Regressive Sin Taxes, with an Application to the Optimal Soda Tax

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

April 29, 2018

Abstract

A common objection to “sin taxes”—corrective taxes on goods that are thought to be over-consumed, 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. We show that the implications of regressivity hinge on why consumption decreases with income. If the consumption-income relationship is driven by income effects, then regressivity is optimally offset by targeted transfers or income tax reforms, not by moderating the level of the sin tax. If the relationship is instead driven by between-income preference heterogeneity, the optimal sin tax depends on the demand elasticity: if demand is more elastic, then progressive benefits from reduced over-consumption can make the optimal sin tax larger than if there were no distributional concerns, while if demand is less elastic, the optimal tax is reduced. 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. Our empirical estimates of elasticities, preference heterogeneity, incorrect beliefs, and self-control imply that current city-level taxes in Berkeley, San Francisco, and elsewhere are actually lower than the social optimum.

*Allcott: New York University and NBER. hunt.allcott@nyu.edu. Taubinsky: Berkeley and NBER. dmitry.taubinsky@berkeley.edu. Lockwood: Wharton and NBER. ben.lockwood@wharton.upenn.edu. We thank Kelly Brownell, Rebecca Diamond, Jean-Pierre Dube, Matt Gentzkow, Anna Grummon, David Laibson, Alex Rees-Jones, Christina Roberto, Claire Wang, Danny Yagan, and seminar participants at Berkeley, Carnegie Mellon and the University of Pittsburgh, Columbia, Davis, Hong Kong University of Science and Technology, the National Tax Association, National University of Singapore, the NBER Public Economics 2018 Spring Meetings, NYU, Princeton, the Society for Benefit-Cost Analysis, University of British Columbia, University of Virginia, and Wharton for helpful feedback. We are grateful to the Sloan Foundation for grant funding. The survey was determined to be exempt from review by the Institutional Review Boards at the University of Pennsylvania (protocol number 828341) and NYU (application FY2017-1123). This paper reflects the authors’ own analyses and calculations based in part on data reported by Nielsen through its Homescan, RMS, and PanelViews services for beverage categories over 2006-2015, for all retail channels in the U.S. market. The conclusions drawn from the Nielsen data are those of the authors and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Replication files are available from https://sites.google.com/site/allcott/research. This paper subsumes and replaces Lockwood and Taubinsky (2017).
“The only way to protect all of us, including the poor, from further harm is through a sugary drink tax....”
– Forbes magazine article (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 of 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 over-consumption of goods such as cigarettes, alcohol, unhealthy foods, and energy inefficient durable goods. Consequently, “sin taxes” that discourage consumption of such goods could increase social welfare.¹ This corrective logic justifies widespread taxation of cigarettes and alcohol, as well as newer taxes on sugar-sweetened beverages in France, Ireland, Mexico, Norway, the United Kingdom, and U.S. cities including Berkeley, Philadelphia, San Francisco, and Seattle.

What is the optimal level of such sin taxes? The existing literature frequently invokes a corrective logic dating to Pigou (1927) and Diamond (1972): the optimal bias-correcting tax equals the average mistake (or “internality”) of marginal consumers.² This principle, however, assumes that consumers do not vary in their marginal utility of money, and thus that policymakers 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 above quote from Senator Bernie Sanders, a common objection to sin taxes is that they are regressive.³

In response to such objections, however, others argue that the harms caused by overconsumption are themselves regressive, so a corrective tax might confer greater benefits on the poor than the rich. For example, smoking and sugary drink consumption cause lung cancer, diabetes, and other

¹See, e.g., Chetty (2015), Mullainathan et al. (2012), and Bernheim and Rangel (2009), for an overview of the various biases and their implications for welfare calculations.
³The poor disproportionately consume cigarettes and sugary drinks, while the rich disproportionately take up energy efficiency subsidies. See, e.g., Gruber and Kőszegi (2004) and Goldin and Homonoff (2013) on cigarettes, and Allcott, Knittel, and Taubinsky (2015), Davis and Borenstein (2016), and Davis and Knittel (2016) on energy efficiency subsidies and fuel economy standards.
health problems that disproportionately affect the poor.\textsuperscript{4} Furthermore, regressivity can be offset through progressive revenue recycling: sin tax revenues can be used to fund progressive initiatives and transfers to the poor. As illustrated in the introductory quotes, this is a central argument in favor of soda taxes in Philadelphia and elsewhere.\textsuperscript{5}

In the first part of this paper, we present a general theoretical model that delivers the first explicit formula for an optimal commodity tax that accounts for each of the above arguments: correction of consumer bias, regressivity, and revenue recycling. In the second part, we estimate the key empirical parameters required by the theory, using sugar-sweetened beverages (hereafter, “SSBs”) as our application. In the third part of the paper, we combine the theory and empirics to calculate the optimal nationwide SSB tax in the U.S.

Our theoretical model in Section II builds on Saez’s (2002a) extension of Atkinson and Stiglitz (1976) by considering an economy of consumers with heterogenous earning abilities and tastes, who choose labor supply and a consumption bundle that exhausts their after-tax income. The policymaker chooses a set of linear commodity taxes and a non-linear income tax which can be used to provide transfers to poor consumers, to raise money for commodity subsidies, or to distribute commodity tax revenue (in a progressive way, if desired). But whereas the standard approach in optimal tax theory assumes that the planner agrees with consumers about what is best for them, we analyze the general case in which the policymaker may disagree with consumers’ choices due to “mistakes” such as present bias, inattention, or incorrect beliefs. Our general approach to modeling bias nests a variety of behavioral economics models of biases as special cases and allows the model to generate insights for many policy-relevant domains. Optimal tax policy in the presence of externalities is obtained as a corollary of our more general results.

