (p + t - \gamma) + c = z - T(z) - s\gamma$, from which the statement follows.

to denote the derivative of $f(x, y)$ with respect to x, and f''_{xy} for the cross-partial derivative with respect to x and y, etc. When no ambiguity arises, we sometimes suppress some arguments and write, for example, $s(\theta)$ for concision.) If $\gamma(\theta) > 0$, this means that type θ consumers overconsume s relative to their normative preferences, whereas $\gamma(\theta) < 0$ means that type θ consumers underconsume. Throughout the article, we assume that the sole source of disagreement between the consumer and policy maker is about the merits of s; we do not focus on labor supply misoptimization.

The statistic can be quantified by comparing consumers' choices in "biased" and "debiased" states, as we do in our empirical application. Other examples that informally employ this definition of bias include Chetty, Looney, and Kroft (2009) and Taubinsky and Rees-Jones (forthcoming), who quantify the (average) value of tax salience as the change in up-front prices that would alter demand as much as a debiasing intervention that displays tax-inclusive prices. Similarly, Allcott and Taubinsky (2015) estimate γ by measuring consumers' demand responses to an experimental intervention that targets informational and attentional biases.

To represent the policy maker's inequality aversion concisely, we employ the notion, common in the optimal taxation literature, of "social marginal welfare weights"—the social value (from the policy maker's perspective) of a marginal unit of consumption for a particular consumer, measured in terms of public funds. We define

$$(4) \qquad g(\theta) := \frac{\alpha(\theta) V'_c}{\lambda},$$

where V'_c represents the derivative of V with respect to its first argument, and λ is the marginal value of public funds (i.e., the multiplier on the government budget constraint at the optimum).[9] These weights are endogenous to the tax system but are useful for characterizing the necessary conditions that must hold at the optimum. We use $\bar{g} = \int_{\Theta} g(\theta) d\mu(\theta)$ to denote the average marginal

9. This definition implies that g represents the social value of a unit of marginal composite consumption c, rather than sin good consumption. When agents make rational decisions about consumption of s, this distinction is immaterial because of the envelope theorem.

social welfare weight. If there are no income effects on consumption and labor supply, then $\bar{g} = 1$ by construction.[10]

II.B. Our Approach

We derive an expression for the optimal sin tax using variational calculus arguments. This generates a first-order (necessary) condition for the optimal sin tax in terms of empirically estimable sufficient statistics and social marginal welfare weights. These statistics are themselves endogenous to the tax system, so this expression should not be understood as a closed-form expression for the optimal tax. However, to the extent that these statistics are stable around modest variations in tax policy, we can approximate the optimal tax by evaluating the statistics at the current tax policy. In Online Appendix M, we calibrate two different structural models that account for the endogeneity, and we show that the resulting optimal taxes are very close to those computed using estimates of the sufficient statistics at the current tax system.

Before formally defining the elasticity concepts and presenting the optimal tax formula, we briefly summarize the core economic forces that correspond to our elasticity concepts and feature in the formula. Intuitively, any variation in the sin tax has three main effects. First, a higher sin tax has a direct ("mechanical") effect on government revenue and on consumers' post-tax incomes. The social welfare consequences of this effect depend on the marginal value of public funds, and on whether the increased tax burden is shouldered more by those with higher or lower marginal utility of money.

Second, an increase in the sin tax leads to substitution away from the sin good, which reduces the revenue collected from taxing the sin good. In the absence of externalities or internalities, the envelope theorem implies the loss in sin tax revenues is the only consequential effect. In the presence of externalities and/or internalities, the behavior change is beneficial because it reduces externalities from consumption and because consumers now

---

10. Because the Pareto weights $\alpha(\theta)$ are exogenous, and because U and V produce identical behavior (and identical choice-based measures of bias) up to monotonic transformations, the social marginal welfare weights reflect a policymaker's or society's normative preference for reducing wealth inequality—they cannot be inferred by observing behavior. As in the rest of the optimal taxation literature, our formulas for optimal taxes will thus depend both on observable behavior (and people's quantifiable mistakes) and on the policy maker's (or society's) inequality aversion, as encoded by these weights.

consume less of a good that they have been overconsuming. The internality correction benefits are highest when the consumers with the highest biases are also most elastic to the sin tax. Moreover, when the low-income consumers are the most responsive and/or the most biased, the internality benefits of behavior change have the additional virtue of being "progressive."

Third, the sin tax could affect consumers' labor supply decisions. Imagine that the sin good is a normal good, so that if a consumer chooses to earn more, she consumes more of the sin good. In this case, increasing the sin tax indirectly increases the marginal tax burden from choosing higher earnings. This disincentive for higher labor supply would then lead to lower income tax revenue. The converse holds for inferior goods. Consequently, when the consumption of a sin good decreases with income, it is crucial to determine whether this is because it is an inferior good, or because preferences for this good are negatively correlated with earnings ability.

This potential earnings response may substantially affect the optimal sin tax even if consumers spend only a small share of their budget on the sin good. Intuitively, what matters is the change in earnings due to the sin tax as a share of expenditures on the sin good—and that share may be substantial even if spending on the sin good is small. This channel of behavioral response is the foundation for the classic Atkinson and Stiglitz (1976) result.[11]

There are also higher-order effects in the general formula in the Online Appendix that are negligible under the simplifying assumptions we make in our main theoretical result. These include considerations such as the fact that changes in labor supply can also affect consumption of the sin good through the income effects channel and that at a given income level there may be a covariance between income effects and internalities.

II.C. Elasticity Concepts, Sufficient Statistics, and Simplifying Assumptions

The optimal sin tax depends on three types of sufficient statistics: elasticities, money-metric measure of bias, and the

---

11. It is also possible that taxes on items that constitute a small share of consumers' budgets might not be salient when labor supply decisions are made. Nonsalience is an alternative form of behavioral misoptimization that may lead to departures from Atkinson and Stiglitz (1976), and though it is beyond the scope of this article, we consider it in our companion work, Allcott, Lockwood, and Taubinsky (2018).

"progressivity of bias correction." We describe these statistics here and collect them in Table I for reference. These statistics are understood to be endogenous to the tax regime (t, T), though we suppress those arguments for notational simplicity. We begin by defining the elasticities related to sin good consumption.

- $\zeta(\theta)$: the price elasticity of demand for s from type θ, formally equal to $-\left(\frac{ds(\theta)}{dt}\right)\frac{p+t}{s(\theta)}$. We assume that the price elasticity of demand equals the tax elasticity of demand.
- $\zeta^c(\theta)$: the compensated price elasticity of demand for s, equal to $-\left(\frac{ds(\theta)}{dt}\Big|_u\right)\frac{p+t}{s(\theta)}$.
- $\eta(\theta)$: the income effect on s expenditure, equal to $\zeta - \zeta^c$.
- $\xi(\theta)$: the causal income elasticity of demand for s, equal to $\frac{d}{dz}s\left(p+t, z-T(z), z; \theta\right) \cdot \frac{z}{s}$.[12]

In addition, we represent the labor supply response to tax reforms using the following parameters, which are defined formally in Online Appendix A. All behavioral responses are defined to include the full sequence of adjustments due to any nonlinearities in the income tax (Jacquet and Lehmann 2016).

- $\zeta_z^c(\theta)$: the compensated elasticity of taxable income with respect to the marginal income tax rate.
- $\eta_z(\theta)$: the income effect on labor supply.

We denote averages of these statistics using "bar" notation; for example, average consumption of s is denoted $\bar{s} := \int_\Theta s(\theta)d\mu(\theta)$, with aggregate elasticity of demand $\bar{\zeta} := -\left(\frac{d\bar{s}}{dt}\right)\frac{p+t}{\bar{s}}$. Similarly, we denote average consumption among consumers with a given income z as $\bar{s}(z)$, with income-conditional elasticities denoted by

---

12. A change in net earnings may come from a change in labor supply or a change in nonlabor income (e.g., due to a tax-level reduction). If sin good consumption and labor are weakly separable in the utility function, sin good consumption will respond identically to either type of change in earnings. In that case, $\xi(\theta)$ is equal to $\frac{\eta(\theta)}{p+t} \cdot \frac{z(\theta)}{s(\theta)}(1 - T'(z(\theta)))$. More generally, ξ quantifies the sin good response to an increase in earnings from labor, for example due to a local reduction in marginal income tax rates. Weak separability implies that the change of s with response to an income shock dz, $\frac{ds}{dz}$, will not depend on whether dz comes from a change in hours or nonlabor earnings, and so the estimated relationship $\frac{ds}{dz}$ will be insensitive to the inclusion of controls for hours worked. Using this test, we find support for weak separability in our empirical application, as discussed in note 20.

TABLE I

NAMES AND DEFINITIONS OF SELECTED STATISTICS

Statistic	Name	Definition
p	Price of sin good	
e	Fiscal externality cost from sin good consumption	
$s(\theta)$	Individual sin good consumption	
$z(\theta)$	Individual labor income	
$\gamma(\theta)$	Individual price metric bias in sin good consumption	See Section II
$\xi(\theta)$	Individual causal income elasticity of sin good demand	$\frac{d}{dz}s\left(p+t, z-T(z), z; \theta\right) \cdot \frac{z}{s}$
$\zeta_z^c(\theta)$	Individual compensated elasticity of taxable income	See Online Appendix A
$h(z)$	Labor income density	$\int_\Theta 1\{z(\theta)=z\}d\mu(\theta)$
$\bar{s}(z)$	Average sin good consumption at income z	$\mathbb{E}\left[s(\theta)\|z(\theta)=z\right]$
$\bar{s}'(z)$	Cross-sectional variation of sin good consumption with income	$\frac{d\bar{s}(z)}{dz}$
$s'_{inc}(z)$	Causal income effect on sin good consumption	$\mathbb{E}\left[\xi(\theta)\frac{s(\theta)}{z} \mid z(\theta)=z\right]$
$s'_{pref}(z)$	Between-income preference heterogeneity	$\bar{s}'(z) - s'_{inc}(z)$
$s_{pref}(z)$	Cumulative between-income preference heterogeneity	$\int_{x=z_{min}}^z s'_{pref}(x)dx$
$g(z)$	Social marginal welfare weight on consumers earning z	See equation (4)
$\hat{g}(z)$	Social marginal utility of income	See note 15
$\bar{\gamma}(z)$	Average marginal bias at income z	$\frac{\int_\Theta \gamma(\theta)\left(\frac{ds(\theta)}{dt}\big|_u\right)1\{z(\theta)=z\}d\mu(\theta)}{\int_\Theta \left(\frac{ds(\theta)}{dt}\big|_u\right)1\{z(\theta)=z\}d\mu(\theta)}$
$\bar{\zeta}^c(z)$	Average compensated elasticity of sin good demand at income z	$\mathbb{E}\left[\left(\frac{ds(\theta)}{dt}\big|_u\right)\|z(\theta)=z\right] \cdot \frac{p+t}{\bar{s}(z)}$
$\bar{s}$	Average sin good consumption	$\int_z \bar{s}(z)h(z)dz$
$\bar{g}$	Average social marginal welfare weight	$\int_z g(z)h(z)dz$
$\bar{\gamma}$	Average marginal bias	$\frac{\int_\Theta \gamma(\theta)\left(\frac{ds(\theta)}{dt}\big|_u\right)d\mu(\theta)}{\int_\Theta \left(\frac{ds(\theta)}{dt}\big|_u\right)d\mu(\theta)}$
σ	Progressivity of bias correction	$Cov\left[g(z), \frac{\bar{\gamma}(z)}{\bar{\gamma}}\frac{\bar{\zeta}^c(z)}{\bar{\zeta}^c}\frac{\bar{s}(z)}{\bar{s}}\right]$

$\bar{\zeta}(z) := -\left(\frac{d\bar{s}(z)}{dt}\right)\frac{p+t}{\bar{s}(z)}$. The income distribution is denoted $H(z) :=$ $\int_{\Theta} 1\left\{z(\theta) \leqslant z\right\} d\mu(\theta)$, with income density denoted $h(z)$.

It is necessary to distinguish between two sources of cross-sectional variation in $\bar{s}(z)$: income effects and (decision) preference heterogeneity. Let $\bar{s}'(z)$ denote the cross-sectional change in s with respect to income z at a particular point in the income distribution. This total derivative can be decomposed into two partial derivatives: the (causal) income effect, $s'_{inc}(z)$, and between-income preference heterogeneity $s'_{pref}(z)$. The causal income effect depends on the empirically estimable income elasticity of s: $s'_{inc}(z) = \mathbb{E}\left[\xi(\theta)\frac{s(\theta)}{z} \mid z(\theta) = z\right]$. Between-income preference heterogeneity is the residual: $s'_{pref}(z) = \bar{s}'(z) - s'_{inc}(z)$. The key sufficient statistic for preference heterogeneity, "cumulative between-income preference heterogeneity" is defined as:

$$(5) \qquad s_{pref}(z) := \int_{x=z_{min}}^{z} s'_{pref}(x)dx.$$

This term quantifies the amount of sin good consumption at income z, relative to the lowest income level z_{min}, that can be attributed to preference heterogeneity rather than income effects.[13]

To aggregate bias across consumers, we follow Allcott, Mullainathan, and Taubinsky (2014) and Allcott and Taubinsky (2015) in defining the average marginal bias

$$(6) \qquad \bar{\gamma} := \frac{\int_{\Theta} \gamma(\theta) \left(\frac{ds(\theta)}{dt}\Big|_{u}\right) d\mu(\theta)}{\int_{\Theta} \left(\frac{ds(\theta)}{dt}\Big|_{u}\right) d\mu(\theta)}.$$

---

13. One question that arises here is the relevant time horizon for income effects, since preferences may themselves be endogenous to income over long periods. For example, consumption patterns of children in poor households—including those causally driven by low income—may affect their preferences later in life, even if their income has increased. Conceptually, the relevant statistic for the optimal tax formula is the labor supply distortion generated by the commodity tax, which is proportional to the income effect on that commodity, measured over the same time horizon as labor supply decisions. Therefore, to the extent that preferences are endogenous to income patterns at longer horizons than labor supply decisions, these should be treated as preference heterogeneity rather than income effects. In practice, our empirical estimates of income effects will be estimated using annual income variation.

Intuitively, this aggregation represents the marginal bias weighted by consumers' marginal responses to a tax reform that raises t while reforming T to offset the average effect on wealth at each income z. In other words, if a tax perturbation causes a given change in total consumption of s, $\bar{\gamma}$ is the average amount by which consumers over- or underestimate the change in utility from that change in consumption. We analogously define $\bar{\gamma}(z)$ as the response-weighted bias conditional on income.[14]

Because our framework considers redistributive motives, unlike Allcott, Mullainathan, and Taubinsky (2014) and Allcott and Taubinsky (2015), we must also account for the progressivity of bias correction:

$$
(7) \qquad \sigma := Cov\left[g(z), \frac{\bar{\gamma}(z)}{\bar{\gamma}} \frac{\bar{\zeta}^c(z)}{\bar{\zeta}^c} \frac{\bar{s}(z)}{\bar{s}}\right].
$$

The term σ is the covariance of welfare weight with the product of consumption-weighted bias and (compensated) elasticity. If this term is positive, it indicates that bias reductions in response to a tax increase are concentrated among consumers with high welfare weights, that is, those with lower incomes.

We impose the following assumptions, common in the optimal commodity taxation literature to focus on the interesting features of sin taxes in a tractable context.

ASSUMPTION 1. Constant social marginal welfare weights conditional on income: $g(\theta) = g(\theta')$ if $z(\theta) = z(\theta')$.

This assumption is analogous to Assumption 1 in Saez (2002a). It holds immediately if types are homogeneous conditional on income. More generally, Saez (2002a) argues this is a reasonable normative requirement even under heterogeneity "if we want to model a government that does not want to discriminate between different consumption patterns." Therefore we sometimes write $g(z)$ to denote the welfare weight directly as a function of earnings.

---

14. Formally,

$$
\bar{\gamma}(z) := \frac{\int_{\Theta} \gamma(\theta) \left(\frac{ds(\theta)}{dt}\Big|_u\right) 1\{z(\theta) = z\} \, d\mu(\theta)}{\int_{\Theta} \left(\frac{ds(\theta)}{dt}\Big|_u\right) 1\{z(\theta) = z\} \, d\mu(\theta)}.
$$

ASSUMPTION 2. U and V are smooth functions that are strictly concave in c, s, and z, and μ is differentiable with full support.

ASSUMPTION 3. The optimal income tax function $T(\cdot)$ is twice differentiable, and each consumer's choice of income z admits a unique global optimum, with the second-order condition holding strictly at the optimum.

Assumptions 2 and 3 ensure that the income distribution does not exhibit any atoms and consumers' labor supply and consumption decisions respond smoothly to perturbations of the tax system (Jacquet and Lehmann 2016).

ASSUMPTION 4. One of the following conditions hold: (a) heterogeneity is unidimensional, so that consumers with a given ability all have the same preferences and behavioral responses, or (b) the sin good s accounts for a small share of all consumers' budgets (so that terms of order $\frac{(p+t)s}{z}$ are negligible) and demand for s is orthogonal to η_z and ζ_z conditional on income.

Assumptions 1–4 are the required conditions for our primary characterization of the optimal sin tax in Proposition 1. However, the expressions simplify further and become empirically more feasible to implement, if we also impose the following assumption.

ASSUMPTION 5. Assumption 4(b) holds, and income effects on labor supply are negligible.

The negligible labor supply income effects assumption is supported by Gruber and Saez (2002), who find small and insignificant income effects on labor supply, and by Saez, Slemrod, and Giertz (2012), who review the empirical literature on labor supply elasticities and argue that "in the absence of compelling evidence about significant income effects in the case of overall reported income, it seems reasonable to consider the case with no income effects."