In developing our theoretical results, we address two challenges. The first challenge is that even in the absence of internalities or externalities, formulas for optimal commodity taxes in the presence of income taxation (and arbitrary preference heterogeneity) have remained an open problem in public economics.\textsuperscript{6} We resolve this longstanding question with a simple and empirically implementable sufficient statistics formula that, in the absence of internalities externalities, resembles the many-person Ramsey tax rule of Diamond (1975). We build on Saez’s (2002) qualitative insight that preference heterogeneity correlated with earnings ability can make it optimal to raise

\textsuperscript{4}Gruber and Kőszegi (2004) make this argument formally in the context of cigarette taxes. Less formally, Huchnergearth (2016) writes, “Here’s why Sanders’ position on sugary drink taxes plays right into the hands of Big Soda, and hurts America’s most vulnerable citizens: Big Soda’s targeted marketing to communities of color and low-income communities is regressive.... Type 2 diabetes, linked to excessive sugary drink consumption, is regressive....”

\textsuperscript{5}Philadelphia directs a share of SSB tax revenues to pre-kindergarten education initiatives in low income neighborhoods. As a result, Blumgart (2016) writes that “By earmarking the revenues to programs that are at the top of the progressive agenda, Kenney has provided a strong counterargument to those concerned by the regressive nature of the tax: the money will go right back into those very neighborhoods that are hit hardest.”

\textsuperscript{6}Previous work by Saez (2002a) has explored the qualitative question of when a “small” commodity tax can increase welfare in the presence of preference heterogeneity, but left to future work the task of deriving an expression for the optimal tax.
revenue through commodity taxes rather than through income taxes. We show that the sufficient statistic for preference heterogeneity that determines the optimal commodity tax is the difference between the cross-sectional variation of the good’s consumption with income and the amount of that cross-sectional variation that is due to (causal) income effects.

The second challenge involves incorporating behavioral biases into the workhorse model of Atkinson and Stiglitz (1976) and Saez (2002). Most papers in the behavioral public economics literature abstract from redistributive concerns and focus purely on the corrective benefits of an “internality tax.” Such settings can be obtained as a special case of our model. Amongst the small number of papers that do consider redistributive concerns (Gruber and Kőszegi, 2004; Bernheim and Rangel, 2004; Farhi and Gabaix, 2015), all focus on specialized models of bias and assume an overly simple tax environment in the spirit of Ramsey. The specialized Ramsey framework does not allow for redistribution through nonlinear income taxation and thus does not allow, e.g., targeted transfers to the poor or the targeting of sin tax revenue toward progressive policy initiatives.

We fill this gap by providing a formula that characterizes the optimal sin tax in the presence of income taxation via a general money-metric definition of bias, and that clearly elucidates the basic economics of how bias interacts with redistributive motives.

We show that the optimal commodity tax depends on two terms: the “corrective benefits” (representing the welfare gains from reducing harmful internalities), and the “regressivity costs” (representing the welfare costs of shifting net resources from poorer to richer consumers). When a harmful “sin good” is consumed primarily by poorer consumers, redistributive motives amplify both corrective benefits and regressivity costs, and so the total directional effect on the optimal sin tax is ambiguous. Two key factors determine which effect dominates. First, the magnitude of regressivity costs depends on whether those costs are optimally offset by changes to the income tax structure, i.e. progressive revenue recycling. This in turn depends on how much of the sin good consumption profile across incomes is driven by income effects vs. preference (or bias) heterogeneity. If all sin good consumption differences are from income effects, it is optimal to fully offset regressivity costs

---

7 See, e.g., the papers in footnote 2.
8 See, in particular, the special case of no redistributive motive in Section II.C.
9 Gruber and Kőszegi (2004) study the incidence of cigarette taxes on low- and high-income consumers with self-control problems, and show that low-income consumers with sufficiently elastic consumption can benefit from such taxes, but they do not characterize the optimal tax implications. Bernheim and Rangel (2004) study the optimal tax on addictive substances in a dynamic framework in which environmental cues trigger mindless consumption, and differences in the marginal utility from wealth arise from differences in spending on the addictive good or on rehabilitation. Farhi and Gabaix (2015) focus on an application of their general framework that assumes two types of consumers—behavioral and rational—with potential differences in the marginal utility from income. 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 lump-sum and cannot, for example, be spent on transfers to the poor or programs that benefit the poor.
10 Also closely related is the work on health insurance by Baicker et al. (2015), Spinnewijn (forthcoming), and Handel et al. (2016). See also Mullainathan et al. (2012) for a model that links the insights from the health insurance literature to commodity tax results; they derive formulas similar to those in Farhi and Gabaix (2015) for an economy in which consumers make a binary choice.
through income tax changes, so that only the corrective benefits are relevant, and the optimal
sin tax is unambiguously higher than the Pigouvian benchmark (assuming elasticities and biases
are constant across incomes). Conversely, when all sin good consumption differences stem from
preference heterogeneity, the sin tax doesn’t alter the equity vs. efficiency tradeoff of the optimal
income tax, and so sin good regressivity costs should not be offset.

In this case, a second factor determines the relative importance of corrective benefits vs. re-
gressivity costs: the price elasticity of demand. When the elasticity is high, consumers shift their
behavior and avoid the tax’s regressive burden so corrective costs dominate. Conversely, regressivity
costs dominate when the elasticity is low.

In Section III, we demonstrate the usefulness of the theory by estimating sufficient statistics
for the optimal nationwide tax on sugar-sweetened beverages (hereafter, “SSBs”). We use Nielsen
Homescan, a 60,000-household, nationally representative panel dataset of grocery purchases, and
Nielsen’s Retail Measurement Services (RMS), a 35,000-store, national panel of UPC-level sales
that covers about 40 percent of all U.S. grocery purchases. As shown in Figure 1, SSB purchases
decline steeply in income: Homescan households with annual income below $10,000 purchase about
99 liters per adult each year, while households with income above $100,000 purchase only half that
amount. This demonstrates how SSB taxes could be regressive.