Assumptions 4 and 5 are not necessary for our proof strategy, and the full optimal commodity tax without these assumptions is derived in Online Appendix C.A, Proposition 6. However, they simplify the optimal tax expressions, and they are realistic for sin goods such as sugary drinks that account for a relatively small share of expenditures. Therefore we impose the full set of Assumptions 1–5 in our empirical implementation in Sections III

and IV. In the empirical implementation we will also assume that the elasticities conditional on income are homogeneous, which is weaker than Assumption 4(a).

II.D. General Expression for the Optimal Sin Tax

To characterize the optimal commodity tax, it is helpful to define the social marginal utility of income, denoted $\hat{g}(z)$, which is defined (as in Farhi and Gabaix 2015) as the average welfare effect of marginally increasing the disposable incomes of consumers currently earning income z. The weights $\hat{g}(z)$ incorporate any fiscal externalities resulting from income effects, and also the social welfare effect from misspending this marginal income due to bias.[15] In Online Appendix B.C, we provide formulas expressed entirely in terms of the social marginal welfare weights $g(z)$. All proofs are contained in Online Appendix C.A.

PROPOSITION 1. Under Assumptions 1–4, the commodity tax t and the income tax T satisfy the following conditions at the optimum:

$$(8) \qquad t = \frac{\bar{\gamma}(\bar{g} + \sigma) + e - \frac{p}{\bar{s}\zeta^c} Cov\left[\hat{g}(z), s_{pref}(z)\right]}{1 + \frac{1}{\bar{s}\zeta^c} Cov\left[\hat{g}(z), s_{pref}(z)\right]}$$

$$T'(z^*) =$$
$$(9) \qquad \frac{\mathbb{E}\left[(g(z^*)\gamma(\theta) + e - t)\frac{\xi(\theta)s(\theta)\zeta_z^c(\theta)}{\zeta_z^c(z^*)z^*}|z(\theta) = z^*\right] + \frac{1}{\zeta_z^c zh(z^*)}\mathbb{E}\left[(1 - \hat{g}(z))|z \geqslant z^*\right]}{1 + \frac{1}{\zeta_z^c(z^*)zh(z^*)}\mathbb{E}\left[(1 - \hat{g}(z))|z \geqslant z^*\right]}.$$

If Assumption 5 also holds, the taxes are approximated by

$$(10) \qquad t \approx \frac{\bar{\gamma}(1 + \sigma) + e - \frac{p}{\bar{s}\zeta^c} Cov\left[g(z), s_{pref}(z)\right]}{1 + \frac{1}{\bar{s}\zeta^c} Cov\left[g(z), s_{pref}(z)\right]}$$

$$(11) \qquad T'(z^*) \approx \frac{\frac{1}{\zeta_z^c zh(z^*)}\mathbb{E}\left[(1 - g(z))|z \geqslant z^*\right]}{1 + \frac{1}{\zeta_z^c(z^*)zh(z^*)}\mathbb{E}\left[(1 - g(z))|z \geqslant z^*\right]}.$$

---

15. Formally, $\hat{g}(z^*) = g(z^*) + \mathbb{E}\left[\eta_z(\theta)\frac{T'}{1-T'}|z(\theta) = z^*\right] + \mathbb{E}\left[(t - g(\theta)\gamma(\theta) - e)\right.$ $\left(\frac{\eta(\theta)}{p+t} + \frac{\eta_z(\theta)}{1-T'}\frac{\xi(\theta)s(\theta)}{z(\theta)}\right)|z(\theta) = z^*\right]$. See Online Appendix C.A for further details. If Assumption 5 holds, then $\hat{g} = g$.

The expression in equation (10) is an approximation because it represents the optimal tax in the limit as the "small" terms in Assumptions 4(b) and 5 go to 0.

For the purpose of building intuition, note that equation (10) can be rearranged as

$$(12) \qquad t = \underbrace{\bar{\gamma}(1+\sigma)+e}_{\text{corrective motive}} + \frac{1}{\frac{d\bar{s}}{dt}} \underbrace{Cov\left[g(z), s_{pref}(z)\right]}_{\text{redistributive motive}}.$$

Equation (12) shows that the optimal tax is the combination of two main terms. The first term, $\bar{\gamma}(1+\sigma)+e$, corresponds to the corrective motive of the tax, and rises with both the negative externality e and internality (average marginal bias) $\bar{\gamma}$. The latter is scaled by $1+\sigma$, illustrating a key difference between externalities and internalities: the magnitude of correction for internalities—but not externalities—depends on whether the bias is bigger for the rich or the poor and on whether the rich or the poor have more elastic demand.[16]

Intuitively, the internality costs from a consumer's overconsumption fall back on that consumer, so they are scaled by the consumer's social marginal welfare weight. In contrast, the externality generated by any given consumer's consumption is borne by the whole population, and thus receives the same weight regardless of whose consumption generates it.

The second term, proportional to $Cov[g(z), s_{pref}(z)]$, corresponds to the redistributive motive of the tax. This term depends on the extent to which sin good consumption acts as a tag for ability. As a result, it depends on the covariation of welfare weights with only that component of consumption which is driven by preference heterogeneity. Intuitively, this term represents the power of the sin tax to accomplish redistribution which cannot already be achieved via income taxation. The importance of this term relative

16. This asymmetry is not an artifact of the assumption that externalities fall on the government's budget. We could alternatively allow for more flexible externalities that are nonlinear in $\bar{s}$ and heterogeneous across agents, reducing each consumer's net income by a type-specific amount $E(\bar{s})$. Then the consumer's budget constraint in equation (3) would instead be written $c + (p+t)\, s \leqslant z - T(z) - E(\bar{s})$. We can then define $e = E'_{\bar{s}}(\bar{s})$, and the optimal tax formulas in Proposition 1 remain the same. That is, under heterogeneous externalities one should set the externality correction equal to the average marginal externality; however, there is no covariance with individual demand elasticity or the level of s consumption, as is the case with internalities.

to the corrective component depends on how responsive consumption of $\bar{s}$ is to the tax: the more consumer behavior responds to the tax, the less important the redistributive motive relative to the corrective motive.

The income tax formula is a slight modification of Jacquet and Lehmann (2016), with the addition of the term $\mathbb{E}\left[(g(z^*)\gamma(\theta) + e - t)\frac{\xi(\theta)s(\theta)\zeta_z^c(\theta)}{\zeta_z^c(z^*)z^*}|z(\theta) = z^*\right]$, which accounts for the way a perturbation of the marginal income tax rate affects consumption of the sin good through the channel of changing consumers' earnings.

II.E. Interpreting the Formula: Special Cases and Additional Intuition

The optimal tax formula in equation (10) has a number of special cases that illustrate important insights about the forces governing the optimal tax. We highlight three special cases of particular interest.

1. Special Case 1: No Inequality Aversion. If social marginal welfare weights are constant (implying the policy maker has no desire to redistribute marginal resources from high- to low-income consumers), then

$$t = \bar{\gamma} + e. \tag{13}$$

This matches the core principle of Pigouvian taxation and the typical sin tax results in the behavioral economics literature (e.g., O'Donoghue and Rabin 2006; Mullainathan, Schwartzstein, and Congdon 2012; Allcott, Mullainathan, and Taubinsky 2014; Allcott and Taubinsky 2015). This special case obtains either if welfare weights are constant across incomes (e.g., if V is linear in c) or if there is no income inequality (so that all consumers have the same marginal utility of consumption). In both cases, the optimal commodity tax must exactly offset the average marginal bias plus the externality.

2. Special Case 2: No (Correlated) Preference Heterogeneity. When differences in consumption are due purely to differences in income, regressive consequences of a sin tax can be perfectly offset by modifications to the income tax. Equivalently, all feasible distribution can be carried out most efficiently through the income

tax itself, and the redistributive motive in equation (12) is 0.[17] Therefore in this case the optimal sin tax is

$$(14) \qquad t = \bar{\gamma}\,(1+\sigma)+e.$$

In contrast to Special Case 1, inequality aversion still plays a role in the size of the optimal commodity tax, as reflected by the σ term. For a given average marginal bias $\bar{\gamma}$, a relative increase in the biases or elasticities of low-income consumers increases σ and thus increases the social welfare benefit of bias correction. When bias and elasticity are constant across incomes, $\sigma > 0$, and the size of the optimal tax will exceed the optimal Pigouvian tax that prevails absent any inequality aversion.

   3. *Special Case 3: No Corrective Concerns.*   A third important special case is when both internalities and externalities are equal to 0, so that only distributional concerns are relevant. In this case,

$$(15) \qquad \frac{t}{p+t} = -\frac{Cov\left[g(z), s_{pref}(z)\right]}{\bar{s}\bar{\zeta}^c}.$$

Equation (15) bears a striking resemblance to Diamond's (1975) "many-person Ramsey tax rule." Diamond (1975) studies a Ramsey framework in which the income tax is constrained to be a lump-sum transfer, and he obtains almost the same expression as equation (15), except with $s_{pref}(z)$ replaced by $\bar{s}(z)$. Equation (15) generalizes that result, showing that in the presence of a nonlinear income tax, the optimal commodity tax still resembles the familiar inverse elasticity rule, with the modification that instead of taxing goods that high earners consume, the planner taxes goods that they prefer. Equation (15) also generalizes the Atkinson-Stiglitz theorem to the case of arbitrary preference heterogeneity. The Atkinson-Stiglitz theorem obtains as a special case of equation (15) when all variation in s consumption is driven by income effects, which then implies that $s_{pref} \equiv 0$ and thus $t = 0$.

---

17. Note that $s'_{pref}(z) = 0$ need not imply that preferences for s are homogeneous, only that they are not correlated with earnings ability. This special case corresponds to Assumption 3 in Saez (2002a), who shows that the Atkinson-Stiglitz result continues to hold under this assumption.

4. Further Insights and Intuition. In addition to the three special cases, there are a few other insights from Proposition 1 that relate to results elsewhere in the literature. First, equation (10) holds when $\bar{\gamma} = 0$, which may arise even with nonzero internalities if the consumers who overconsume (or underconsume) the good are inelastic to the tax (or subsidy). A striking implication of the result then is that when lower-income consumers prefer the good more, the optimal sin tax will be negative (a sin subsidy). This captures the spirit of a key result of Bernheim and Rangel (2004) about the optimality of subsidizing addictive goods when the marginal utility of income is increasing with the consumption of the addictive good.[18] Although the Bernheim and Rangel (2004) result that the sin good should be subsidized is seemingly in stark contrast to the sin tax results of O'Donoghue and Rabin (2006) and elsewhere, our general tax formula clarifies the economic forces that lead to each result.

A second key insight from Proposition 1 involves the role of the demand elasticity in governing the relative importance of corrective and redistributive concerns. As is evident from equation (10), when the demand elasticity grows large, the redistributive motive becomes small and the optimal tax t approaches $\bar{\gamma}(\bar{g} + \sigma) + e$, corresponding to Special Case 2. At the opposite extreme, when the elasticity grows small, the corrective motive becomes negligible and the optimal tax approaches the expression in equation (15). More generally, if preference heterogeneity accounts for any share of the decrease in s consumption across incomes, then for a sufficiently low elasticity, the optimal tax becomes negative (a subsidy). Intuitively, if consumers do not respond to commodity taxes, then such taxes become a powerful instrument to enact redistribution through targeted subsidies.

Together, these results show how the price elasticity of demand modulates the role of consumer bias in determining the sign and magnitude of the optimal commodity tax. Perhaps most important, the demand elasticity also provides practical guidance on how sensitive the optimal tax is to different values of the bias $\bar{\gamma}$ and the externality e. A lower elasticity dampens the responsiveness of the optimal tax to the bias $\bar{\gamma}$, because the corrective benefits in equation (10) depend on the products $\bar{\zeta}^c \bar{\gamma}$ and $\bar{\zeta}^c e$.

18. Because in Bernheim and Rangel (2004) overconsumption of the good is a consequence of cue-triggered neural processes that render the consumer inelastic to prices, the average bias of consumers who are elastic to the tax is zero.

Simply put, learning that the average marginal bias or the externality are \$1 per unit higher than previously thought does not imply that the optimal tax should increase by \$1—the optimal adjustment could be higher or lower, depending on the demand elasticity.

Finally, Proposition 1 clarifies the role of revenue recycling— the possibility of using sin tax revenues to offset their regressivity. By including a nonlinear income tax, this model allows for tax revenues to be redistributed in a means-tested fashion. Yet our results also explain why such recycling may not be optimal. If sin good consumption differences are driven by income effects, then the sin tax and income tax cause similar labor supply distortions. In this case, when a corrective sin tax is implemented, the optimal income tax should be jointly reformed to be more progressive, effectively recycling sin tax revenues in a progressive manner. (This corresponds to Special Case 2 and accords with the argument behind the quote from Jim Kenney.) However, if sin good consumption differences are driven by preference heterogeneity, then sin good consumption serves as a tag that is useful for redistribution, even in the presence of the optimal income tax. In this case the optimal sin tax is reduced to effectively subsidize the sin good for redistributive reasons relative to the pure Pigouvian benchmark.

II.F. Optimal Sin Tax at a Fixed Income Tax

Tax authorities may not be able to optimize the income tax system at the same time that a sin tax is imposed and, indeed, the income tax may be suboptimal from their perspective. For example, in the United States many SSB taxes are set by cities that do not control the income tax structure. In such cases, the optimal sin tax satisfies the following condition:

PROPOSITION 2. If Assumptions 1–5 hold, then the optimal commodity tax is approximated by:

$$(16) \qquad t \approx \frac{\bar{\gamma}(1+\sigma) + e - \frac{p}{\bar{s}\bar{\zeta}^c}\left(Cov\left[g(z), s(z)\right] + \mathcal{A}\right)}{1 + \frac{1}{\bar{s}\bar{\zeta}^c}\left(Cov\left[g(z), s(z)\right] + \mathcal{A}\right)},$$

where $\mathcal{A} = \mathbb{E}\left[\frac{T'(z(\theta))}{1 - T'(z(\theta))}\zeta_z(\theta)s(\theta)\xi(\theta)\right]$.

The expression for the constrained optimal SSB tax resembles that in equation (10), but it replaces the term $Cov[g(z), s_{pref}(z)]$ with $Cov[g(z), s(z)] + \mathcal{A}$. This reflects the fact that under the optimal income tax, the fiscal externality from income adjustments in response to the tax (captured by the term $\mathcal{A}$ in Proposition 2) is exactly equal to $Cov[g(z), s_{inc}(z)]$, leaving $Cov[g(z), s_{pref}(z)]$ as the residual term. The term $\mathcal{A}$ is proportional to the income elasticity ξ. Intuitively, when the sin good is a normal good, a higher sin tax is equivalent to an increase in the marginal income tax rate. The converse obtains when the sin good is an inferior good.

II.G. *Multiple Sin Goods and Substitution*

Thus far, our model has involved only a single sin good. We extend our results in a number of ways to account for substitution between sin goods. In Online Appendix B.A we derive a general formula for the optimal set of sin taxes, allowing for a general relationship between cross-price elasticities and biases. In Online Appendix B.B, we also characterize results when the sin good s is a composite good consisting of several different items, such as soft drinks of various sizes.

A third important and practical case is when the policy maker is constrained to tax only one of the sin goods. For example, the policy maker may impose a sin tax on sugary drinks, while not taxing other sugary foods like ice cream and candy bars. We focus on this case here.

Formally, let s denote consumption of the taxed sin good sold at price p and let $r_1, \ldots, r_N$ denote consumption of other sin goods sold at prices $p_1, \ldots p_N$. Let $x_s = ps$ and $x_n = p_n r_n$ denote the pretax expenditures of the respective goods. Define $\varphi(\theta) := -\frac{\sum_n \frac{dx_n(\theta)}{dt}}{\frac{dx_s(\theta)}{dt}}$ to measure how much of the reduction in pretax expenditures on s is reallocated to expenditures on the other sin goods r_n in response to a local increase in the sin tax.