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 stores where
the household usually shops. This instrument may be useful for estimating demand in other ap-
lications, because it delivers high power while relaxing stronger assumptions made in prior work
such as Hausman (1996) and Nevo (2001). We estimate the income elasticity of demand using
within-household income variation. We find a small positive income elasticity, which means that
the downward-sloping consumption-income profile illustrated in Figure 1 is driven by strong pref-
erence heterogeneity, not by causal income effects. This strong preference heterogeneity reduces
the socially optimal SSB tax.

To quantify consumer bias, we measured bias proxies—specifically, nutrition knowledge and
self-control—as well as taste and health preferences using a survey of 18,000 Homescan households
that we designed specifically for this paper. Figure 2 presents the unconditional relationships
between Homescan households’ SSB purchases and the two bias proxies. Panel (a) shows that higher
nutrition knowledge is strongly unconditionally associated with lower SSB purchases: households
in the lowest decile of nutrition knowledge purchase more than twice as much SSBs as households
in the highest decile. The vertical line reflects that the 24 dietitians and nutritionists in our survey
sample scored an average of 91 percent correct. Panel (b) shows that higher perceived self-control is
also strongly associated with lower SSB consumption: households whose survey respondents answer
that they “definitely” drink SSBs “more often than I should” purchase almost three times more
SSBs than households whose respondents answer “not at all.”
These striking relationships suggest that lack of nutrition knowledge and self-control could be important contributors to SSB consumption. We formalize this intuition using a framework inspired by Handel and Kolstad (2015) and Bronnenberg, Dubé, Gentzkow, and Shapiro (2015), which we call the “predicted normative consumer” strategy. To implement the approach, we estimate the relationships in Figure 2 after conditioning on preferences and demographics and correcting for measurement error. We then predict “normative” consumption—that is, the consumption each household would have if its survey respondents had the nutrition knowledge of dietitians and nutritionists and no reported self-control problems. This prediction requires the strong assumption that any unobserved preferences are conditionally uncorrelated with bias, and this is the key weakness of this approach. We divide the quantity effect of bias by the demand elasticity to derive the crucial theoretical parameter, the magnitude of bias in units of cents per ounce of SSB consumption.

We estimate that American households consume 0.54 to 0.90 log points more SSBs than they would if they had the nutrition knowledge of nutritionists and dietitians and had no self-control problems. Put differently, we estimate that 42 to 60 percent of American households’ SSB consumption is explained by imperfect nutrition knowledge and self-control. Furthermore, this predicted overconsumption is highly regressive: bias explains 51 to 72 percent of consumption at household incomes below $10,000, while it explains only 37 to 53 percent of the (much smaller) consumption at household incomes above $100,000. This will imply a higher optimal soda tax, because the tax counteracts a bias that is itself regressive.

Finally, in Section IV, we implement our optimal tax formulas using our empirical results. Quantitatively, our results suggest that for a broad range of specifications, the optimal soda tax lies between 1.5 and 2.5 cents per ounce—slightly higher than the levels which have been implemented in most U.S. cities to date.

In addition to its contributions to optimal tax theory and behavioral public economics, our work also connects to a large and growing literature on SSB taxes. One set of papers estimates the price elasticity of SSB demand and the effects of SSB taxes on consumption. Our approach offers transparent estimates from a plausible identifying assumption in a large nationwide sample, whereas most previous papers require more restrictive identifying assumptions or deliver less precise estimates 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 (2014), Wang (2015), and Zhen et al. (2014). These papers estimate consumer surplus without attempting to quantify consumer bias or formally address regressivity.

11Bronnenberg et al. (2015) show that sophisticated shoppers—in their application, doctors and pharmacists—are more likely to buy generic instead of branded drugs, and they conduct welfare analysis assuming that only sophisticates’ choices are welfare-relevant. Analogously, we show that “unbiased” consumers—in our application, people who have as much nutrition knowledge as nutritionists and dietitians and report no self-control problems—purchase less SSBs, and we assume that only choices predicted in the absence of bias are welfare-relevant.

12This includes Bollinger and Sexton (2017), Rojas and Wang 2017, Silver et al. 2017, Tiffin, Kehlbacher, and Salois (2015), Zhen et al. (2011), and others; see Powell et al. (2013), and Thow, Downs, and Jan (2014) for reviews.
concerns, leaving unanswered core questions in the policy debate. Relative to this literature, our paper is novel in that it is the only one to ask the following basic question: what is the optimal soda tax? More broadly, our paper illustrates how optimal commodity taxes can be empirically calibrated in a wide variety of contexts.

II A Theory of Regressive Sin Taxes

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

Individuals make their choices to maximize “decision utility” $U(c, s, z; \theta)$ subject to this budget constraint. $U$ is assumed to be increasing and weakly concave in its first two arguments, and decreasing and strictly concave in the third. The policymaker believes consumers should instead maximize “normative utility” $V(c, s, z; \theta)$, which may differ from from $U$. The policymaker selects a tax $T(\cdot)$ and $t$ to maximize aggregate normative utility, subject to a government budget constraint and to individual optimization. That is, the policymaker’s problem is to maximize aggregate experienced utility, weighted by type-specific Pareto weights $\alpha(\theta)$, less any externalities $E$:

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

subject to the government budget constraint

$$\int_{\Theta} (ts(\theta) + T(z(\theta))) \mu(\theta) \geq R$$

and to individual optimization

$$\{c(\theta), s(\theta), z(\theta)\} = \arg \max_{\{c,s,z\}} U(c, s, z; \theta) \quad \text{s.t.} \quad c + (p + t)s \leq z - T(z) \text{ for all } \theta.$$ 

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, as documented for food choices and energy-efficiency choices by Allcott (2013), Attari et al. (2010), Bollinger et al. (2011), and Schofield (2015). As another example, consumers may have limited attention or salience bias with respect to certain attributes of $s$, as modeled in Gabaix and Laibson (2006), DellaVigna (2009), Gabaix (2014), and Bordalo et al. (2013), and as documented by Allcott and Taubinsky (2015) for energy-efficient appliances. Finally, present bias may lead consumers to underweight the future health costs of some goods.
(e.g., potato chips or cigarettes) as in O’Donoghue and Rabin (2006). Our framework allows us to treat $\beta$ 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 $\kappa$ on the “future-oriented self” and weight $1 - \kappa$ 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 allowing for empirical measurement through an intuitive and operationalizable definition of bias. We thus adopt a general, yet empirically grounded definition of consumer bias in the spirit of the sufficient statistics approach. We do this by constructing a price metric for consumer bias.

Formally, let $s(\theta, y, p, t, T)$ be the sin good consumption chosen by a type $\theta$ individual with after-tax earnings $y = z - T(z)$ at price $p$ and tax $t$. Analogously, define $s^V(\theta, y, p, t, T)$ to be the amount of $s$ that would be chosen if the individual were maximizing $V$ instead. We define the bias, denoted $\gamma(\theta, y, p, t, T)$ as the value for which $s(\theta, y, p, t, T) = s^V(\theta, y - s\gamma, p - \gamma, t, T)$. In words, $\gamma$ is equal to the compensated price change that produces the same effect on demand as the bias does. In terms of primitives, $\gamma = \frac{U'_s}{U'_c} - \frac{V'_s}{V'_c}$.

Throughout, we will use the notation $f'_x$ to denote the derivative of $f(x, y)$ with respect to $x$, and $f''_{xy}$ for the the cross-partial derivative with respect to $x$ and $y$, etc. When no ambiguity arises, we sometimes suppress some arguments and write, for example, $\gamma(\theta)$ for concision. If $\gamma(\theta) > 0$, this means that type $\theta$ consumers “over-consume” relative to if they were maximizing normative utility, whereas $\gamma(\theta) < 0$ means that type-$\theta$ consumers “under-consume.”

The statistic can be quantified by comparing consumers’ choices in “biased” and “debiased” states, as we will do in our empirical application. Other examples that informally employ our definition of bias include Chetty et al. (2009) and Taubinsky and Rees-Jones (forthcoming), who estimate the (average) value of $\gamma$ when bias arises from lack of tax salience by estimating 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) run an experiment that directly estimates $\gamma$ based on consumer’s biased valuation of energy efficient compact fluorescent lightbulbs (CFLs) Note that by being tied directly to choice, the bias measure $\gamma$ is not tied to the policymaker’s redistributive preferences and is invariant to monotonic transformations of the utility functions $U$ and $V$.

To represent redistributive motives concisely, we employ the notion, common in the optimal taxation literature, of “social marginal welfare weights”—the social value (from the policymaker’s perspective) of a marginal unit of consumption for a particular individual, measured in terms of

---

13 See Bernheim and Rangel (2009) and Bernheim (2016) for a discussion of why treating $\beta < 1$ fully as a bias may be problematic.

14 To see this, note that by definition, $U'_c / U'_s = p + t$ and $V'_s / V'_c = p + t - \gamma$ when the arguments $c$, $s$, and $z$ are the same in $U$ and $V$. This also shows that the equivalence only holds under the assumption that the tax $t$ is fully salient; otherwise, $\gamma$ may deviate from $\frac{U'_s}{U'_c} - \frac{V'_s}{V'_c}$.

15 Throughout the paper, we will 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.
public funds. We therefore define \( g(\theta) = \alpha(\theta) V'_c / \lambda \), where \( V'_c \) represents the derivative of \( V \) with respect to its first argument, and \( \lambda \) is the marginal value of public funds (i.e., the multiplier on the government budget constraint at the optimum). These weights are endogenous to the tax system, but are useful for characterizing the necessary conditions which must hold at the optimum. We use \( \bar{g} = \int g(\theta) d\mu(\theta) \) to denote the average marginal social welfare weight—if there are no income effects on consumption and labor supply, then \( \bar{g} = 1 \) by construction.17

II.A Two Simple Examples with Discrete Types

We first use two examples to illustrate an key insight: the optimal sin tax depends crucially on why sin good consumption declines with income. There are two broad possibilities. On the one hand, individuals with different ability levels (and thus different incomes) may have different preferences for the sin good. On the other hand, the sin good may be inferior, so that individuals choose to consume less as their budget increases. In what follows, we describe a simple model with two types of consumers to illustrate the implications of preference heterogeneity vs. income effects for the optimal sin tax.

Our model economy consists of two consumer types, \( \theta \in \{L, H\} \), in equal proportion, with wages \( w(L) < w(H) \). Utility functions:

\[
U(c, s, z; \theta) = G\left(c + v(s; c, \theta) - \psi\left(\frac{z}{w(\theta)}\right)\right)
\]

and

\[
V(c, s, z; \theta) = G\left(c + v(s; c, \theta) - \psi\left(\frac{z}{w(\theta)}\right) - \gamma s\right)
\]

We’ll consider two special cases (though will take the first steps of the derivation using a common approach).

- Pure income effects: \( v(s; c, \theta) = v_{inc}(s; c) \), doesn’t depend directly on type, only on numeraire consumption. Assume \( \frac{\partial^2 v_{inc}}{\partial s \partial c} < 0 \), so \( s(L) > s(H) \), meaning that \( s \) is an inferior good.

---

16 Note that 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, since the policymaker places the same value on marginal spending on either good due to the envelope theorem.

17 Note that 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, redistributive motives 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) normative inequality aversion, as encoded by these weights.

18 These insights, and the resulting tax formulas, extend transparently to settings where consumption rises with income, but for concreteness we discuss a declining consumption profile throughout this section, for consistency with the assumption that sin taxes are regressive.
Pure preference heterogeneity: \( v(s; c, \theta) = v_{\text{pref}}(s; \theta) \), doesn’t depend on numeraire consumption, only on type. Assume \( v'_s(s, L) > v'_s(s, H) \) for all \( s \), so \( s(L) > s(H) \), meaning that high earnings ability is correlated with a higher taste for \( s \).