We continue to let γ and e denote the money-metric measures of the bias and externality on s, respectively, and we define γ_n and e_n analogously for the other sin goods r_n. These measures are hard to compare, however, because each is in units of dollars per unit of the respective sin good (e.g., dollars per ounce, dollars per pack). We therefore convert these to unitless measures by dividing by the pretax price of the respective sin good: $\tilde{\gamma}_s := \frac{\gamma}{p}$, $\tilde{e}_s := \frac{e}{p}$, $\tilde{\gamma}_n := \frac{\gamma_n}{p_n}$, $\tilde{e}_n := \frac{e_n}{p_n}$. We then define $\tilde{\gamma}_r(\theta) := \frac{\sum_n \tilde{\gamma}_n(\theta)\frac{dx_n(\theta)}{dt}}{\sum_n \frac{dx_n(\theta)}{dt}}$ as the

expenditure elasticity-weighted average bias of type θ, and we define $\tilde{\gamma}(\theta) := \tilde{\gamma}_s(\theta) - \varphi(\theta)\tilde{\gamma}_r(\theta)$. In words, $\tilde{\gamma}(\theta)$ measures the extent to which consumers overestimate the value of the marginal change in consumption of all sin goods that is induced by an increase in t. Using this definition of $\tilde{\gamma}(\theta)$, we define $\bar{\tilde{\gamma}}$, $\bar{\tilde{\gamma}}(z)$, and $\tilde{\sigma}$ in exact analogue to the single sin good case. Finally, we define $\tilde{e}_r := \frac{\sum_n \tilde{e}_n \frac{d\bar{x}_n}{dt}}{\sum_n \frac{d\bar{x}_n}{dt}}$ as the expenditure-elasticity-weighted externality for untaxed sin goods for the whole population, and we set $\bar{\tilde{e}} := \tilde{e}_s - \frac{\sum_n \frac{d\bar{x}_n}{dt}}{\frac{d\bar{x}_s}{dt}}\tilde{e}_r$ to denote the expenditure-weighted externality, per unit change in pretax expenditures on s. Using these definitions, the formula for the optimal commodity tax is analogous to that in Proposition 1:

PROPOSITION 3. If Assumptions 1–5 hold and all sin goods are a small share of the consumers' total expenditures, then the optimal commodity tax at any fixed tax is approximated by

$$(17) \qquad t \approx p\frac{\bar{\tilde{\gamma}}(1+\tilde{\sigma}) + \bar{\tilde{e}} - \frac{1}{\bar{s}\bar{\zeta}^c}\left(Cov\left[g(z), s(z)\right] + \mathcal{A}\right)}{1 + \frac{1}{\bar{s}\bar{\zeta}^c}Cov\left[g(z), s(z)\right] + \mathcal{A}},$$

where $\mathcal{A} = \mathbb{E}\left[\frac{T'(z(\theta))}{1-T'(z(\theta))}\zeta_z(\theta)s(\theta)\xi(\theta)\right]$. If the income tax is optimal, then the optimal commodity tax is approximated by

$$(18) \qquad t \approx p\frac{\bar{\tilde{\gamma}}(1+\tilde{\sigma}) + \bar{\tilde{e}} - \frac{1}{\bar{s}\bar{\zeta}^c}Cov\left[g(z), s_{pref}(z)\right]}{1 + \frac{1}{\bar{s}\bar{\zeta}^c}Cov\left[g(z), s_{pref}(z)\right]}.$$

We provide a proof of this result, as well as a formula for the optimal income tax, in Online Appendix C.C. The key difference between this formula and the formulas with a single sin good is in the construction of the bias and externality variables $\tilde{\gamma}(\theta)$ and $\tilde{e}$. In the absence of cross-price effects, $p\bar{\tilde{\gamma}} = p\gamma$ and $p\bar{\tilde{e}} = e$, and thus the formula reduces to our initial result. In general, substitution to untaxed sin goods will reduce the total average marginal bias $\bar{\tilde{\gamma}}$ and the average marginal externality $\bar{\tilde{e}}$. For example, if the normalized values of bias and externality are all equal across the sin goods, and if the cross-price expenditure outflow φ is equal to 30%, we would have $p\bar{\tilde{\gamma}} = 0.7\gamma$ and $p\bar{\tilde{e}} = 0.7e$; that is, the corrective benefits are reduced by 30%. On the other hand, if untaxed sin goods are complements—for example, if drinking lowers inhibitions to smoking or using drugs—then the optimal sin tax is higher.

The formula can also be applied to study "leakage": cases where consumers shop outside the jurisdiction to avoid a local sin tax. In this case, we can think of r as the sin good available across the border, with $\gamma_r = \gamma$. Then φ is simply the change in demand for the sin good across the border divided by the change in demand for the sin good within the taxed jurisdiction. Because leakage is typically relevant at the city level but not the national level, optimal tax rates will typically be lower for city-level taxes than for nationwide taxes.

III. Estimating Key Parameters for the Optimal Soda Tax

In this section, we gather the empirical parameters needed to calibrate the optimal nationwide tax on SSBs. First, we describe our data sources. Second, we estimate the price and income elasticities ζ and ξ and how elasticity varies by income. Third, we decompose the SSB consumption versus income relationship into causal income effects $s'_{inc}(z)$ and between-income preference heterogeneity $s'_{pref}(z)$. Fourth, we estimate bias γ, and how this varies by income. Fifth, we discuss the externality e.

See Online Appendix D.A for additional notes on data preparation, and see Online Appendix D.B for an assessment of the strengths and weaknesses of our data and empirical approaches. In Online Appendix D.C we show how our estimating equations can be derived from a class of utility functions.

III.A. Data

1. Nielsen Retail Measurement Services and Homescan Data. The Nielsen Retail Measurement Services (RMS) data include sales volumes and sales-weighted average prices at the UPC-by-store-by-week level at about 37,000 stores each year from 106 retail chains for 2006–2016. RMS includes 53%, 32%, 55%, 2%, and 1% of national sales in the grocery, mass merchandiser, drug, convenience store, and liquor channels, respectively. For a rotating subset of stores that Nielsen audits each week, we also observe merchandising conditions: whether each UPC was "featured" by the retailer in the market where each store is located (through newspaper or online ads and coupons), and whether the UPC was on "display" inside each store.

To measure household grocery purchases, we use the Nielsen Homescan Panel for 2006–2016. Homescan includes about 38,000

TABLE II

DESCRIPTIVE STATISTICS

	Obs.	Mean	Std. dev.	Min	Max
Panel A: Homescan household-by-year data					
Household income ($000s)	653,554	68.00	47.59	3	167
Years education	653,554	13.92	2.06	6	18
Age	653,554	52.24	14.41	18	95
$\mathbb{1}$(White)	653,554	0.77	0.42	0	1
$\mathbb{1}$(Black)	653,554	0.12	0.33	0	1
$\mathbb{1}$(Have children)	653,554	0.33	0.47	0	1
Household size (adult equivalents)	653,554	2.48	1.36	1	11
$\mathbb{1}$(Employed)	653,554	0.61	0.44	0	1
Weekly work hours	653,554	22.84	16.73	0	40
SSBs purchased (liters)	653,554	155.90	192.57	0	13,257
Average price ($/liter)	633,136	1.14	1.45	0	228
Panel B: PanelViews respondent-level data					
Nutrition knowledge	20,640	0.70	0.15	0	1
Self-control	20,640	0.77	0.34	0	1
Other head self-control	13,066	0.67	0.38	0	1
Taste for juice drinks	20,640	0.49	0.32	0	1
Taste for soda	20,640	0.52	0.36	0	1
Taste for tea/coffee	20,640	0.45	0.36	0	1
Taste for sports drinks	20,640	0.29	0.32	0	1
Taste for energy drinks	20,640	0.17	0.28	0	1
Taste for fruit juice	20,640	0.72	0.29	0	1
Taste for diet drinks	20,640	0.32	0.37	0	1
Health importance	20,640	0.84	0.18	0	1
SSB consumption (liters)	20,640	87.70	146.13	0	1,735
$\mathbb{1}$(Male)	20,640	0.28	0.45	0	1
$\mathbb{1}$(Primary shopper)	20,640	0.88	0.33	0	1

Notes. Panel A presents descriptive statistics on the Nielsen Homescan data, which are at the household-by-year level for 2006–2016. If there are two household heads, we use the two heads' mean age, education, employment status, and weekly work hours. We code weekly work hours as 0 for people who are not employed. For people who are employed, weekly work hours is reported in three bins: < 30, 30–34, and ⩾ 35, which we code as 24, 32, and 40, respectively. The U.S. government Dietary Guidelines include calorie needs by age and gender; we combine that with Homescan household composition to get each household member's daily calorie need. Household size in "adult equivalents" is the number of household heads plus the total calorie needs of all other household members divided by the nationwide average calorie need of household heads. Prices and incomes are in real 2016 dollars. Panel B presents descriptive statistics on the Homescan PanelViews data, with one observation for each respondent. See Online Appendix E for the text of the PanelViews survey questions. Observations are weighted for national representativeness.

households in 2006, and about 61,000 households each year for 2007–2016.

Each year, Homescan households report demographic variables such as household income (in 16 bins), educational attainment, household composition, race, binary employment status, and weekly hours worked (in three bins). Table II, Panel A

presents descriptive statistics for Homescan households at the household-by-year level. All households report either one or two heads. If there are two heads, we use their average age, years of education, employment status, and weekly work hours. The U.S. government Dietary Guidelines provide calorie needs by age and gender; we combine that with Homescan household composition to get each household member's daily calorie need. Household size in "adult equivalents" is the number of household heads plus the total calorie needs of all other household members divided by the nationwide average calorie consumption of household heads. In all tables and figures, we weight the sample for national representativeness.

Nielsen groups UPCs into product modules. We define SSBs as the product modules that have typically been included in existing SSB taxes: fruit drinks (which includes sports drinks and energy drinks), premade coffee and tea (for example, bottled iced coffee and iced tea), carbonated soft drinks, and noncarbonated soft drinks (which includes cocktail mixes, breakfast drinks, ice pops, and powdered soft drinks). Fruit and vegetable juice and artificially sweetened drinks such as diet soda are not included. The bottom two rows of Table II, Panel A show that the average Homescan household purchases 156 liters of SSBs per year, at an average price of $1.14 per liter. (Average price paid is undefined for the 3.1% of household-by-year observations with no SSB purchases.) We deflate all prices and incomes to real 2016 dollars.

There are two important differences between Homescan grocery purchase data and total SSB consumption. First, Homescan does not include data on beverages purchased and consumed away from home, such as at restaurants and vending machines. Second, people might give soda to others or throw it out instead of drinking it themselves. For these reasons, we also record total SSB consumption in the survey described below.

2. Homescan PanelViews Survey. We designed a special survey to measure total SSB consumption as well as biases and preferences affecting consumption. Using its PanelViews survey platform, Nielsen fielded the survey in October 2017 to all adult heads of the approximately 60,000 eligible households that were in the 2015 or 2016 Homescan data. We have complete responses from 20,640 people at 18,159 households; there are 2,481 households where both heads responded. Table II, Panel B summarizes the

respondent-level data. Online Appendix E gives the exact text of the survey questions.

We quantify two classes of consumer bias that might drive a wedge between consumers' decisions and normative utility: imperfect nutrition knowledge and imperfect self-control. To measure nutrition knowledge, we delivered 28 questions from the General Nutrition Knowledge Questionnaire (GNKQ).[19] The GNKQ is widely used in the public health literature; see Kliemann et al. (2016) for a validation study. The nutrition knowledge variable is the share correct of the 28 questions; the average score was approximately 0.70 out of 1.

To measure self-control, we asked respondents to state their level of agreement with the following statements: "I drink soda pop or other sugar-sweetened beverages more often than I should," and, if the household has a second head, "The other head of household in my house drinks soda pop or other sugar-sweetened beverages more often than they should." There were four responses: "Definitely," "Mostly," "Somewhat," and "Not at all." To construct the self-control variable, we code those responses as 0, $\frac{1}{3}$, $\frac{2}{3}$, and 1, respectively.

To measure taste and preference heterogeneity, we asked, "Leaving aside any health or nutrition considerations, how much would you say you like the taste and generally enjoy drinking the following?" We asked this question for five types of SSBs (sweetened juice drinks, regular soft drinks, premade coffee and tea, sports drinks, and caffeinated energy drinks) and two non-SSBs (100% fruit juice and diet soft drinks). To measure health preferences, we asked, "In general, how important is it to you to stay healthy, for example by maintaining a healthy weight, avoiding diabetes and heart disease, etc.?" Responses to each question were originally on a scale from 0 to 10, which we rescale to between 0 and 1.

To measure total SSB consumption, we asked people to report how many 12-ounce servings of seven different types of beverages

---

19. One example question is, "If a person wanted to buy a yogurt at the supermarket, which would have the least sugar/sweetener?" The four possible responses were "0% fat cherry yogurt," "Plain yogurt," "Creamy fruit yogurt," and "Not sure." A second example is, "Which is the main type of fat present in each of these foods?" The five possible responses were "Polyunsaturated fat," "Monounsaturated fat," "Saturated fat," "Cholesterol," and "Not sure." This question was asked about olive oil (correct answer: monounsaturated), butter (saturated), sunflower oil (polyunsaturated), and eggs (cholesterol).

they drink in an average week. Finally, we asked gender, age, occupation, and whether the respondent makes the majority of the grocery purchase decisions.

III.B. Price and Income Elasticities

1. Empirical Model. In this section, we estimate the price and income elasticities of demand, ζ and ξ, and how they vary by income. Let s_{it} denote Homescan SSB purchases (in liters per adult equivalent) by household i in quarter t. Let p_{it} denote the price per liter of household i's SSBs in quarter t, and let $\boldsymbol{f}_{it}$ denote the vector of feature and display variables; we detail these variables below. z_{ct} is the mean per capita income reported by the Bureau of Economic Analysis (2017) for county c in the calendar year that contains quarter t, ω_t is a vector of quarter of sample indicators, and μ_{ic} is a household-by-county fixed effect. Our base regression specification to estimate uniform elasticities that do not vary by income is

$$(19) \qquad \ln s_{it} = -\zeta \ln p_{it} + \xi \ln z_{ct} + \nu \boldsymbol{f}_{it} + \omega_t + \mu_{ic} + \varepsilon_{it},$$

with standard errors clustered by county and with $\ln p_{it}$ instrumented in a manner we describe below.[20] To allow elasticities to vary by income, we add linear interaction terms.

Because SSBs are storable, previous purchases could affect current stockpiles and thus current purchases, and Hendel and Nevo (2006) and others document stockpiling in weekly data. In our quarterly data, however, there is no statistically detectable effect of lagged prices and merchandising conditions on current purchases, and it is statistically optimal not to include lags in equation (19) according to the Akaike and Bayesian information criteria. See Online Appendix F for details.

We use county mean income z_{ct} instead of Homescan panelists' self-reported income because we are concerned about measurement error in the within-household self-reported income variation. This is for three reasons: there is likely to be

20. As we show in Online Appendix Table A9, employment status and weekly hours worked are not statistically significantly associated with SSB consumption when included in equation (3), so we cannot reject weak separability of SSB consumption and labor. Thus, it does not matter whether variation in z_{ct} results from nonlabor windfalls such as government benefits or from wage changes, nor does it matter whether such a wage change results from a shift in local labor supply or demand.

measurement error in self-reported year-to-year changes in household income, there is uncertainty as to the time period for which the self-reported incomes apply, and variation in income that is not due to variation in labor market conditions is less likely to be exogenous to preferences for SSBs. However, county income shares some of the same problems and could in principle be correlated with other market prices or consumer preferences. See Online Appendix G for more detailed discussion and alternative estimates using self-reported income.

2. Price, Merchandising Conditions, and the Local Price Deviation Instrument. Household i's SSB price p_{it} is the average price of the UPCs they usually buy at the stores where they usually buy them. Specifically, define p_{ijkt} as the average price that household i pays for UPC k at store j in quarter t. Define s_{ijkc} as household i's total purchases of UPC k at store j while living in county c, and define s_{ic} as household i's total SSB purchases while living in county c, both measured in liters. Then $\pi_{ijkc} = \frac{s_{ijkc}}{s_{ic}}$ is the share of household i's SSB liters purchased while in county c that are of UPC k at store j. Household i's price variable is $p_{it} = \sum_{k,j} \pi_{ijkc} \ln p_{ijkt}$. p_{it} differs from the average price paid per liter because p_{it} does not vary with the household's quantity choices in quarter t, although the results are very similar if we simply use the average price paid.

The feature and display variables are constructed analogously, except using RMS data. Let the two-vector $\boldsymbol{f}_{jkt}$ denote the number of weeks in which UPC k is featured at RMS store j in quarter t, divided by the number of weeks in which feature is observed for that store in that quarter, as well as the analogous share of weeks in which UPC k is observed to be on display at store j. The feature and display variables we use in the household-by-quarter regressions are

$$(20) \qquad \boldsymbol{f}_{it} = \sum_{k,j \in RMS} \pi_{ijkc} \boldsymbol{f}_{jkt}.$$

A key challenge in demand estimation is addressing simultaneity bias: omitted variables bias generated by a potential correlation between price and unobserved demand shifters. We address simultaneity bias using a price instrument leveraging two facts documented by DellaVigna and Gentzkow (forthcoming) and Hitsch, Hortaçsu, and Lin (2017). First, retail chains vary prices

over time in a highly coordinated way across their stores: if retailer X is offering Gatorade on sale right now in Toledo, it's probably also offering Gatorade on sale in Topeka. Second, different chains vary their prices independently of each other over time: retailer X's current sale has little relationship to what other retailers are doing. Online Appendix H illustrates these patterns in more detail.

To construct the instrument, define $\ln p_{jkw}$ as the natural log of the price charged at store j for UPC k in week w. Further define $\ln p_{kw}$ as the national average of natural log price of UPC k in week w, unweighted across stores. Then, let $\ln p_{krt, -c}$ denote the unweighted average of $\ln p_{jkw} - \ln p_{kw}$ at all of retail chain r's stores outside of county c during quarter t. The leave-out construction guarantees that our instrument is not contaminated by store-specific responses to local demand shocks, although in practice the leave-out construction makes little difference because price variation is so coordinated within chains. Differencing out the national average price helps remove responses to national-level demand shocks that might influence the price of the specific UPC k, which could still be a concern even after we condition on time fixed effects ω_t that soak up shocks to overall SSB demand.

To construct an instrument for the average SSB price faced by each household, we fit the leave-out price deviations $\ln p_{krt, -c}$ to the household's average purchasing patterns. Household i's predicted local price deviation in quarter t is

$$(21) \qquad Z_{it} = \sum_{k, j \in RMS} \pi_{ijkc} \ln p_{krt, -c}.$$

Price deviations $\ln p_{krt, -c}$ are only observed at RMS stores, so Z_{it} sums only over purchases at RMS stores; approximately 34% of SSB purchases are at RMS chains. Because π_{ijkc} is the purchase share across all SSB purchases (at both RMS and non-RMS stores), each household's quantity-weighted prices paid, $\ln p_{it}$, should vary approximately one-for-one with Z_{it}.

The exclusion restriction is that the local price instrument Z_{it} is uncorrelated with demand shifters ε_{it}, conditional on the set of controls in equation (19). The economic content of this assumption is that when retail chains vary prices across weeks and quarters, they do not observe and respond to chain-specific demand shocks. One threat to this assumption would be price cuts coordinated with retailer-specific advertising, but retailers do little advertising beyond the newspaper and online ads and coupons that are

already captured by the RMS feature variable (DellaVigna and Gentzkow, forthcoming). Furthermore, we show below that the estimates are largely unaffected by alternative instrument constructions and fixed effect controls that address other types of regional and city-specific demand shocks.