To characterize the optimal \( t \), we will employ a perturbation argument, wherein we consider a small change to the sin tax \( t \), paired with a corresponding change to the income tax \( T \) which exactly preserves labor supply choices. At the optimum, any such joint reform must have no first-order effect on social welfare, allowing us to write down a first-order condition for the optimal commodity tax.

To understand the rationale for this approach, and the nature of the labor-choice-preserving income tax reform, note that, as is standard in two-type models of optimal nonlinear taxation, the optimum will feature two levels of income, denoted \( z^L \) and \( z^H \), and corresponding income taxes \( T^L \) and \( T^H \) which redistribute as much as possible from \( H \) to \( L \) subject to a constraint that \( H \) willingly select \( z^H \) rather than \( z^L \). In the context of this model, letting \( \hat{s}(t; y, \theta) \) denote type \( \theta \)'s preferred choice of sin good consumption when endowed with net income \( y = z - T(z) \) and facing a sin tax \( t \), this “incentive compatibility” constraint implies19

\[
y^H - (1 + t)\hat{s}(t; y^H, H) + v\left( \hat{s}(t; y^H, H) ; y^H - (1 + t)\hat{s}(t; y^H, H), H \right) - \psi \left( \frac{z^H}{w(H)} \right) \geq 0
\]

\[
y^L - (1 + t)\hat{s}(t; y^L, H) + v\left( \hat{s}(t; y^L, H) ; y^L - (1 + t)\hat{s}(t; y^L, H), H \right) - \psi \left( \frac{z^L}{w(H)} \right) \geq 0
\]

This optimization problem can be solved using a Lagrangian \( L = W + \lambda \cdot RC + \nu \cdot IC \), where \( RC \) denotes the nonnegative resource constraint and \( IC \) denotes the nonnegative incentive compatibility constraint (generated from 2 and 4 by subtracting the right-hand side of each equation from the left-hand side). Under the optimal policy, \( \frac{dC}{dt} = \frac{dC}{dT} = \frac{dC}{dT} = 0 \). The multiplier \( \lambda \) on \( RC \) represents the marginal value of public funds. Since the multiplier \( \nu \) lacks such an intuitive economic interpretation, it is useful to consider a perturbation which preserves 4 with equality (i.e., a perturbation in which the income tax is adjusted to preserve labor supply choices). By doing so, we can express the optimal \( t \) solely in terms of the effects on social welfare, and fiscal externalities (through the marginal value of public funds).

Note that the income tax compensation necessary to preserve equality of 4 depends on whether soda demand differences are driven by preference heterogeneity or income effects. We consider these each case in turn.

First, consider the case where soda consumption varies solely due to income effects, so that \( v(s; c, \theta) = v_{\text{inc}}(s; c) \), and as a result \( \hat{s}(t; y, \theta) = \hat{s}_{\text{inc}}(t; y) \) depends only on \( y \), and not on \( \theta \).

\(^{19}\)Formally, \( \hat{s}(t; y, \theta) \) is implicitly defined by \( (1 + t)U'_c = U'_s \).
Consider now how a sin tax perturbation $dt$ alter the high types’ desire to select $z^L$ instead of $z^H$. An increase $dt$ in the commodity increases the relative tax burden of choosing $z^H$ instead of $z^L$ by $[\hat{s}_{inc}(t; y^L) - \hat{s}_{inc}(t; y^H)]dt$; that is, high types are encouraged to earn less now since their tax expenditures on $s$ will be lower the less they earn. The joint income tax reform which exactly preserves equality of 4 is

$$dT^L = -dt \cdot s(L)$$

and

$$dT^H = -dt \cdot s(H).$$

That is, the income tax is reformed to exactly rebate to each type the additional amount that they pay in sin taxes. Therefore, the total effect on the Lagrangian of this joint reform is equal to

$$\frac{dL}{dt} = \frac{dW}{dt} + \lambda \cdot \frac{dRC}{dt} = \lambda \sum_\theta \left[ -g(\theta)\gamma \left. \frac{\partial s(\theta)}{\partial t} \right|_u + t \cdot \left. \frac{\partial s(\theta)}{\partial t} \right|_u \right],$$

where the notation $\left. \frac{\partial s(\theta)}{\partial t} \right|_u$ denotes the compensated sin good demand response to a change in $t$. Rearranging this expression gives

$$t = \frac{\sum_\theta g(\theta)\gamma \left. \frac{ds(\theta)}{dt} \right|_u}{\sum_\theta \left. \frac{ds(\theta)}{dt} \right|_u}.$$

Second, suppose that soda consumption varies solely due to preference heterogeneity, so that $v(s; c, \theta) = v_{inc}(s; \theta)$, and as a result $\hat{s}(t; y, H)$ does not depend on $y^\theta$. Consequently the tax expenditures on $s$ will not depend on whether the choice of earnings is $z^H$ or $z^L$, and thus an increase in the commodity tax will not affect the high type’s incentive to choose $z^L$ instead of $z^H$. As a result, a sin tax perturbation $dt$ requires no income tax reform to preserve labor supply choice, and so has the following total effect on the Lagrangian:

$$\frac{dL}{dt} = \frac{dW}{dt} + \lambda \cdot \frac{dRC}{dt} = \lambda \sum_\theta \left[ -g(\theta) \left( s(\theta) + \gamma \cdot \left. \frac{\partial s(\theta)}{\partial t} \right|_u \right) + s(\theta) + t \cdot \left. \frac{\partial s(\theta)}{\partial t} \right|_u \right].$$

At the optimum this derivative must equal zero, implying the optimal tax satisfies

$$t = \frac{\sum_\theta g(\theta)\gamma \left. \frac{ds(\theta)}{dt} \right|_u - \sum_\theta \left. \frac{ds(\theta)}{dt} \right|_u \sum_\theta s(\theta)(1 - g(\theta))}{\sum_\theta \left. \frac{ds(\theta)}{dt} \right|_u \sum_\theta \left. \frac{ds(\theta)}{dt} \right|_u}.$$

More generally, this examples illustrate the role of preference heterogeneity vs. income effects when both forces are present. Note that we can decompose the total difference in $s$ consumption between $L$ and $H$ as follows.

$$\Delta s := \hat{s}(t; y^H, H) - \hat{s}(t; y^L, L) = \Delta s_{pref} + \Delta s_{inc},$$

(where $\Delta s$ is endogenous to the tax policy, but we suppress dependence on policy variables for
brevity) with
\[ \Delta s_{\text{pref}} := \hat{s}(t; y^L, H) - \hat{s}(t; y^L, L) \]
and
\[ \Delta s_{\text{inc}} := \hat{s}(t; y^H, H) - \hat{s}(t; y^L, H). \]

In the special case of pure preference heterogeneity, \( \Delta s_{\text{pref}} = \Delta s \) and \( \Delta s_{\text{inc}} = 0 \). Now define \( s_{\text{pref}}(\theta) = \hat{s}(t; y^L, \theta) \) to be the soda consumption which would be consumed by type \( \theta \) if that type had income \( y^L \). Finally, define \( \hat{g}(\theta) \) to be the marginal social value of additional net income \( y \) for a given type, accounting for the effects on constraints and consumption of \( s \) due to income effects. Note that in the presence of both income effects and preference heterogeneity, an income tax reform which preserves equality of 4 is
\[ dT^\theta = -dt \cdot \hat{s}(t; y^\theta, H) \]
—that is, each type’s income tax is reduced in proportion to the amount of \( s \) that would be consumed by type \( H \), if \( H \) were to earn that amount. (Note this follows directly from the incentive compatibility constraint, and is a generalization of the special cases of pure income effects or preference heterogeneity.) Similarly, a perturbation which sets
\[ dT^\theta = -dt \cdot (\hat{s}(t; y^\theta, H) - \hat{s}(t; y^L, H)) \]
also preserves incentive compatibility. Under such a joint perturbation, the welfare transfer to each type (which is not offset by the income tax reform) is
\[ -dt \cdot (s(\theta) - \hat{s}(t; y^\theta, H) + \hat{s}(t; y^L, H)) = -dt \cdot s_{\text{pref}}(\theta). \]

Since this change in net income generates income effects, it is weighted by \( \hat{g}(\theta) \), and therefore the net effect on welfare of such a joint perturbation is
\[
\frac{dL}{dt} = dt \cdot \lambda \sum_\theta \left[ -g(\theta) \gamma \frac{\partial s(\theta)}{\partial t} \bigg|_u + t \cdot \frac{\partial s(\theta)}{\partial t} \bigg|_u + s_{\text{pref}}(\theta) (1 - \hat{g}(\theta)) \right],
\]
and the resulting optimal tax satisfies
\[
t = \frac{\sum_\theta g(\theta) \gamma \frac{ds(\theta)}{dt} \bigg|_u}{\sum_\theta \frac{ds(\theta)}{dt} \bigg|_u} - \frac{\sum_\theta s_{\text{pref}}(\theta) (1 - \hat{g}(\theta))}{\sum_\theta \frac{ds(\theta)}{dt} \bigg|_u}.
\]

Again, note this is a generalization of the two special cases—pure income effects and pure preference heterogeneity—already discussed.

II.B Optimal Sin Tax in a General Setting

The insights from the preceding examples extend to a more general setting with a continuum of types and multidimensional heterogeneity in ability, bias, demand elasticities, and preferences, under some weak assumptions. This section presents that more general setup, and the resulting optimal tax formulas.
II.B.1 Elasticity Concepts, Sufficient Statistics, and Simplifying Assumptions

The optimal sin tax depends on three types of sufficient statistics: elasticities, money-metric bias, and the "progressivity of bias correction." We begin by defining the elasticities. These statistics are understood to be endogenous to the tax regime \((t, T)\), though we suppress those arguments for notational simplicity.

- \(\zeta(\theta)\): the price elasticity of demand for \(s\) from type \(\theta\), formally equal to \(-\left(\frac{ds(\theta)}{dt}\right)\frac{p + t}{s(\theta)}\).
- \(\zeta^c(\theta)\): the compensated price elasticity of demand for \(s\), equal to \(-\left(\frac{ds(\theta)}{dt}\right)\bigg|_u\frac{p + t}{s(\theta)}\).
- \(\eta(\theta)\): the income effect on \(s\) expenditure, equal to \(\zeta - \zeta^c\).
- \(\xi(\theta)\): the causal income elasticity of \(s\), equal to \(\eta(\theta)\frac{z(\theta)}{s(\theta)}(1 - T'(z(\theta)))\).

Additionally, we represent the labor supply response to tax reforms using the following parameters, which are defined formally in Appendix A:

- \(\zeta^c_z(\theta)\): the compensated elasticity of taxable income with respect to the marginal income tax rate.
- \(\eta_z(\theta)\): the income effect on labor supply. These responses are defined to include the full sequence of adjustments due to any nonlinearities in the income tax. See Jacquet and Lehmann (2014) for an extensive discussion of such "circularities."

It will prove convenient to average these statistics across all agents; such averages are denoted using "bar" notation, for example aggregate (that is, 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 as \(\bar{s}(z)\), with behavioral income-conditional responses denoted \(\bar{\zeta}(z) = -\left(\frac{d\bar{s}(z)}{dt}\right)\frac{p + t}{\bar{s}(z)}\), and similarly for other parameters. The income distribution is denoted \(H(z) = \int_{\Theta} 1 \{z(\theta) \leq z\} d\mu(\theta)\), with income density denoted \(h(z)\).