3. Estimation Results. Figure II, Panels A and B, present binned scatterplots of the first stage and reduced form of our instrumental variables (IV) estimates of uniform elasticities, conditioning on the other controls in equation (19). Dividing the reduced-form slope by the first-stage slope implies a price elasticity of approximately $\frac{1.66}{1.21} \approx 1.37$.

Table III presents estimates of equation (19). The first four columns evaluate robustness of the uniform elasticity estimates, and the final column presents estimates allowing elasticities to vary by income. Column (1) presents the primary IV estimates of equation (19), which give estimated price elasticity $\hat{\zeta} \approx 1.37$ and income elasticity $\hat{\xi} \approx 0.20$.

The exclusion restriction would be violated if chains vary prices in response to chain-specific demand shocks. For example, retailers might respond to local economic downturns in cities where they operate or to seasonal variation in soft drink demand that could vary across warm and cold cities. Column (2) addresses these concerns by adding city-by-quarter fixed effects. Demand shocks could also vary across chains serving different demographic groups, for example, if an economic downturn primarily affects low-income households that shop at some retailers more than others. Column (3) addresses this by allowing the city-by-quarter fixed effects to differ for above- versus below-median household income. In both columns, the point estimates move slightly but are statistically indistinguishable.

Although these control strategies can address demand shocks that are common across SSB UPCs, they cannot address UPC-specific demand shocks. For example, a warm spring on the East Coast might increase demand for soft drinks more than it increases demand for bottled coffee. If retailers were to recognize and respond to this, then the subgroup of east coast households that often buy soft drinks would have a positive demand shock and an instrument Z_{it} that is correlated with that shock, even conditional on city-by-time fixed effects. Column (4) addresses this concern by using an instrument constructed with deviations from the census region average log price instead of the national

(A) First Stage
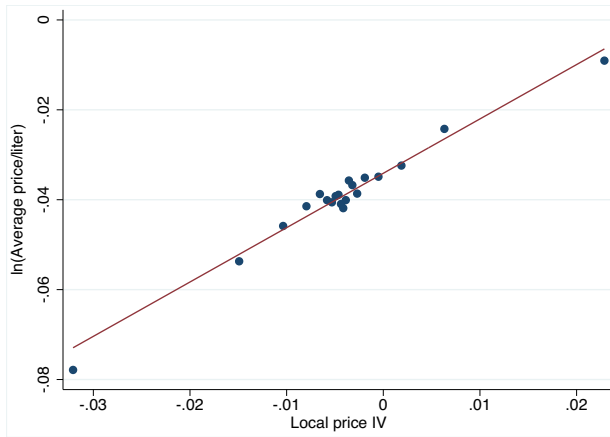
(B) Reduced Form
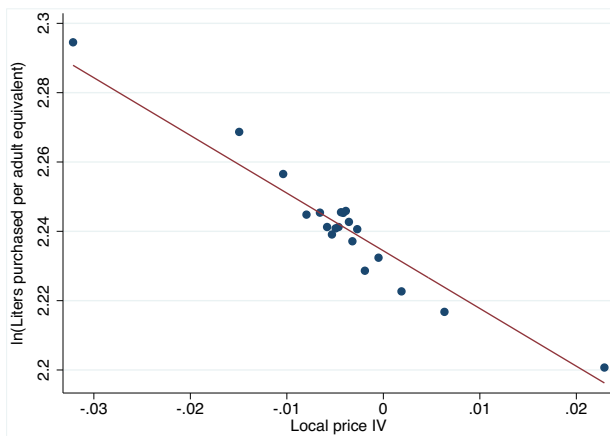
FIGURE II

Contemporaneous First Stage and Reduced Form of the Local Price Instrument

These figures present binned scatterplots of the first stage (in Panel A) and reduced form (in Panel B) of the instrumental variables estimates of equation (19). Both relationships are residual of the other variables in equation (19): feature and display, natural log of county mean income, quarter of sample indicators, and household-by-county fixed effects. Purchases are measured in liters per "adult equivalent," where household members other than the household heads are rescaled into adult equivalents using the recommended average daily consumption for their age and gender group. Observations are weighted for national representativeness.

TABLE III
INSTRUMENTAL VARIABLES ESTIMATES OF PRICE AND INCOME ELASTICITIES

	(1)	(2)	(3)	(4)	(5)
ln(average price/liter)	−1.373***	−1.463***	−1.481***	−1.354***	−1.406***
	(0.089)	(0.098)	(0.099)	(0.091)	(0.168)
ln(county income)	0.204***	0.234***	0.197**	0.201***	0.340***
	(0.073)	(0.080)	(0.080)	(0.073)	(0.126)
Feature	1.154***	1.164***	1.150***	1.161***	1.229***
	(0.065)	(0.067)	(0.066)	(0.067)	(0.136)
Display	0.503***	0.524***	0.510***	0.504***	0.445***
	(0.089)	(0.088)	(0.088)	(0.089)	(0.162)
Income ($100k) × ln(average price/liter)					0.050
					(0.234)
Income ($100k) × ln(county income)					−0.203
					(0.158)
Income ($100k) × feature					−0.106
					(0.172)
Income ($100k) × display					0.088
					(0.217)
Market-quarter fixed effects	No	Yes	No	No	No
Market-quarter-income fixed effects	No	No	Yes	No	No
Kleibergen-Paap first-stage F stat	272.3	256.1	259.3	302.7	42.3
N	2,219,344	2,219,344	2,219,296	2,219,344	2,219,344

Notes. This table presents estimates of equation (19). All regressions include quarter of sample indicators and household-by-county fixed effects. Columns (1)–(3) present instrumental variables estimates using the primary IV. Column (4) constructs the instrument using deviations from regional average prices instead of national average prices. In columns (2) and (3), "market" is Nielsen's Designated Market Area (DMA). Column (5) includes interactions with household i's average income over all years it appears in the sample. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90%, 95%, and 99% confidence, respectively.

average log price. The estimates are again very similar. Thus, for the exclusion restriction to be violated, there must be some specific form of endogeneity not addressed by these multiple alternative specifications.[21]

To measure whether elasticities vary by income, Table III, column (5) presents estimates of equation (19) that include interactions with household i's mean income over the years it appears in the sample. The interaction with price is not statistically significant, although the point estimate suggests that lower-income households are slightly more price elastic: the fitted elasticity is 1.40 at \$5,000 household income and 1.34 at \$125,000. Although low-income consumers are more price elastic in many other product markets, SSBs may be different because lower-income households have much higher demand; SSB demand slopes $\frac{ds}{dp}$ are much steeper at lower incomes. The interaction with income is also not statistically significant, although the point estimate suggests the intuitive result that SSB purchases are less responsive to additional income at higher income levels. For the analysis that follows, we use the fitted values from this column as household-specific price and income elasticities $\hat{\zeta}_i$ and $\hat{\xi}_i$.

4. Substitution to Untaxed Sin Goods. In Section II.G, we derived the optimal SSB tax when complement or substitute sin goods are not taxed. A key statistic for that formula is φ, the share of SSB expenditures that are reallocated to other sin goods in response to an SSB tax increase. We now estimate that statistic.

To keep the scope manageable, we first define a set of goods that are both unhealthy and plausible substitutes or complements to SSBs. We consider all Nielsen product modules averaging more than 15% sugar content by weight. This definition includes everything from the highest-sugar modules (sugar, syrups, sweeteners, etc.) down to moderate-sugar modules such as sauces (pickle relish, ketchup, etc.) and crackers (graham crackers, wafers, etc.).

---

21. The estimates include only observations with positive SSB consumption, as price paid p_{it} is undefined for the 15% of quarterly observations with no SSB purchases. In theory, this can bias our estimates, because high prices are more likely to cause zero-purchase observations. Online Appendix Table A8 addresses this by presenting Tobit estimates (thereby formally accounting for latent demand that is censored at 0) of the reduced form (thereby giving an instrumented price for every observation), with SSB purchases in levels instead of logs (thereby giving a dependent variable for every observation). Price elasticity estimates are economically similar and statistically indistinguishable.

We add diet drinks, because these are likely substitutes even though the health harms are uncertain. We also add alcohol and cigarettes, which could be substitutes or complements to sugary drinks if consumers think of them together as a class of tempting pleasures. We group these into 12 groups, indexed by n, and construct household i's grams purchased in quarter t, r_{nit}, as well as price p_{nit}, instrument Z_{nit}, and feature and display $\boldsymbol{f}_{nit}$ analogous to the SSB variables described above. We estimate the following regression:

(22)
$$\ln r_{nit} = \tilde{\zeta} \ln p_{it} + \zeta_n \ln p_{nit} + \tilde{\xi} \ln z_{ct} + \nu \boldsymbol{f}_{it} + \nu_n \boldsymbol{f}_{nit} + \omega_t + \mu_{ic} + \varepsilon_{it},$$

instrumenting for $\ln p_{it}$ and $\ln p_{nit}$ with Z_{it} and Z_{nit}.

Table IV presents results. The first three columns present substitute beverages, the next eight columns present substitute foods, and the final column presents tobacco. The estimated own-price elasticities $\hat{\zeta}_n$ are in a reasonable range between 0.5 and 2. Unsurprisingly, we find that SSBs and diet drinks are substitutes. Only 1 of the other 11 groups has a statistically significant cross-price elasticity. The average of the 12 cross-price elasticities is a statistically insignificant $-.02$, suggesting that if anything, these other goods are slight complements on average.

Using these estimates, we construct an estimated $\hat{\varphi}_i$ for each household and get the population average, which is $\hat{\varphi} \approx -.03$.[22] Because the health effects of diet drinks are under debate in the public health literature, we construct a second $\hat{\varphi}$ excluding diet drinks ($\hat{\varphi} \approx -.18$). Finally, because diet drinks are the most natural substitute to sugary drinks, while the other estimates may simply be imprecise zeros, we construct a third $\hat{\varphi}$ with only diet drinks ($\hat{\varphi} \approx .15$). In the first two cases, the point estimate of $\hat{\varphi}$ is negative, meaning that if anything, an SSB tax reduces expenditures on these other goods, and accounting for this complementarity will slightly increase the optimal SSB tax. In the final case, substitution to diet drinks will decrease the optimal SSB tax if diet drinks generate internalities and externalities.

---

22. Specifically, we assume $dt = dp_{it}$ and construct $\widehat{\frac{dx_{nit}}{dt}} = \hat{\zeta}_n \frac{x_{nit} p_{nit}}{p_{it}}$, $\widehat{\frac{dx_{si}}{dt}} = \hat{\zeta}_i s_{it}$, and $\hat{\varphi}_{it} = -\frac{\sum_n \widehat{\frac{dx_{nit}}{dt}}}{\frac{dx_s}{dt}}$.

TABLE IV

ESTIMATES OF SUBSTITUTION TO OTHER PRODUCT GROUPS

	Alcohol (1)	Diet drinks (2)	Fruit juice (3)	Baked goods (4)	Baking supplies (5)	Breakfast foods (6)
ln(average price/liter) (SSBs)	0.058	0.248**	0.095	−0.137	0.009	−0.129
	(0.131)	(0.120)	(0.077)	(0.088)	(0.116)	(0.079)
ln(average price/kg)	−1.332***	−0.953***	−1.150***	−1.767***	−1.197***	−1.127***
	(0.250)	(0.120)	(0.153)	(0.167)	(0.106)	(0.119)
ln(county income)	0.131	0.140*	0.131**	0.086	0.063	0.051
	(0.095)	(0.079)	(0.066)	(0.060)	(0.058)	(0.058)
Feature (SSBs)	0.054	−0.307***	0.002	0.055	0.049	0.029
	(0.106)	(0.091)	(0.073)	(0.070)	(0.085)	(0.062)
Display (SSBs)	0.017	−0.130	0.006	0.077	0.029	0.058
	(0.127)	(0.140)	(0.113)	(0.105)	(0.102)	(0.085)
Feature	0.480***	0.830***	1.793***	2.196***	0.962***	1.544***
	(0.100)	(0.063)	(0.094)	(0.126)	(0.099)	(0.060)
Display	0.208*	0.306***	0.759***	1.181***	0.414***	1.379***
	(0.120)	(0.081)	(0.198)	(0.163)	(0.141)	(0.160)
Kleibergen-Paap first-stage F stat	26.6	96.7	143.5	94.0	118.7	131.4
N	913,107	1,128,236	1,701,540	2,004,353	1,408,264	1,816,889
Expenditures ($/adult-quarter)	19.66	6.80	7.23	10.69	2.87	8.89

TABLE IV
CONTINUED

	Candy	Canned, dry fruit	Desserts	Sauces, condiments	Sweeteners	Tobacco
	(7)	(8)	(9)	(10)	(11)	(12)
ln(average price/liter) (SSBs)	−0.113	−0.192**	−0.033	−0.057	−0.069	0.071
	(0.100)	(0.083)	(0.091)	(0.086)	(0.115)	(0.335)
ln(average price/kg)	−1.997***	−1.023***	−0.926***	−0.529***	−1.093***	−1.453
	(0.173)	(0.140)	(0.163)	(0.075)	(0.125)	(1.136)
ln(county income)	0.111*	0.210***	0.035	0.014	−0.122**	0.021
	(0.065)	(0.061)	(0.060)	(0.052)	(0.057)	(0.195)
Feature (SSBs)	0.065	0.053	0.126*	−0.058	−0.015	0.298
	(0.078)	(0.077)	(0.067)	(0.061)	(0.067)	(0.189)
Display (SSBs)	−0.022	−0.136	0.131	−0.026	0.044	0.103
	(0.119)	(0.112)	(0.088)	(0.092)	(0.081)	(0.241)
Feature	3.366***	1.895***	1.753***	0.863***	0.681***	0.547***
	(0.151)	(0.147)	(0.115)	(0.098)	(0.199)	(0.160)
Display	1.800***	0.817***	1.023***	0.078	0.181	0.000
	(0.181)	(0.163)	(0.251)	(0.084)	(0.137)	(.)
Kleibergen-Paap first-stage *F* stat	54.8	53.4	71.8	131.7	45.5	9.9
N	1,969,025	1,453,957	1,931,588	1,077,121	1,595,505	239,106
Expenditures ($/adult-quarter)	11.08	3.22	9.49	0.89	3.13	12.85

Notes. This table presents instrumental variables estimates of equation (22). The product groups are as described in the column headers. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90%, 95%, and 99% confidence, respectively.
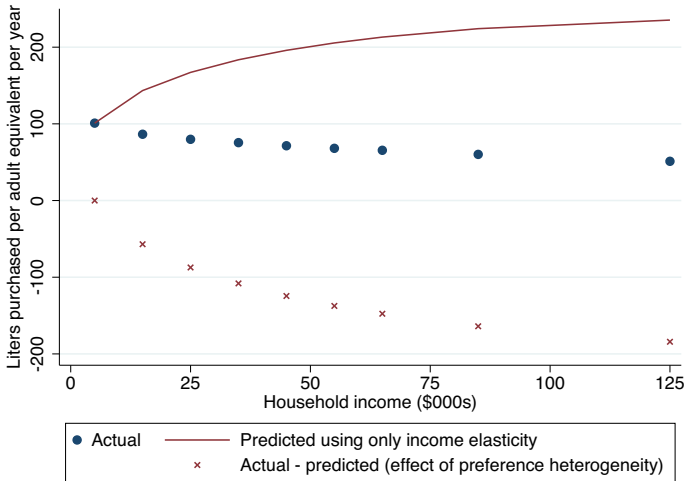
Causal Income Effects and Between-Income Preference Heterogeneity

The circles plot average purchases of sugar-sweetened beverages by household income, using Nielsen Homescan data for 2006–2016. Purchases are measured in liters per "adult equivalent," where household members other than the household heads are rescaled into adult equivalents using the recommended average daily consumption for their age and gender group. The curve at the top of the figure uses the income elasticity estimates from Table III, column (5) to predict the causal effects of income increases on the SSB consumption of households earning less than \$10,000 a year: $s_{inc}(z_d) = \bar{s}(z < \$10k) \prod_{h=2}^{d} \left(\frac{z_h}{z_{h-1}} \right)^{\frac{\hat{\xi}_h + \hat{\xi}_{h-1}}{2}}$, where d and h index income groups. The x's are $\bar{s}_{pref}(z) = \bar{s}(z) - s_{inc}(z)$, the difference between actual consumption and consumption predicted only using income elasticity. Observations are weighted for national representativeness.

III.C. Causal Income Effects versus Between-Income Preference Heterogeneity

The second key empirical statistic needed to determine the optimal sin tax is between-income preference heterogeneity $s'_{pref}(z)$. The dark circles in Figure III repeat the consumption-income relationship from Figure I; this is now compressed due to an expanded y-axis range. The curve at the top of the figure uses the income elasticity estimates from Table III, column (5) to predict the causal effects of income increases on the SSB consumption of households earning less than \$10,000 a year: $s_{inc}(z_d) = \bar{s}(z < \$10,000) \prod_{h=2}^{d} \left(\frac{z_h}{z_{h-1}} \right)^{\frac{\hat{\xi}_h + \hat{\xi}_{h-1}}{2}}$, where d and h

index income groups. Between-income preference heterogeneity is $s_{pref}(z) = \bar{s}(z) - s_{inc}(z)$, the difference between actual consumption and consumption predicted by income effects. On the graph, $s_{pref}(z)$ is thus the vertical difference between the dark circles and the curve.

The estimate of $s_{pref}(z)$ indicates large between-income preference heterogeneity. If all households were exogenously reassigned to earn the same income, households currently making over \$100,000 a year would purchase 184 liters fewer SSBs than households currently making under \$10,000 a year. This difference is about 2.7 times average consumption. This result that lower-income households have stronger preferences for SSBs—regardless of whether they have higher consumption—means that in the absence of internalities and externalities, a policy maker would want to subsidize SSBs.