We also define the analogue of \(s_{pref}\) in the continuous type context. As discussed in Section II.A, it is helpful to distinguish between two sources of cross-sectional variation in \(\bar{s}(z)\): income effects, and 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 can be defined in terms of the empirically estimable income elasticity of \(s\), as \(s'_{inc}(z) = E\left[\xi(\theta)\frac{s(\theta)}{z} | z(\theta) = z\right]\), and between-income preference heterogeneity as the residual: \(s'_{pref}(z) = \bar{s}'(z) - s'_{inc}(z)\).\(^{20}\)

\(^{20}\)We emphasize this computation of \(s'_{pref}(z)\) as a residual, since that is the estimation strategy we use in Section III, although in principle one could attempt to quantify \(s'_{pref}(z)\) directly using data on preference heterogeneity, letting \(s'_{inc}(z)\) be a residual.
Using these definitions, we define our key sufficient statistic for preference heterogeneity, “cumulative between-income preference heterogeneity:

\[ \bar{s}_{\text{pref}}(z) := \int_{x = z_{\text{min}}}^{z} \bar{s}'_{\text{pref}}(x) dx. \]

This term quantifies the amount of sin good consumption at income \( z \) (relative to the lowest income level \( z_{\text{min}} \)) that can be attributed to preference heterogeneity.

To aggregate bias across individuals, we will follow Allcott, Mullainathan, and Taubinsky (2014) and Allcott and Taubinsky (2015) in defining the average marginal bias (i.e., money-metric bias weighted by compensated demand response):

\[ \bar{\gamma} = \frac{\int_{\Theta} \gamma(\theta) \left( \frac{ds(\theta)}{dt} \right)_{u} \mu(\theta) \mu(\theta)}{\int_{\Theta} \left( \frac{ds(\theta)}{dt} \right)_{u} \mu(\theta) \mu(\theta)} . \] (5)

Intuitively, this aggregation represents the marginal bias weighted by individuals’ marginal responses to a tax reform which raises the tax \( t \) 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 under-estimate the change in utility from that change in consumption. We additionally define this response-weighted bias conditional on income, denoted \( \bar{\gamma}(z) \).

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:

\[ \sigma := \text{Cov} \left[ g(z), \frac{\bar{\gamma}(z)}{\bar{\gamma}}, \frac{\bar{s}'_{\text{c}}(z)}{\bar{s}} \right] \]

The term \( \sigma \) 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 are concentrated among individuals with high welfare weights, i.e., those with lower incomes. (Covariances and expectations are understood to be taken over the observed income distribution throughout.)

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

\[ \text{Formally,} \]

\[ \bar{\gamma}(z) = \frac{\int_{\Theta} \gamma(\theta) \left( \frac{ds(\theta)}{dt} \right)_{u} 1 \{ z(\theta) = z \} d\mu(\theta)}{\int_{\Theta} \left( \frac{ds(\theta)}{dt} \right)_{u} 1 \{ z(\theta) = z \} d\mu(\theta)}. \]
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, it seems reasonable to assume that the primitive conditions on utilities and social weights have been specified so that Assumption 1 is true at the optimum.” Therefore we sometimes write \( g(z) \) to denote the welfare weight directly as a function of earnings.

2. \( U \) and \( V \) are smooth functions that are strictly concave in \( c, s, \) and \( z, \) and \( \mu \) is differentiable with full support.

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.

4. At each before-tax income \( z \): i) \( s \) is orthogonal to the labor supply income effect \( \eta_z \), labor supply elasticity \( \zeta^c_z \) and the income-elasticity \( \xi \), ii) the income elasticity \( \xi \) is orthogonal to bias \( \gamma \) and labor-supply elasticity \( \zeta^c_z \).

Assumptions 2 and 3 ensure that the income distribution does not exhibit any atoms and individuals’ labor supply and consumption decisions respond smoothly to perturbations of the tax system (see Jacquet and Lehmann (2014) for the role of these assumptions). Assumption 4 simplifies the derivations and resulting formulas, but is not necessary for this approach. (We derive analytic results without this assumption in Appendix D.B.) A particularly simple case in which Assumption 5 holds is when all individuals earning the same income \( z \) have the same preferences for \( s \); however, our assumption is much weaker than this condition.

II.B.2 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 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 affect from mis-spending this marginal income due to bias.\(^{22}\) We then show that under empirically realistic assumptions, we can replace \( \hat{g}(z) \) with \( g(z) \) in the commodity tax formula. In the appendix, we show how these expressions can be expressed entirely in terms of the standard social marginal welfare weights \( g(z) \). Additionally, we define \( e = E/\lambda \) to denote externality costs from \( s \) consumption,

\[ \hat{g}(z) = g(z) - (1 + \eta_z(z)) \frac{g(z)\xi(z)}{\rho + 1} - \eta_z(z) + \eta_z(z) \frac{T'}{1 - \gamma}. \]

\(^{22}\)Formally, \( \hat{g}(z) = g(z) - (1 + \eta_z(z)) \frac{g(z)\xi(z)}{\rho + 1} - \eta_z(z) + \eta_z(z) \frac{T'}{1 - \gamma}. \)
in terms of public funds. Finally, letting $\Sigma^{(z^*)}_{a,b} = Cov_{a|z^*} [a(\theta), b(\theta) | z(\theta) = z^*]$ denote the income-conditional covariance between two variables $a(\theta)$ and $b(\theta)$ at the optimum, we can prove the following proposition:

**Proposition 1.** The commodity tax $t$ satisfies the following at the optimum:

$$
\begin{align*}
t & = \frac{s\gamma^c(\bar{g} + \sigma) + E \left[ \frac{\bar{g}}{\bar{z}} g(z) \Sigma^{(z)}_{s,s} \right]}{s\gamma^c + E \left[ \frac{\bar{g}}{\bar{z}} \Sigma^{(z)}_{s,s} \right]} - \frac{(p + t)Cov[\hat{g}(z), \bar{s}_{pref}(z)]}{s\gamma^c + E \left[ \frac{\bar{g}}{\bar{z}} \Sigma^{(z)}_{s,s} \right]} \tag{6}
\end{align*}
$$

$$
\begin{align*}
t & = \frac{s\gamma^c(\bar{g} + \sigma) + E \left[ \frac{\bar{g}}{\bar{z}} g(z) \Sigma^{(z)}_{s,s} \right]}{s\gamma^c + E \left[ \frac{\bar{g}}{\bar{z}} \Sigma^{(z)}_{s,s} \right]} - pCov[\hat{g}(z), \bar{s}_{pref}(z)] \tag{7}
\end{align*}
$$