III.D. *Measuring Bias*

1. The Counterfactual Normative Consumer Estimation Strategy. In Section II, we defined bias γ as the compensated price cut that would induce the counterfactual normative self to consume as much of the sin good as the actual biased self. Our counterfactual normative consumer empirical strategy directly implements this definition, using an approach that builds on Bronnenberg et al. (2015), Handel and Kolstad (2015), and other work.[23] The process is to use surveys to elicit proxies of bias, estimate the relationship between bias proxies and quantity consumed, use that relationship to predict the counterfactual

---

23. Bartels (1996), Cutler et al. (2015), 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. All of these papers require the same identifying assumption: that preferences are conditionally uncorrelated with measures of informedness. Bronnenberg et al. (2015) show that sophisticated shoppers—in their application, doctors and pharmacists—are more likely to buy generic instead of branded drugs and use this to infer that unsophisticated shoppers are making mistakes by not buying generics. The Bronnenberg et al. (2015) identifying assumptions may initially seem more plausible because branded versus generic drugs are close substitutes, whereas consumer tastes for SSBs vary substantially. But if generic drugs are perfect substitutes, then the sophisticated shoppers' decisions are not needed to identify consumer mistakes. The reason to study sophisticated shoppers is to avoid the assumption that generics are perfect substitutes, at which point one must maintain the same assumption that sophisticated and unsophisticated shoppers do not have heterogeneous preferences for the attributes that differentiate branded drugs.

quantity that would be consumed if consumers instead maximized normative utility, and finally transform the quantity difference into dollar units using the price elasticity.

To formalize the approach, recall that money-metric bias γ is defined to satisfy $s(p, y, \theta) = s^V(p - \gamma, y - s\gamma, \theta)$. We log-linearize this equation as described in Online Appendix D.C.1, and we now use an i subscript for each household in the data, recognizing that each i maps to a (p, y, θ) triple. This gives

$$(23) \qquad \ln s_i = \ln s_i^V + \frac{\zeta_i^c \gamma_i}{p_i},$$

where $\ln s_i^V$ denotes the log of the quantity that household i would consume in the absence of bias, s_i and p_i are observed in the Home-scan data, and ζ_i^c is the compensated price elasticity of demand, which we obtain from the Slutsky equation using our estimates of the uncompensated price elasticity and income effects. As an example, imagine that bias increases quantity demanded by 15% and that the compensated demand elasticity is 1.5. Then the impact of bias on consumption is the same as a 10% price reduction: $\gamma_i = p_i \cdot \frac{15\%}{1.5} = 10\% \cdot p_i$.

Let $\boldsymbol{b}_i = [b_{ki}, b_{si}]$ denote a vector of indices measuring household i's bias: nutrition knowledge b_{ki} and self-control b_{si}, as measured in the PanelViews survey. Let $\boldsymbol{b}^V = [b_k^V, b_s^V]$ denote the value of $\boldsymbol{b}$ for a "normative" consumer that maximizes V. $\boldsymbol{a}_i$ is the vector of preferences (beverage tastes and health preferences) measured in the PanelViews survey, $\boldsymbol{x}_i$ is the vector of household characteristics introduced in Table II, and μ_c is a county fixed effect.

We assume that a household's SSB purchases depend on the average biases, preferences, and demographics of all (one or two) household heads.[24] For two-head households, b_{si} is the average of the primary shopper's self-control assessments for herself and the other head. In two-head households where only one head responded, we impute household average nutrition knowledge b_{ki} and preferences $\boldsymbol{a}_i$ based on the observed head's bias proxies and preferences; see Online Appendix I for details.

---

24. Online Appendix Table A11 presents estimates under the alternative assumption that a household's SSB purchases depend on the biases and preferences of the primary shopper only. The pattern of results is very similar, but the coefficient estimates and resulting bias magnitudes are attenuated by about 15%.

Our empirical strategy for estimating γ_i requires three assumptions.

ASSUMPTION 6. Normative consumers: $b_k^V = \mathbb{E}[b_{ki} | dietitian, nutritionist]$, $b_s^V = 1$.

For nutrition knowledge, we set b_k^V equal to the average nutrition knowledge score of the 24 dietitians and nutritionists in the PanelViews survey, which is 0.92. For self-control, we set $b_s^V = 1$: normative consumers are those for whom "not at all" is the correct response to the statement, "I drink soda pop or other sugar-sweetened beverages more often than I should."

ASSUMPTION 7. Linearity: $\frac{\zeta_i^c \gamma_i}{p_i} = \boldsymbol{\tau} \cdot (\boldsymbol{b}^V - \boldsymbol{b}_i)$, where $\boldsymbol{\tau}$ comprises two parameters scaling the effects of nutrition knowledge and self-control.

In our data, linearity is a realistic assumption, as demonstrated in Online Appendix Figure A5.[25]

ASSUMPTION 8. Unconfoundedness: $\boldsymbol{b}_i \perp (\ln s_i^V | \boldsymbol{a}_i, \boldsymbol{x}_i, \mu_c)$.

In words, bias is conditionally independent of normative consumption. Although such unconfoundedness assumptions are often unrealistic, this is more plausible in our setting because of our tailor-made survey measures of beverage tastes and health preferences.

Equation (23) and Assumptions 7 and 8 imply our estimating equation:

$$(24) \qquad \ln(s_i + 1) = \boldsymbol{\tau}\boldsymbol{b}_i + \beta_a \boldsymbol{a}_i + \beta_x \boldsymbol{x}_i + \mu_c + \varepsilon_i.$$

We add 1 to SSB purchases before taking the natural log to include households with zero purchases.

Inserting our parameter estimates into equation (23) and Assumption 7, we obtain estimates of counterfactual normative consumption and money-metric bias:

$$(25) \qquad \log \hat{s}_i^V = \log s_i - \hat{\boldsymbol{\tau}}(\boldsymbol{b}^V - \hat{\boldsymbol{b}}_i)$$

---

25. Linearity is also theoretically plausible, because it results from any "structural" behavioral model in which $\boldsymbol{b}_i$ scales the share of costs that are misperceived—for example, a β, δ model in which consumers downweight future health effects and b_i is proportional to β.

$$(26) \qquad \hat{\gamma}_i = \frac{\hat{\boldsymbol{\tau}}(\boldsymbol{b}^V - \hat{\boldsymbol{b}}_i)p_i}{\hat{\zeta}_i^c}.$$

For these empirical analyses, we use each household's most recent year in the Homescan data, which for about 98% of households is 2016.

2. Descriptive Facts. Figure IV shows that there is a strong unconditional relationship between our bias proxies and SSB purchases. Panel A shows that households whose primary shoppers are in the lowest decile of nutrition knowledge purchase more than twice as many SSBs as households in the highest decile. Panel B shows that households whose primary shoppers answer that they "definitely" drink SSBs "more often than I should" purchase more than twice as many SSBs as households whose primary shoppers answer "not at all." After conditioning on other controls, this is the variation that identifies τ in equation (24).

Figure V shows that nutrition knowledge and self-control are strongly correlated with income. Panel A shows that people with household income above $100,000 score 0.12 higher (0.82 standard deviations) than people with income below $10,000 on the nutrition knowledge questionnaire. Panel B shows that people with income above $100,000 also report about 0.14 higher (0.40 standard deviations) self-control. These relationships suggest that bias is regressive, which augments the corrective benefits of SSB taxes.

Figure VI shows that preferences entering normative utility also differ systematically by income. Panel A shows that relative to people with household income above $100,000, people with income below $10,000 average about 0.09 higher (0.24 standard deviations) in terms of how much they "like the taste and generally enjoy drinking" regular soft drinks. Panel B shows that relative to that highest-income group, the lowest-income group averages about 0.06 points lower (0.36 standard deviations) in their reported importance of staying healthy. Both results imply that lower-income consumers have stronger normative preferences for SSBs. This corroborates the result illustrated in Figure III that the declining consumption–income relationship is driven by preference heterogeneity, not income effects.

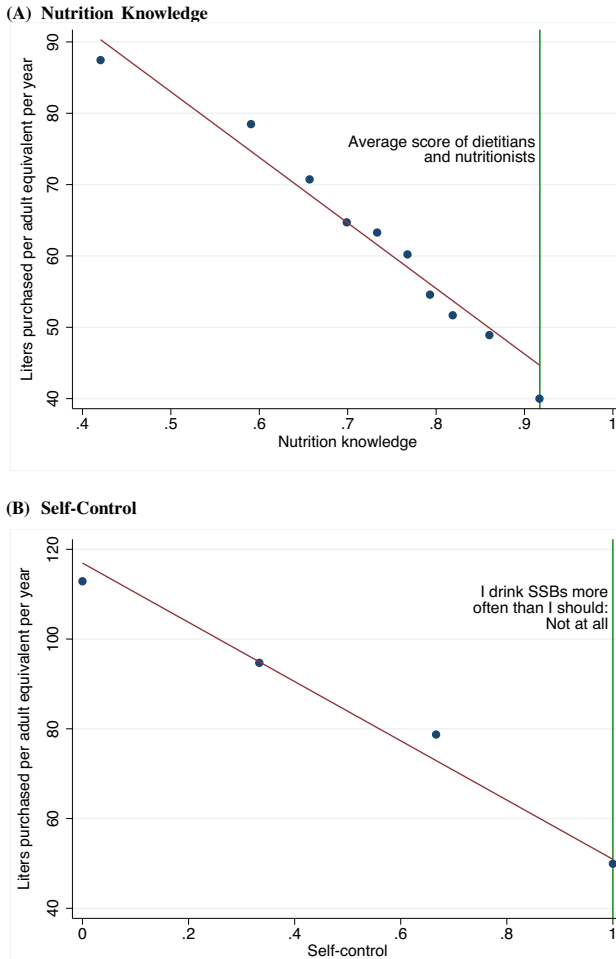(A) Nutrition Knowledge

(B) Self-Control

FIGURE IV

Nutrition Knowledge and Self-Control versus Sugar-Sweetened Beverage Consumption

These figures present average purchases of sugar-sweetened beverages for each household's most recent year in the Nielsen Homescan data against the primary shopper's nutrition knowledge (in Panel A) and self-control (in Panel B). Nutrition knowledge is the share correct out of 28 questions from the General Nutrition Knowledge Questionnaire (Kliemann et al. 2016). Self-control is level of agreement with the statement, "I drink soda pop or other sugar-sweetened beverages more often than I should," with answers coded as "Definitely" = 0, "Mostly" = $\frac{1}{3}$, "Somewhat" = $\frac{2}{3}$, and "Not at all" = 1. Purchases are measured in liters per "adult equivalent," where household members other than the household heads are rescaled into adult equivalents using the recommended average daily consumption for their age and gender group. Observations are weighted for national representativeness.

(A) Nutrition Knowledge
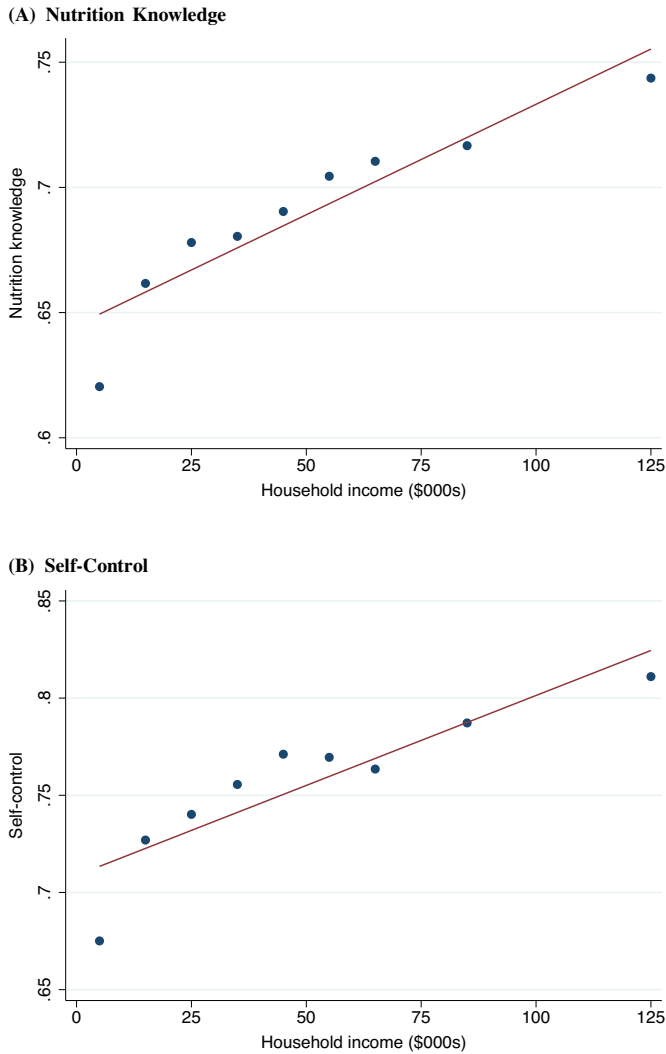
(B) Self-Control

FIGURE V

Nutrition Knowledge and Self-Control by Income

These figures present average nutrition knowledge (in Panel A) and self-control (in Panel B) by household income. Nutrition knowledge is the share correct out of 28 questions from the General Nutrition Knowledge Questionnaire (Kliemann et al. 2016). Self-control is level of agreement with the statement, "I drink soda pop or other sugar-sweetened beverages more often than I should," with answers coded as "Definitely" = 0, "Mostly" = $\frac{1}{3}$, "Somewhat" = $\frac{2}{3}$, and "Not at all" = 1. Observations are weighted for national representativeness.

(A)   Taste for Soda
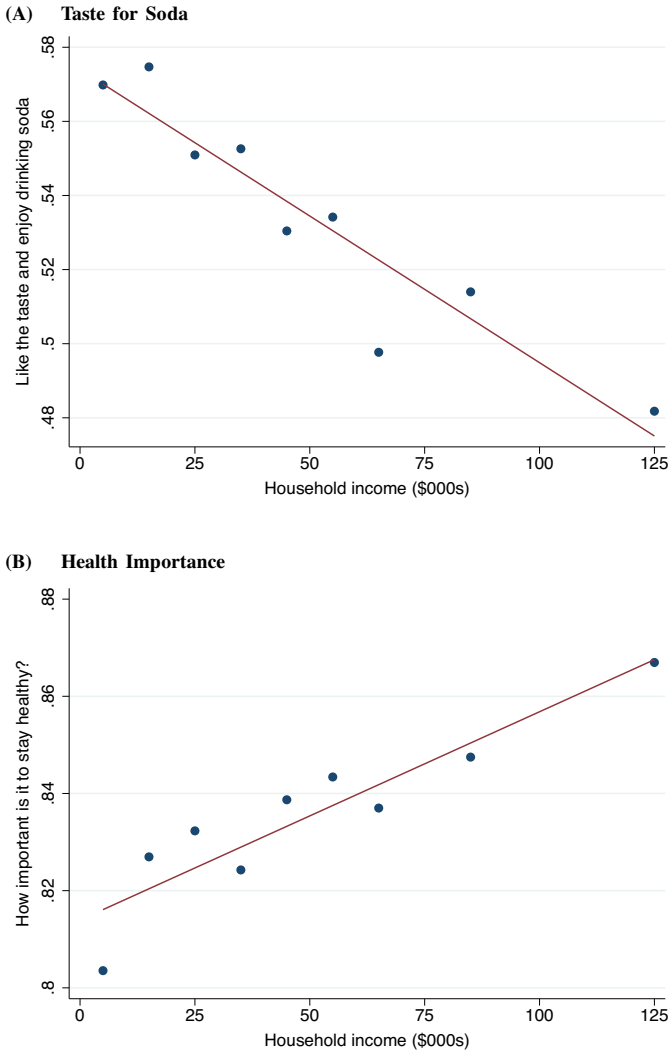

(B)   Health Importance

FIGURE VI

Preferences by Income

These figures present average taste for soda (in Panel A) and health importance (in Panel B) by household income. Taste for soda is the response to the question, "Leaving aside any health or nutrition considerations, how much would you say you like the taste and generally enjoy drinking [Regular soft drinks (soda pop)]?" Health importance is the response to the question, "In general, how important is it to you to stay healthy, for example by maintaining a healthy weight, avoiding diabetes and heart disease, etc.?" Responses to each question were originally on a scale from 0 to 10, which we rescale to between 0 and 1. Observations are weighted for national representativeness.

3. Regression Results. Table V presents estimates of equation (24). Column (1) is our primary specification. Paralleling the unconditional relationships illustrated in Figure IV, both nutrition knowledge and self-control are highly conditionally associated with lower SSB purchases.

There are at least three important reasons to be concerned about this empirical strategy, some of which can be partially addressed in Table V. First, a central concern is our unconfoundedness assumption.[26] Our demographics and taste variables are potentially noisy and incomplete measures of normatively valid preferences, meaning that unobserved preferences might bias the estimated $\hat{\tau}$. To explore this, columns (2)–(4) illustrate coefficient movement: how the $\hat{\tau}$ estimates change with the exclusion of different controls. Preferences, income, and education are correlated with SSB purchases and bias proxies, so it is unsurprising that their exclusion increases the $\hat{\tau}$ in column (2) relative to the primary estimates in column (1). Other demographics (age, race, the presence of children, household size, employment status, and weekly work hours) and county indicators, however, have relatively little effect on $\hat{\tau}$ in columns (3) and (4). This limited coefficient movement in columns (3) and (4) is consistent with the idea that unobservables also have limited impacts on $\hat{\tau}$, although this is certainly not dispositive due to the low R^2 values (Oster 2017).

A second concern is measurement error in the self-control variable. For example, survey respondents with different incomes and SSB demands might not interpret the response categories (not at all, mostly, etc.) in the same way, as highlighted in a related setting by Bond and Lang (2018). As another example, respondents might have interpreted the "more often than I should" phrasing of the question in different ways that don't necessarily reflect bias, for example, that they are optimizing but would aspire to

---

26. An alternative approach that would identify the causal effect of nutrition knowledge would be to run an information provision field experiment, as in Allcott and Taubinsky (2015) or the nutrition education interventions reviewed by Vargas-Garcia et al. (2017). However, this is also an imperfect way to measure γ, as it requires the assumption that the intervention is sufficiently comprehensive and well understood to remove all bias from the treatment group. Furthermore, such experiments in practice involve additional challenges around demand effects and external validity. The finding in Vargas-Garcia et al. (2017) that nutrition information interventions have limited effects could be because lack of nutrition knowledge has little impact on purchases, or it could be because the interventions were incomplete or easily forgotten.

TABLE V

REGRESSIONS OF SUGAR-SWEETENED BEVERAGE CONSUMPTION ON BIAS PROXIES

	(1)	(2)	(3)	(4)	(5)	(6)
Nutrition knowledge	−0.854***	−1.187***	−0.939***	−0.851***	−1.030***	−0.659***
	(0.086)	(0.083)	(0.086)	(0.079)	(0.087)	(0.083)
Self-control	−0.825***	−1.163***	−0.775***	−0.865***		−1.408***
	(0.042)	(0.039)	(0.043)	(0.039)		(0.068)
Taste for soda	0.560***		0.547***	0.553***	0.894***	0.390***
	(0.044)		(0.045)	(0.042)	(0.042)	(0.046)
Health importance	−0.258***		−0.121	−0.275***	−0.388***	−0.184**
	(0.075)		(0.075)	(0.072)	(0.076)	(0.072)
ln(Household income)	−0.045**		−0.077***	−0.066***	−0.055***	−0.024
	(0.018)		(0.017)	(0.017)	(0.019)	(0.017)
ln(Years education)	−0.708***		−0.718***	−0.851***	−0.753***	−0.681***
	(0.101)		(0.101)	(0.096)	(0.103)	(0.096)
Other beverage tastes	Yes	No	Yes	Yes	Yes	Yes
Other demographics	Yes	Yes	No	Yes	Yes	Yes
County indicators	Yes	Yes	Yes	No	Yes	Yes
Self-control 2SLS	No	No	No	No	No	Yes
R^2	0.285	0.250	0.272	0.166	0.263	0.285
N	18,568	18,568	18,568	18,568	18,568	18,568

Notes. This table presents estimates of equation (24). Data are at the household level, and the dependent variable is the natural log of SSB purchases per adult equivalent in the most recent year that the household was in Homescan. Column (6) corrects for measurement error in self-control using two-sample 2SLS, with standard errors calculated per Chodorow-Reich and Wieland (2016). Taste for soda is the response to the question, "Leaving aside any health or nutrition considerations, how much would you say you like the taste and generally enjoy drinking [Regular soft drinks (soda pop)]?" "Other beverage tastes" are the responses to parallel questions for other beverages. Health importance is the response to the question, "In general, how important is it to you to stay healthy, for example by maintaining a healthy weight, avoiding diabetes and heart disease, etc.?" Responses to each question were originally on a scale from 0 to 10, which we rescale to between 0 and 1. "Other demographics" are natural log of age, race, an indicator for the presence of children, household size in adult equivalents, employment status, and weekly work hours. Observations are weighted for national representativeness. Robust standard errors are in parentheses. *, **, ***: statistically significant with 90%, 95%, and 99% confidence, respectively.

something different in the absence of financial or other constraints.[27] In general, measurement error could bias our estimates of γ up or down. One natural model is a type of classical measurement error: noise is uncorrelated with SSB purchases, uncorrelated with income, and uncorrelated across different survey responses within the same household. In this model, our estimated $\hat{\tau}$ for self-control in column (1) would be attenuated toward 0, but we can recover unbiased estimates by instrumenting for self-control.

We address measurement error through several sensitivity analyses. First, we construct γ by halving or doubling the $\hat{\tau}$ coefficient on self-control from column (1). Second, Table V, column (5) simply omits the self-control variable, allowing an estimate of γ that depends only on nutrition knowledge. Because knowledge and self-control are positively correlated, the nutrition knowledge coefficient is stronger in this column. Third, we instrument for self-control using the repeated measurements in the households where two heads responded to the PanelViews survey, using a two-sample two-stage least squares procedure detailed in Online Appendix I. Column (6) presents results. Comparing with column (1), we see that the measurement error correction addresses what would otherwise be substantial attenuation bias in the self-control coefficient. We use the results in this column to construct yet another alternative estimate of γ, which is unbiased under classical measurement error. Positive (negative) correlation in measurement error across household heads would imply that the column (6) estimates are lower (upper) bounds.

A final important concern with the empirical strategy is that we assume our survey measures fully capture the only types of biases that affect SSB consumption. In reality, our measures may be incomplete measures of all types of imperfect knowledge and self-control that could affect SSB consumption. Furthermore, if other biases increase SSB consumption—for example, projection bias or inattention to health harms—then we could understate the optimal SSB tax. We chose these two biases and these specific

27. Online Appendix Table A10 presents additional estimates including an interaction term between knowledge and self-control. This interaction term is highly significant, perhaps because it takes knowledge of health damages to believe that one "should" consume less. Including this interaction in the model, however, does not materially change the estimates of γ.
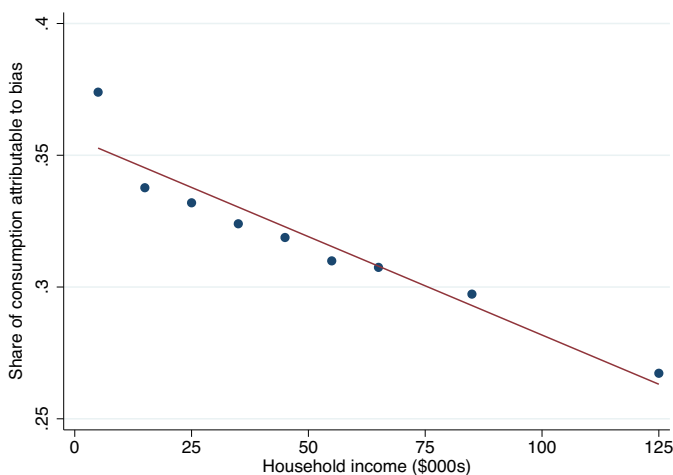
FIGURE VII

Share of Consumption Attributable to Bias by Income

This figure presents the share of sugar-sweetened beverage consumption attributable to bias, that is, the unweighted average of $\frac{s_i - \hat{s}_i^V}{s_i}$, by income, using Homescan purchase data. Observations are weighted for national representativeness.

survey questions because we thought that these best reflected the most relevant and plausible sources of bias.

4. Estimates of Bias. For the average U.S. household, predicted normative SSB consumption from our primary estimates in Table V, column (1) is only $\frac{\hat{s}^V}{s_i} \approx 69\%$ of actual consumption. Put differently, we predict that the average U.S. household would consume $\frac{s_i - \hat{s}_i^V}{s_i} \approx 31\%$ fewer SSBs if they had the nutrition knowledge of dietitians and nutritionists and no self-control problems. Figure VII plots the share of consumption attributable to bias, that is, the unweighted average of $\frac{s_i - \hat{s}_i^V}{s_i}$ across households, by income. Predicted overconsumption is much larger for low-income households: it is 37% and 27%, respectively, for households with income below $10,000 and above $100,000.

Figure VIII plots our primary estimates of the demand slope–weighted average marginal bias $\hat{\bar{\gamma}} = \frac{\sum_i \hat{\zeta}_i^c \frac{s_i}{p_i} \hat{\gamma}_i}{\sum_i \hat{\zeta}_i^c \frac{s_i}{p_i}}$ by income. The average marginal bias across all U.S. households is 0.91 cents per ounce. Since nutrition knowledge and self-control increase with
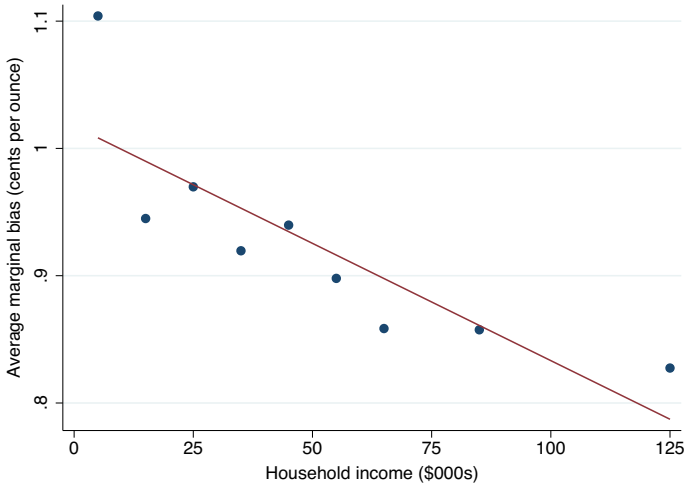
FIGURE VIII

Average Marginal Bias by Income

This figure presents the demand slope–weighted average marginal bias by income, using Homescan purchase data. Observations are weighted for national representativeness.

income and elasticities and prices do not differ much by income, we know from equation (26) that money-metric bias $\hat{\bar{\gamma}}$ will decline in income. Indeed, average marginal biases are 1.10 and 0.83 cents per ounce, respectively, for households with income below \$10,000 and above \$100,000.

In Online Appendix J, we present alternative bias estimates using the PanelViews self-reported SSB consumption. The $\hat{\tau}$ parameters (the associations between SSB consumption and bias proxies) are larger, which makes the bias estimates larger: with the PanelViews data, 37% of the average household's consumption is attributable to bias (48% and 32%, respectively, for household incomes below \$10,000 and above \$100,000), and average marginal bias is 2.14 cents per ounce.

III.E. Externalities

We import an externality estimate from outside sources. Using epidemiological simulation models, Wang et al. (2012) estimate that one ounce of soda consumption increases health care costs by an average of approximately one cent per ounce.

Yong, Bertko, and Kronick (2011) estimate that for people with employer-provided insurance, about 15% of health costs are borne by the individual, while 85% are covered by insurance. Similarly, Cawley and Meyerhoefer (2012) estimate that 88% of the total medical costs of obesity are borne by third parties, and obesity is one of the primary diseases thought to be caused by SSB consumption. Accordingly, we approximate the health system externality at $e \approx 0.85$ cents per ounce.

There are two caveats to this calculation. First, Bhattacharya and Bundorf (2009) find that obese people in jobs with employer-provided health insurance bear the full health costs of obesity through lower wages. However, this result may or may not apply to the other diseases caused by SSB consumption, including diabetes and cardiovascular disease, and it does not apply to people with government-provided health insurance through Medicaid or Medicare. Second, the diseases caused by SSBs might decrease life expectancy, reducing the amount of social security benefits that people claim and thereby imposing a positive fiscal externality (Fontaine et al. 2003; Bhattacharya and Sood 2011). Accounting for these two factors would reduce the externality estimate. Section IV presents optimal tax estimates under alternative assumptions that illustrate the impact of externalities.

IV. Computing the Optimal SSB Tax

We combine the theoretical results from Section II with the empirical estimates from Section III to compute the optimal nationwide tax on SSBs. We compute the optimal tax across a range of specifications, under two different assumptions about the income tax. First, we compute the optimal SSB tax assuming the income tax is held fixed at the current status quo in the United States, using Proposition 2. Second, we compute the optimal SSB tax assuming the income tax is also reformed to be optimal, using equation (10) in Proposition 1.

These computations require an assumption about inequality aversion. We use a schedule of social marginal welfare weights common in the optimal taxation literature (see, for example, Saez 2002b) proportional to $y_{US}^{-\nu}$, where y_{US} is posttax income in the United States, and ν is a parameter that governs the strength of inequality aversion. We use $\nu = 1$ as our baseline, and $\nu = 0.25$ and $\nu = 4$ as our "weak" and "strong" redistributive preferences, respectively. We also report optimal taxes computed under the

assumption that redistributive preferences rationalize the observed U.S. income tax. Calibrations of the status quo U.S. income distribution and income tax are drawn from Piketty, Saez, and Zucman (2018); see Online Appendix M.B for details.

The sufficient statistics formulas for the optimal tax depend on a number of statistics, as well as their covariances with welfare weights. These statistics are reported in Table VI. Panel A presents estimates of the key population-level statistics estimated in Section III, and Panel B presents estimates of statistics within each Homescan income bin. Details of these calculations are reported in the table notes for Table VI. We compute the sufficient statistics involving covariances using the discrete covariance formula reported in the table notes.

Equation (27) shows how these statistics enter the theoretical formula from Proposition 2 for the optimal sin tax under a fixed income tax—this represents our baseline calculation of the optimal SSB tax, which is 1.42 cents per ounce.

$$
\begin{aligned}
t &\approx \frac{\bar{\gamma}(1+\sigma)+e-\frac{p}{\bar{s}\bar{\zeta}^c}\left(Cov\left[g(z),s(z)\right]+\mathbb{E}\left[\frac{T'(z(\theta))}{1-T'(z(\theta))}\zeta_z(\theta)s(\theta)\xi(\theta)\right]\right)}{1+\frac{1}{\bar{s}\bar{\zeta}^c}\left(Cov\left[g(z),s(z)\right]+\mathbb{E}\left[\frac{T'(z(\theta))}{1-T'(z(\theta))}\zeta_z(\theta)s(\theta)\xi(\theta)\right]\right)}, \\
&\approx \frac{\left[0.93(1+0.2)+0.85\right]-\frac{3.63}{46.48\cdot1.39}\left(6.72+0.26\right)}{1+\frac{1}{46.48\cdot1.39}\left(6.72+0.26\right)} \\
&\approx 1.42.
\end{aligned}
$$

(27)

This calculation also provides intuition for the key determinants of the optimal tax. The denominator is close to 1. In the numerator, the corrective motive is equal to $\bar{\gamma}(1+\sigma)+e \approx 0.93(1+0.2)+0.85 \approx 1.97$ cents per ounce. Less than half of the corrective motive is driven by externality correction, as the average marginal bias $\bar{\gamma}$ is larger than the externality. Moreover, the internality correction is further inflated by about 20% due to the bias correction progressivity term σ, reflecting the fact that the benefits of bias correction accrue disproportionately to poorer consumers.

Counteracting this corrective motive, the redistributive motive pushes toward a smaller optimal SSB tax, because the poor have much stronger preferences for SSBs than the wealthy. Using Table VI, we can calculate $Cov[g(z), s(z)] \approx 6.72$ and

TABLE VI

BASELINE OPTIMAL TAX CALCULATION

								Value
Panel A: Population sufficient statistics								
SSB consumption (ounces per week): $\bar{s}$								46.48
SSB price (cents per ounce): p								3.63
SSB demand elasticity: $\bar{\zeta}^c$								1.39
Elasticity of taxable income: $\bar{\zeta}_z$								0.33
Average marginal bias (cents per ounce): $\bar{\gamma}$								0.93
Externality (cents per ounce): e								0.85
z	f	$\bar{s}(z)$	$\bar{\zeta}^c(z)$	$\xi(z)$	$\bar{\gamma}(z)$	$g(z)$	$T'(z)$	$s_{pref}(z)$
Panel B: Calculating covariances								
5,000	0.11	63.1	1.40	0.33	1.07	2.75	−0.19	0.0
15,000	0.16	56.7	1.40	0.31	0.92	1.42	−0.05	−32.9
25,000	0.14	53.3	1.39	0.29	0.91	1.03	0.08	−51.1
35,000	0.10	47.2	1.39	0.27	0.90	0.82	0.15	−67.5
45,000	0.08	44.8	1.38	0.25	0.91	0.69	0.19	−77.6
55,000	0.07	42.9	1.38	0.23	0.90	0.60	0.21	−85.5
65,000	0.09	39.3	1.37	0.21	0.91	0.53	0.21	−93.8
85,000	0.09	35.2	1.36	0.17	0.91	0.43	0.22	−104.8
125,000	0.15	30.3	1.34	0.09	0.85	0.31	0.23	−116.8

$\sigma \approx 0.2$, $Cov[g(z), s_{pref}(z)] \approx 24.8$, $\mathbb{E}\left[\frac{T'(z)}{1-T'(z)} \bar{\zeta}_z \bar{s}(z) \bar{\xi}_{inc}(z)\right] \approx 0.26$

Panel C: Optimal SSB tax

Under existing income tax	$t \approx \dfrac{\bar{s}\bar{\zeta}^c(\bar{\gamma}(\bar{g}+\sigma)+e)-p(Cov[g(z),s(z)]+\mathcal{A})}{\bar{s}\bar{\zeta}^c+Cov[g(z),s(z)]+\mathcal{A}} \approx 1.42$
Under optimal income tax	$t \approx \dfrac{\bar{s}\bar{\zeta}^c(\bar{\gamma}(1+\sigma)+e)-pCov[g(z),s_{pref}(z)]}{\bar{s}\bar{\zeta}^c+Cov[g(z),s_{pref}(z)]} \approx 0.41$

Notes. Panel A reports estimates of population-level sufficient statistics required to compute the optimal SSB tax. All statistics are computed using the data described in Section III, except for the externality e, the calculation of which is described in Section III.E, and the elasticity of taxable income ζ_z^c, which is drawn from Chetty (2012). SSB consumption and price data are computed for 2016. Panel B reports sufficient statistics by income bin, which are used to compute covariances using the formula $Cov[a, b] = \sum_d f_d(a_d b_d) - \sum_d f_d a_d \sum_d f_d b_d$, where d indexes rows. Income bins (z) are those recorded in the Homescan data discussed in Section III.A; f represents the U.S. population share with pretax incomes in ranges bracketed by midpoints between each income bin, according to Piketty, Saez, and Zucman (2018). The statistics $\bar{s}(z)$, $\bar{\zeta}^c(z)$, $\xi(z)$, and $\bar{\gamma}(z)$ represent SSB consumption (in ounces per week), the compensated SSB demand elasticity, the SSB income elasticity, and average marginal money-metric bias estimated within each income bin, as described in Sections III.B to III.D. $\bar{\zeta}^c(z)$ and $\xi(z)$ are computed across incomes using the regression specification reported in Table III, column (5). The column $g(z)$ reports our assumed marginal social welfare weights, while $T'(z)$ represents estimated net marginal tax rates (see text for discussion of each). The column $s_{pref}(z)$ is computed as $\bar{s}(z) - s_{inc}(z)$, where $s_{inc}(z_d) = \bar{s}(z_1) \prod_{h=2}^{d} \left(\frac{z_h}{z_{h-1}}\right)^{\frac{\xi_h+\xi_{h-1}}{2}}$. Panel C uses each of these values to compute the optimal tax estimates using equation (10) and Proposition 2, where $\mathcal{A} = \mathbb{E}\left[\frac{T'(z)}{1-T'(z)} \bar{\zeta}_z \bar{s}(z) \bar{\xi}(z)\right]$.

$\mathbb{E}\left[\frac{T'(z)}{1-T'(z)}\bar{\zeta}_z\bar{s}(z)\bar{\xi}(z)\right] \approx 0.26$. The first term represents the mechanical distributional effect of the tax based on actual SSB consumption, and the second represents the change in income tax revenues due to the effect of the SSB tax on labor supply. Thus the redistributive motive reduces the tax by about 20% relative to the pure corrective motive.

The baseline calculation above holds fixed the status quo income tax. If the optimal income tax is allowed to adjust to be optimal, the impact of the redistributive motive is instead proportional to $-Cov[g(z), s_{pref}(z)]$. This statistic can be computed directly from Table VI, Panel B, where $s_{pref}(z)$ is constructed as the difference between observed SSB consumption $\bar{s}(z)$ and consumption predicted from estimated income elasticities, similarly to Figure III. In this case, the estimated covariance is $Cov[g(z), s_{pref}(z)] \approx 24.8$. Proposition 1 therefore implies an optimal sin tax of 0.41 cents per ounce.

The optimal SSB tax is higher under the status quo income tax than under the optimal income tax because under our assumed welfare weights, status quo marginal income tax rates are "too low" relative to the optimum. Since SSB taxes distort labor supply downward when SSBs are a normal good, they create a negative fiscal externality through the income tax. That negative fiscal externality is much larger under the optimal income tax than under the status quo because the marginal income tax rates of the optimal income tax are much larger than those of the status quo. Consequently, the optimal SSB tax is lower under the optimal income tax.

These estimates of the optimal SSB tax are reported in Table VII, along with calculations under several alternative assumptions that we now summarize.

The second row of Table VII reports the optimal tax estimated using self-reported SSB consumption from our PanelViews survey, rather than data captured by Homescan. This specification results in a higher optimal SSB tax of 2.13 cents per ounce. (All other rows use the Homescan data used for the baseline calculation.)

The next three rows consider alternative assumptions about the policy maker's preference for redistribution. The "Pigouvian" specification reports the optimal tax in the absence of any inequality aversion, in which case the tax is simply equal to $\bar{\gamma} + e$. We report the optimal tax under weaker preferences for redistribution, which lead to a higher tax than in the baseline specification, since

TABLE VII

OPTIMAL SUGAR-SWEETENED BEVERAGE TAX UNDER ALTERNATIVE ASSUMPTIONS

	Existing income tax	Optimal income tax
Baseline	1.42	0.41
Self-reported SSB consumption	2.13	0.96
Pigouvian (no redistributive motive)	1.78	—
Weaker redistributive preferences	1.66	1.35
Stronger redistributive preferences	1.10	−0.64
Redistributive preferences rationalize U.S. income tax	1.73	1.68
Higher demand elasticity ($\zeta^c(\theta) = 2$)	1.57	0.78
Lower demand elasticity ($\zeta^c(\theta) = 1$)	1.23	0.01
Demand elasticity declines faster with income	1.44	0.44
Pure preference heterogeneity	1.44	1.44
Pure income effects	1.49	1.97
Measurement error correction for self-control	1.70	0.64
Internality from nutrition knowledge only	1.00	0.08
Self-control bias set to 50% of estimated value	1.16	0.20
Self-control bias set to 200% of estimated value	1.93	0.82
With substitution: untaxed goods equally harmful	1.48	0.45
With substitution: untaxed goods half as harmful	1.45	0.43
With substitution: untaxed goods doubly harmful	1.53	0.50
With substitution: diet drinks not harmful	1.73	0.66
With substitution: only to diet drinks, equally harmful	1.16	0.20
No internality	0.41	−0.40
No corrective motive	−0.36	−1.01
Optimal local tax, with 25% cross-border shopping	0.97	—
Optimal local tax, with 50% cross-border shopping	0.53	—

Notes. This table reports the optimal sweetened beverage tax, as computed using the sufficient statistics formulas for t^* under the status quo U.S. income tax (using equation (10)) and under the optimal income tax (using Proposition 2) across a range of assumptions. The first row reports our baseline calculations, which employ the sufficient statistics by income bin displayed in Table VI. The second row reports the Pigouvian optimal tax, equal to $\bar{\gamma} + e$. The next two rows report the optimal tax under weaker and stronger redistributive social preferences than the baseline. (Social marginal welfare weights are computed to be proportional to $y_{US}^{-\nu}$, where y_{US} is posttax income in each bin—see Online Appendix M.B for details—with $\nu = 1$, $\nu = 0.25$, and $\nu = 4$ in the baseline, and under "weaker" and "stronger" redistributive preferences, respectively.) "No internality" assumes zero bias for all consumers, and "No corrective motive" assumes zero bias and zero externality. "Self-reported SSB consumption" reports results using SSB consumption data from our PanelViews survey, rather than from Homescan.

the redistributive motive of the tax is weakened. (The corrective motive also decreases, due to smaller bias correction progressivity, but this is a smaller effect.) Similarly, stronger inequality aversion leads to a lower tax than in the baseline. In both cases, the deviation from baseline is greater under the optimal income tax than under the status quo, for the same reason that the baseline SSB tax is lower under the optimal income tax: SSB tax fiscal externalities are more costly under stronger redistributive preferences.

Finally, we consider the assumption that the policy maker's redistributive preferences exactly rationalize the observed U.S. income tax. This assumption implies very weak redistributive motives and social marginal welfare weights that are not even everywhere decreasing with income.[28] Consequently, this assumption raises the tax toward the "Pigouvian" case.[29] However, we are skeptical that the implied redistributive preferences represent deep normative judgments—as opposed to political economy constraints or other factors—and so we use our more conventionally chosen "baseline" weights for the other rows of Table VII.

We consider alternative assumptions about the SSB demand elasticity. A higher elasticity scales down the redistributive motive, raising the optimal tax. Conversely, a lower elasticity reduces the optimal tax. The next two rows of Table VII report the optimal tax assuming a (constant) demand elasticity of either 2 or 1, rather than our heterogeneous empirical estimate. The following row explores the effect of assuming elasticities decline more steeply with income—there we assume that the interaction term on elasticities and household income is four times as large as our estimate from Online Appendix Table A4, while adjusting the intercept to leave the population average elasticity unchanged. This raises the optimal tax, through the bias concentration progressivity term σ, but the effect is muted.

The next two rows consider different possible roles of preference heterogeneity versus income effects in accounting for cross-sectional variation in SSB consumption. The first case, "Pure preference heterogeneity," assumes that all SSB consumption differences are driven by between-income preference heterogeneity. In this case, the optimal SSB tax is independent of the income tax, and so it is the same in both columns. The "Pure income effects" case assumes preferences are homogeneous, implying that SSBs are highly inferior goods. In this case, redistribution is more efficiently carried out through the optimal income tax. This does not substantially alter the optimal SSB tax under the

---

28. See Lockwood and Weinzierl (2016) for a description of this inversion procedure and a discussion of the implied preferences for redistribution.

29. Theory predicts that in a continuous model, the optimal SSB tax would be the same in the two columns, since the existing U.S. income tax is optimal by assumption. Under this discretized calculation, the two values differ slightly, but they are much closer than under the alternative assumptions about redistributive preferences.

(suboptimal) status quo income tax. If the income tax is optimal, however, then the optimal SSB tax is higher than the Pigouvian case of no inequality aversion, as in Special Case 2.

The next four specifications report alternative sets of assumptions about our internality estimates from limited self-control. First, we use a measurement error correction for our estimate of bias from self-control, as in Table V, column (6). The measurement error correction raises the estimated bias from self-control problems, which increases the optimal tax. Second, we report the optimal tax assuming consumers have no self-control problems—that is, assuming bias is driven solely by incorrect nutrition knowledge. This reduces the optimal tax, relative to our baseline. Finally, to reflect the relative uncertainty about the precision of bias due to limited self-control, we report the optimal tax assuming that bias due to limited self-control is either one-half or twice as large as our baseline estimate.

The next five specifications compute the optimal tax accounting for substitution patterns across sweetened goods, using the theoretical formula presented in Proposition 3. These substitution patterns are based on the estimates reported in Table IV. In the specification "With substitution: untaxed goods equally harmful," we assume the categories reported in Table IV are equally harmful to SSBs (in terms of price-normalized externalities and internalities). Since these other categories of goods are estimated to be slightly complementary to SSB consumption on average, accounting for substitution raises the corrective motive of the tax, resulting in a higher optimal tax. We report analogous exercises on the following two lines, assuming the other categories of goods are either half or twice as harmful as SSBs, respectively. Finally, because diet drinks are the one category that is estimated to be a significant substitute for SSBs in Table IV, we consider two possible assumptions about their role. In the specification "With substitution: diet drinks not harmful," we assume that diet drinks are unharmful, with no internalities or externalities, while all other categories are as harmful as SSBs. Because the other categories (excluding diet drinks) are a stronger complement to SSBs as a whole, this assumption implies that the corrective benefits of SSB taxes are larger, resulting in a higher optimal tax. For a contrasting assumption, we note that the insignificant (or barely significant) substitution patterns estimated for goods other than diet drinks in Table IV could be due to statistical noise. Therefore, in the specification "With substitution: only to diet drinks, equally

harmful" we assume that all categories other than diet drinks are neither substitutes nor complements, and that diet drinks are equally harmful to regular SSBs. Because diet drinks are a strong substitute for SSBs, this reduces the corrective strength of an SSB tax, resulting in a lower optimal tax rate.

We consider two more extreme assumptions about internalities and externalities. The row labeled "No internality" reports the optimal tax if consumer bias is assumed to be 0, which substantially reduces the optimal tax. The specification "No corrective motive" also assumes that externalities (in addition to internalities) are 0. In this case, only the redistributive motive is active, resulting in an optimal subsidy for SSBs.

The final two rows beginning with "Optimal local tax" report SSB tax calculations assuming that the tax is implemented at a local level, with some leakage due to cross-border shopping. As noted in Section II.G, the optimal tax in the presence of leakage can be derived from Proposition 3, interpreting cross-border goods as a substitute untaxed good that generates identical internalities and externalities. Although we focus on the optimal nationwide SSB tax for our benchmark analysis, allowing for such leakage may be informative for determining the optimal city-level policy. There are several existing estimates of leakage. Roberto et al. (2019) estimate that about 25% of the total consumption change due to the Philadelphia beverage tax was offset by increased purchases in the surrounding Pennsylvania counties. Seiler, Tuchman, and Song (2019) estimate that offset to be 50% in Philadelphia. Leakage estimates from Philadelphia may be an upper bound, however, as that beverage tax applied to diet beverages, so some cross-border shopping may be replaced by substitution from regular to diet beverages under a conventional SSB tax. Bollinger and Sexton (2019) study the SSB tax in Berkeley and find about half of the consumption reduction is offset by cross-border shopping. We compute the optimal local SSB tax under the assumption that either 25% or 50% of the consumption reduction is offset by cross-border shopping. We do not compute these specifications under the optimal income tax, as it is unclear what assumption should be made about local income taxes.

We compute the welfare gains from SSB taxes in Online Appendix L. In our baseline specification, the optimal tax generates an estimated increase in social welfare of $7.86 per adult equivalent consumer per year, or about $2.4 billion in aggregate across

the United States.[30] This generates \$100 million more in annual welfare gains than a nationwide 1-cent-per-ounce tax—the modal policy among U.S. cities with SSB taxes—and \$1 billion more than the tax that would incorrectly be deemed "optimal" if policy makers do not account for consumer bias. These gains from the optimal tax highlight the importance of using empirical estimates of internalities and externalities, rather than assuming them away or using round number "rules of thumb" to design policy. In the specification using self-reported SSB consumption from the second line of Table VII, the optimal tax generates welfare gains of \$21.86 per adult equivalent consumer per year, or \$6.8 billion across the United States.

The welfare gains can be decomposed into four distinct components, plotted in Figure IX across the income distribution and described in Online Appendix L. The net gains vary across income groups for two competing reasons. First, groups that consume more SSBs have a larger decision utility equivalent-variation loss due to the financial burden of the tax. Second, groups that are more biased and more elastic experience a larger benefit from bias correction because of the tax. As a result, the profile of net benefits from a seemingly "regressive" sin tax can be increasing or decreasing with income. The former force tends to dominate in our baseline estimate, generating an upward-sloping profile of net benefits across the income distribution. That slope is modest, however, and a similar exercise using the PanelViews self-reports—which give a steeper negative slope of bias in income—suggests that the poor benefit nearly as much as the rich; see Online Appendix Figure A6.

A potential concern about the estimates in Table VII, and the corresponding welfare calculations, is that our implementation of the sufficient statistics formulas yield only an approximation to the optimal tax, for two reasons. First, our formulas assume that income effects on labor supply and the budget share of SSBs

---

30. For intuition on the magnitudes, consider the efficiency gains from a purely corrective tax based on our estimated population-level statistics. The sum of the estimated average internality and externality is $\bar{\gamma} + e \approx 0.93 + 0.85 \approx 1.78$ cents per ounce, and the (absolute inverse) slope of the population demand curve is approximated by $\bar{\zeta}^c \frac{\bar{s}}{p} \approx 17.72$. Thus the deadweight loss triangle eliminated by a fully corrective tax is approximated by $0.5(17.72)(1.78)^2 \approx 28.04$ cents per week, or \$14.58 per year. The actual welfare gains are somewhat smaller because of redistributive considerations.
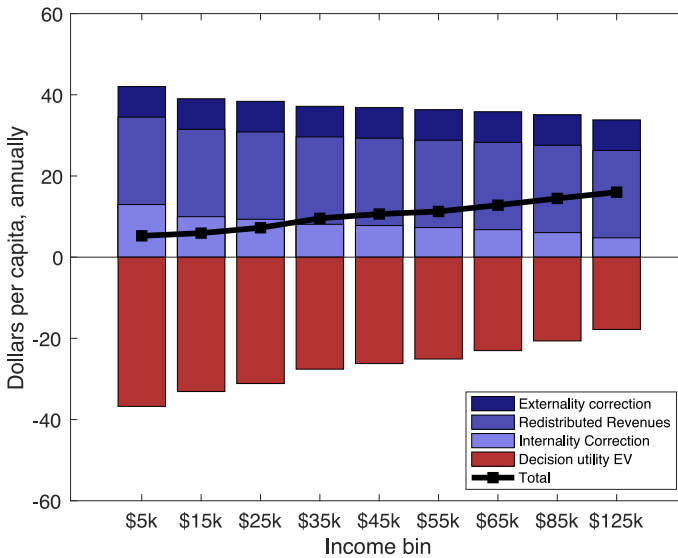
FIGURE IX

Welfare Consequences of Optimal Sugar-Sweetened Beverage Tax

This figure plots the decomposition of welfare changes resulting from the baseline optimal sugar-sweetened beverage tax, across income bins (color version available online). "Redistributed Revenues" and "Externality correction" are money-metric values, assumed to be distributed equally across the income distribution. "Internality Correction" is the increase in (money-metric) welfare due to the change in consumption resulting from the tax, at each income level. "Decision utility EV" is the value dy such that consumers are indifferent between a change in net income dy and the introduction of the optimal SSB tax.

are negligible. Second, we estimate the statistics at the status quo equilibrium, rather than under the optimal tax—that is, we do not account for how a new tax regime would change the consumption of s by income or how it would affect the elasticities. To explore the importance of these sources of error, Online Appendix M presents estimates of the optimal tax using two different structural models, with taxes computed using sufficient statistics at the optimum and fully accounting for all behavioral responses. Those estimates exhibit the same qualitative patterns and are quantitatively close to the values reported here, particularly in the case where the income tax is held fixed, providing additional evidence that an empirically feasible implementation of our sufficient statistics formulas provides a close approximation to the optimal tax.

V. Conclusion

This article provides a tractable theoretical and empirical framework for setting and evaluating optimal commodity taxes in the presence of internalities and externalities. We provide the first optimal commodity tax formula that takes account of the three key elements of public policy debates around sin taxes: correcting consumer bias and externalities, distributional concerns, and revenue recycling through income taxes or income-targeted transfers. Prior work in behavioral economics and public economics has considered only subsets of these three issues or imposed unrealistic assumptions around preference heterogeneity and other parameters.

We demonstrate the usefulness of the theoretical results by focusing on a particularly timely and controversial public policy question: what is the optimal soda tax? Our PanelViews survey data provide novel insights about the relationship between nutrition knowledge, self-control, income, and SSB consumption, and we provide a credible estimate of the price elasticity of demand for SSBs using a new and broadly usable instrument. Our results suggest that externalities and internalities each provide about half of the corrective rationale for SSB taxes, highlighting the importance of attempting to measure internalities. We calculate that the socially optimal nationwide SSB tax is between 1 and 2.1 cents per ounce, or between 28% and 59% of the quantity-weighted average price of SSBs recorded in Homescan. Our preferred estimates imply that the optimal federal tax would increase welfare by $2.4 billion to $6.8 billion a year.

Although we take seriously the possibility that consumers might make mistakes, our methodology fundamentally relies on revealed preference. Our methods are designed to identify the choices people would make if they were fully informed and consumed sugary drinks as much as they feel they actually should. In parallel with other work in behavioral public economics, our approach allows us to continue to use standard tools of public economics to evaluate policies. One alternative approach implicit in much of the public health literature is to assume that the only social objective is to maximize positive health outcomes. However, it is difficult to justify why one should ignore all other factors, such as the benefits of enjoying sugary drinks. A second alternative is to elicit subjective well-being or other measures of "experienced utility," as in the Gruber and Mullainathan (2005) study of the

impact of cigarette taxes on smokers' happiness. However, this approach is fundamentally retrospective, so it cannot be used to evaluate policies that have not yet been implemented. A central problem shared by both alternative approaches is that they do not generate consumer surplus estimates in units of dollars, which are necessary for a comprehensive welfare analysis that also includes producer surplus, externalities, government revenue, and redistributive concerns.

Our theoretical results and empirical methodology could immediately be applied to study optimal taxes (or subsidies) on cigarettes, alcohol, unhealthy foods such as sugar or saturated fat, and consumer products such as energy-efficient appliances. Leaving aside internalities, our results can be used to clarify active debates about the regressivity of externality correction policies such as carbon taxes and fuel economy standards. Our theory is also applicable to questions about capital income taxation or subsidies on saving, and with some appropriate modification our empirical methods could be extended to quantify taxes in those domains as well. Finally, our theory could be extended in a number of potentially fruitful directions, such as allowing for issues of tax salience or incorporating endogenous producer pricing and product line choice.

As we discuss throughout the article, our approach has its weaknesses. One should be cautious about advocating any particular optimal tax estimate too strongly, and we encourage further work extending, generalizing, and critiquing our approach. But by leveraging robust economic principles tied closely to data, our methods almost surely provide valuable input into thorny public policy debates that often revolve around loose intuitions, unsubstantiated assumptions, personal philosophies, or political agendas.

NEW YORK UNIVERSITY, MICROSOFT RESEARCH, AND NBER
UNIVERSITY OF PENNSYLVANIA, WHARTON SCHOOL AND NBER
UNIVERSITY OF CALIFORNIA BERKELEY AND NBER

SUPPLEMENTARY MATERIAL

REFERENCES

Allcott, Hunt, "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market," *American Economic Journal: Economic Policy*, 5 (2013), 30–66.

Allcott, Hunt, Christopher Knittel, and Dmitry Taubinsky, "Tagging and Targeting of Energy Efficiency Subsidies," *American Economic Review*, 105 (2015), 187–201.

Allcott, Hunt, Benjamin B. Lockwood, and Dmitry Taubinsky, "Ramsey Strikes Back: Optimal Commodity Taxes and Redistribution in the Presence of Salience Effects," *American Economic Association Papers and Proceedings*, 108 (2018), 88–92.

———, "Replication Data for: 'Regressive Sin Taxes, with an Application to the Optimal Soda Tax," Harvard Dataverse (2019), doi:10.7910/DVN/STKTU5.

Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky, "Energy Policy with Externalities and Internalities," *Journal of Public Economics*, 112 (2014), 72–88.

Allcott, Hunt, and Dmitry Taubinsky, "Evaluating Behaviorally-Motivated Policy: Experimental Evidence from the Lightbulb Market," *American Economic Review*, 105 (2015), 2501–2538.

Andreyeva, Tatiana, Michael W. Long, and Kelly D. Brownell, "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 (2010), 216–222.

Atkinson, Anthony B., and J. E. Stiglitz, "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics*, 6 (1976), 55–75.

Attari, Shahzeen, Michael DeKay, Cliff Davidson, and Wandi Bruine de Bruin, "Public Perceptions of Energy Consumption and Savings," *Proceedings of the National Academy of Sciences*, 107 (2010), 16054–16059.

Bartels, Larry M., "Uninformed Votes: Information Effects in Presidential Elections," *American Journal of Political Science*, 40 (1996), 194–230.

Bernheim, B. Douglas, and Antonio Rangel, "Addiction and Cue-Triggered Decision Processes," *American Economic Review*, 94 (2004), 1558–1590.

Bhattacharya, Jay, and M. Kate Bundorf, "The Incidence of the Healthcare Costs of Obesity," *Journal of Health Economics*, 28 (2009), 649–658.

Bhattacharya, Jay, and Neeraj Sood, "Who Pays for Obesity?," *Journal of Economic Perspectives*, 25 (2011), 139–158.

Blumgart, Jake, "A Soda Tax is Fundamentally Regressive. What If It's Philadelphia's Best Hope for Progressive Change?," *The Nation* (May 23, 2016).

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen, "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 3 (2011), 91–128.

Bollinger, Bryan, and Steven Sexton, "Local Excise Taxes, Sticky Prices, and Spillovers: Evidence from Berkeley's Soda Tax," SSRN working paper, 2019.

Bond, Timothy N., and Kevin Lang, "The Sad Truth about Happiness Scales: Empirical Results," NBER Working Paper No. 24853, 2018.

Bovenberg, A. Lans, and Lawrence H. Goulder, "Optimal Environmental Taxation in the Presence of Other Taxes: General-Equilibrium Analyses," *American Economic Review*, 86 (1996), 985–1000.

Bronnenberg, Bart, Jean-Pierre Dubé, Matthew Gentzkow, and Jesse Shapiro, "Do Pharmacists Buy Bayer? Sophisticated Shoppers and the Brand Premium," *Quarterly Journal of Economics*, 130 (2015), 1669–1726.

Bureau of Economic Analysis, *Personal Income Summary: Personal Income, Population, Per Capita Personal Income* (Washington: BEA, 2017).

Cawley, John, and Chad Meyerhoefer, "The Medical Care Costs of Obesity: An Instrumental Variables Approach," *Journal of Health Economics*, 31 (2012), 219–230.

Chetty, Raj, "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 80 (2012), 969–1018.

———, "Behavioral Economics and Public Policy: A Pragmatic Perspective," *American Economic Review*, 105 (2015), 1–33.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99 (2009), 1145–1177.

Chodorow-Reich, Gabriel, and Johannes Wieland, "Secular Labor Reallocation and Business Cycles," NBER Working Paper No. 21864, 2016.

Cutler, David, Amber Jessup, Donald Kenkel, and Martha Starr, "Valuing Regulations Affecting Addictive or Habitual Goods," *Journal of Benefit Cost Analysis*, 6 (2015), 247–280.

Davis, Lucas, and Severin Borenstein, "The Distributional Effects of U.S. Clean Energy Tax Credits," *Tax Policy and the Economy*, 30 (2016), 191–234.

Davis, Lucas, and Christopher Knittel, "Are Fuel Economy Standards Regressive?," NBER Working Paper No. 22925, 2016.

DellaVigna, Stefano, and Matthew Gentzkow, "Uniform Pricing in US Retail Chains," *Quarterly Journal of Economics* (forthcoming).

Diamond, Peter A., "Consumption Externalities and Imperfect Corrective Pricing," *Bell Journal of Economics and Management Science*, (1973), 526–538.

———, "A Many-Person Ramsey Tax Rule," *Journal of Public Economics*, 4 (1975), 335–342.

Diamond, Peter, and Johannes Spinnewijn, "Capital Income Taxes with Heterogeneous Discount Rates," *American Economic Journal: Economic Policy*, 3 (2011), 52–76.

Dubois, Pierre, Rachel Griffith, and Martin O'Connell, "How Well-Targeted Are Soda Taxes?," Institute for Fiscal Studies working paper, 2017.

Duffey, Kiyah J., Penny Gordon-Larsen, James M. Shikany, David Guilkey, David R. Jacobs, and Barry M. Popkin, "Food Price and Diet and Health Outcomes," *Archives of Internal Medicine*, 170 (2010), 420–426.

Farhi, Emmanuel, and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," NBER Working Paper No. 21524, 2015.

Finkelstein, Eric A., Chen Zhen, Marcel Bilger, James Nonnemaker, Assad M. Farooqui, and Jessica E. Todd, "Implications of a Sugar-Sweetened Beverage (SSB) Tax When Substitutions to Non-Beverage Items Are Considered," *Journal of Health Economics*, 32 (2013), 219–239.

Fletcher, Jason M., David E. Frisvold, and Nathan Teft, "The Effects of Soft Drink Taxes on Child and Adolescent Consumption and Weight Outcomes," *Journal of Public Economics*, 94 (2010), 967–974.

Fontaine, Kevin R., David T. Redden, Chenxi Wang, Andrew O. Westfall, and David B. Allison, "Years of Life Lost Due to Obesity," *Journal of the American Medical Association*, 289 (2003), 187–193.

Gauthier, Stephane, and Fanny Henriet, "Commodity Taxes and Taste Heterogeneity," *European Economic Review*, 101(C) (2018), 284–296.

Global Food Research Program (GFRP), "Sugary Drink Taxes around the World," GFRP, 2019.

Goldin, Jacob, and Tatiana Homonoff, "Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity," *American Economic Journal: Economic Policy*, 5 (2013), 302–336.

Goulder, Lawrence H., "Environmental Taxation and the Double Dividend: A Reader's Guide," *International Tax and Public Finance*, 2 (1995), 157–183.

———, "Climate Change Policy's Interactions with the Tax System," *Energy Economics*, 40 (2013), 3–11.

Goulder, Lawrence H., and Roberton C. Williams, "The Substantial Bias from Ignoring General Equilibrium Effects in Estimating Excess Burden, and a Practical Solution," *Journal of Political Economy*, 111 (2003), 898–927.

Gruber, Jonathan, and Botond Kőszegi, "Is Addiction 'Rational'? Theory and Evidence," *Quarterly Journal of Economics*, 116 (2001), 1261–1303.

———, "Tax Incidence When Individuals Are Time-Inconsistent: The Case of Cigarette Excise Taxes," *Journal of Public Economics*, 88 (2004), 1959–1987.

Gruber, Jonathan, and Sendhil Mullainathan, "Do Cigarette Taxes Make Smokers Happier?," *B.E. Journal of Economic Analysis & Policy*, 5 (2005).

Gruber, Jon, and Emmanuel Saez, "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 84 (2002), 1–32.

Handel, Benjamin R., and Jonathan T. Kolstad, "Health Insurance for 'Humans': Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 105 (2015), 2449–2500.

Harding, Matthew, and Michael Lovenheim, "The Effect of Prices on Nutrition: Comparing the Impact of Product- and Nutrient-Specific Taxes," *Journal of Health Economics*, 53 (2015), 53–71.

Hausman, Jerry, "Valuation of New Goods under Perfect and Imperfect Competition," In *The Economics of New Goods*, vol. 58, Timothy Bresnahan and Roger Gordon, eds. (Chicago: University of Chicago Press, 1996), 207–248.

Hendel, Igal, and Aviv Nevo, "Sales and Consumer Inventory," *RAND Journal of Economics*, 37 (2006), 543–561.

Heutel, Garth, "Optimal Policy Instruments for Externality-Producing Durable Goods under Present Bias," *Journal of Environmental Economics and Management*, 72 (2015), 54–70.

Hitsch, Günter J., Ali Hortaçsu, and Xiliang Lin, "Prices and Promotions in U.S. Retail Markets: Evidence from Big Data," Chicago Booth Research Paper No. 17-18, 2017.

Huehnergarth, Nancy Fink, "Senator Sanders Is Wrong: A Soda Tax Is Progressive, Not Regressive," *Forbes* (April 25, 2016).

Jacobs, Bas, and Robin Boadway, "Optimal Linear Commodity Taxation under Optimal Non-Linear Income Taxation," *Journal of Public Economics*, 117 (2014), 201–210.

Jacobs, Bas, and Ruud A. de Mooij, "Pigou Meets Mirrlees: On the Irrelevance of Tax Distortions for the Second-Best Pigouvian Tax," *Journal of Environmental Economics and Management*, 71 (2015), 90–108.

Jacquet, Laurence, and Étienne Lehmann, "Optimal Taxation with Heterogeneous Skills and Elasticities: Structural and Sufficient Statistics Approaches," THEMA Working Paper No. 2016-04, 2016.

Johnson, Erin M., and M. Marit Rehavi, "Physicians Treating Physicians: Information and Incentives in Childbirth," *American Economic Journal: Economic Policy*, 8 (2016), 115–141.

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

Kliemann, N., J. Wardle, F. Johnson, and H. Croker, "Reliability and Validity of a Revised Version of the General Nutrition Knowledge Questionnaire," *European Journal of Clinical Nutrition*, 70 (2016), 1174–1180.

Levitt, Steven D., and Chad Syverson, "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions," *Review of Economics and Statistics*, 90 (2008), 599–611.

Lockwood, Benjamin B., and Dmitry Taubinsky, "Regressive Sin Taxes," NBER Working Paper No. 23085, 2017.

Lockwood, Benjamin B., and Matthew Weinzierl, "Positive and Normative Judgments Implicit in US Tax Policy, and the Costs of Unequal Growth and Recessions," *Journal of Monetary Economics*, 77 (2016), 30–47.

Mullainathan, Sendhil, Joshua Schwartzstein, and William J. Congdon, "A Reduced-Form Approach to Behavioral Public Finance," *Annual Review of Economics*, 4 (2012), 511–540.

Nevo, Aviv, "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69 (2001), 307–342.

O'Donoghue, Ted, and Matthew Rabin, "Optimal Sin Taxes," *Journal of Public Economics*, 90 (2006), 1825–1849.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 37 (2017), 1–18.

Pigou, Arthur Cecil, "*The Economics of Welfare*" (London: Macmillan, 1920).

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, "Distributional National Accounts: Methods and Estimates for the United States," *Quarterly Journal of Economics*, 133 (2018), 553–609.

Powell, Lisa M., Jamie F. Chriqui, Tamkeen Khan, Roy Wada, and Frank J. Chaloupka, "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 (2013), 110–128.

Roberto, Christina A., Hannah G. Lawman, Michael T. LeVasseur, Nandita Mitra, Ana Peterhans, Bradley Herring, and Sara N. Bleich, "Association of a Beverage Tax on Sugar-Sweetened and Artificially Sweetened Beverages With Changes in Beverage Prices and Sales at Chain Retailers in a Large Urban Setting," *Journal of the American Medical Association*, 321 (2019), 1799–1810.

Rojas, Christian, and Emily Wang, "Do Taxes for Soda and Sugary Drinks Work? Scanner Data Evidence from Berkeley and Washington," SSRN working paper, 2017.

Saez, Emmanuel, "The Desirability of Commodity Taxation under Non-Linear Income Taxation and Heterogeneous Tastes," *Journal of Public Economics*, 83 (2002a), 217–230.

———, "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117 (2002b), 1039–1073.

Saez, Emmanuel, Joel Slemrod, and Seth H. Giertz, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 50 (2012), 3–50.

Sanders, Bernie, "A Soda Tax Would Hurt Philly's Low-Income Families," *Philadelphia Magazine* (April 24, 2016).

Seiler, Stephan, Anna Tuchman, and Song Yao, "The Impact of Soda Taxes: Passthrough, Tax Avoidance, and Nutritional Effects," Stanford University Graduate School of Business Research Paper No. 19-12, 2019.

Silver, Lynn D., Shu Wen Ng, Suzanne Ryan-Ibarra, Lindsey Smith Taillie, Marta Induni, Donna R. Miles, Jennifer M. Poti, and Barry M. Popkin, "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 (2017).

Smith, Travis A., Biing-Hwan Lin, and Jonq-Ying Lee, "Taxing Caloric Sweetened Beverages: Potential Effects on Beverage Consumption, Calorie Intake, and Obesity," USDA Economic Research Report No. 100, 2010.

Taubinsky, Dmitry, and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies* (forthcoming).

Thow, Anne Marie, Shauna Downs, and Stephen Jan, "A Systematic Review of the Effectiveness of Food Taxes and Subsidies to Improve Diets: Understanding the Recent Evidence," *Nutrition Reviews*, 72 (2014), 551–565.

Tiffin, Richard, Ariane Kehlbacher, and Matthew Salois, "The Effects of A Soft Drink Tax in the UK," *Health Economics*, 24 (2015), 583–600.

Vargas-Garcia, E. J., C. E. L. Evans, A. Prestwich, B. J. Sykes-Muskett, J. Hooson, and J. E. Cade, "Interventions to Reduce Consumption of Sugar-Sweetened Beverages or Increase Water Intake: Evidence from a Systematic Review and Meta-Analysis," *Obesity Reviews*, 18 (2017), 1350–1363.

Wang, Emily Yucai, "The Impact of Soda Taxes on Consumer Welfare: Implications of Storability and Taste Heterogeneity," *RAND Journal of Economics*, 46 (2015), 409–441.

Wang, Y. Claire, Pamela Coxson, Yu-Ming Shen, Lee Goldman, and Kirsten Bibbins-Domingo, "A Penny-per-Ounce Tax on Sugar-Sweetened Beverages Would Cut Health and Cost Burdens of Diabetes," *Health Affairs*, 31 (2012), 199–207.

Yong, Pierre L., John Bertko, and Richard Kronick, "Actuarial Value and Employer-Sponsored Insurance," Research brief, Office of the Assistant Secretary for Planning and Evaluation, 2011.

Zhen, Chen, Eric A. Finkelstein, James Nonnemaker, Shawn Karns, and Jessica E. Todd, "Predicting the Effects of Sugar-Sweetened Beverage Taxes on Food and Beverage Demand in a Large Demand System," *American Journal of Agricultural Economics*, 96 (2014), 1–25.

Zhen, Chen, Michael K. Wohlgenant, Shawn Karns, and Phillip Kaufman, "Habit Formation and Demand for Sugar-Sweetened Beverages," *American Journal of Agricultural Economics*, 93 (2011), 175–193.