When the sin good is a small small share of expenditures, or when consumption of the sin good is a homogeneous conditional on income, the optimal tax $t$ satisfies

$$
\begin{align*}
t & \approx \frac{s\gamma^c(\bar{g} + \sigma)}{s\gamma^c} - \frac{(p + t)Cov[\hat{g}(z), \bar{s}_{pref}(z)]}{s\gamma^c} \tag{8}
\end{align*}
$$

$$
\begin{align*}
t & = \frac{s\gamma^c(\bar{g} + \sigma) - pCov[\hat{g}(z), \bar{s}_{pref}(z)]}{s\gamma^c + Cov[\hat{g}(z), \bar{s}_{pref}(z)]} \tag{9}
\end{align*}
$$

The optimal income tax satisfies

$$
\begin{align*}
T'(z) = \frac{1}{1 - T'(z)} & = E \left[ \frac{g(z)\gamma(\theta) + e - t}{p + t} \eta(\theta) | z(\theta) = z \right] + \frac{1}{\zeta^zh(z)} \int_{z}^{\infty} (1 - \hat{g}(x))dH(x)
\end{align*}
$$

Proposition 1, proved in Appendix C.A, shows that the optimal tax is the combination of two main terms, which appear in equations (8) and (9). The term $\gamma(\bar{g} + \sigma) + e$ corresponds to the corrective benefits of the tax. The corrective benefits are increasing in (1) the average marginal bias $\gamma$, (2) the average social welfare weight $\bar{g}$, and (3) the extent to which bias correction is concentrated on the low-income consumers—quantified by $\sigma$. This term illustrates a key difference between externalities and internalities. In the case of consumer bias, it is necessary to know whether the bias is bigger for the rich or for the poor, and whether the rich or the poor are more elastic, since the costs from a given individual’s consumption fall back on that individual, and thus are scaled by their marginal social welfare weight. The externality, on the other hand, is born by the population in general, and thus receives the same weight regardless of whose consumption generates it.

The term $Cov[\hat{g}(z), \bar{s}_{pref}(z)]$ scales the redistributive motive of the tax. Importantly, this term depends on the extent to which the variation of sin good consumption by income is due to preference heterogeneity rather than income effects.

The additional terms proportional to $\hat{g}(z)\Sigma^{(z)}_{s,s}$ in equations (6,7) result from the fact that changes
in commodity taxes also generate income effects, which will be heterogeneous when consumption of the sin good is heterogeneous (conditional on income). This leads to heterogeneous changes in behavior and thus heterogeneous consequences for both the corrective benefits and fiscal externalities. However, these effects are negligible when the sin good constitutes a small share of expenditures.

We consider the simple case of negligible income effects in equation (8), which is written in a form that echoes the result in Proposition ?? with two types. The equation illustrates that the optimal tax can be decomposed into marginal corrective benefits—given by $\gamma (g + \sigma) + e$—less marginal regressivity costs—given by $\frac{p + t}{\zeta} Cov[\hat{g}(z), s_{pref}(z)]$. In (9) we rearrange so that $t$ does not appear on the right side. This formulation provides guidance about the relative influence of the corrective benefits versus the regressivity costs in the optimal tax formula. The higher is the elasticity, the higher are the corrective benefits per unit change in the tax. Correspondingly, the relative importance of corrective benefits, as well as the sensitivity of the optimal tax $t$ to bias $\gamma$ depends on the price elasticity of demand.

A complication in the formulas in Proposition 1 that lead to expressions involving $\hat{g}$ instead of $g$ is the potential relevance of labor supply income effects $\eta_z$. However, Gruber and Saez (2002) find small and insignificant income effects and Saez et al. (2012) 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.” This gives rise to an important special case. If there are no labor supply income effects, and if consumers’ expenditures on the sin good are sufficiently small to have a negligible impact on the marginal utility from numeraire consumption, then the weights $\hat{g}$ can be replaced by $g$ in Proposition 1. This claim is formalized by the following corollary; the proof is in the appendix.

**Corollary 1.** Assume that $\eta_z \approx 0$. And assume that expenditures on $s$ are a small share of expenditures. Then

$$t \approx \frac{s\zeta c (\gamma (1 + \sigma) + e) - pCov[g(z), s_{pref}(z)]}{s\zeta c + Cov[g(z), s_{pref}(z)]}$$

The proof follows immediately from Proposition 1, noting that $\eta$ is negligible when $s$ is a small portion of expenditures, and noting from footnote 22 that when $\eta \approx 0$ and $\eta_z \approx 0$, then $g(z) \approx \hat{g}(z)$. We use this formula as our primary sufficient statistics approximation of the optimal tax (in the presence of the optimal income tax) in Section IV.

**II.B.3**

**Corollary 2.**
II.B.4 General Expression for the Optimal Sin Tax at a Fixed Income Tax

In the United States, many soda taxes are being set by cities that do not control the income tax structure, and in general it may not be possible to re-optimize the income tax system at the same time that a sin tax is established. What is the optimal sin tax in that case? The causal income elasticity always plays an important role, even when the income tax is not necessarily optimal. As before, we define \( g(z) \) to be the marginal utility from the numeraire divided by the marginal value of public funds.

**Proposition 2.** Assuming that \( \eta \) and \( \eta_z \) are negligible, and that \( \zeta_z, \xi_{inc}, \) and \( s \) are uncorrelated conditional on income, the optimal commodity tax satisfies:

\[
t = \bar{\gamma} (\bar{g} + \sigma) + e - \frac{(p + t) [(\bar{g} - 1)\bar{s} + \text{Cov}[g(z), s(z)]]}{\bar{s} \bar{\Sigma}^c} - \frac{p + t}{\bar{s} \bar{\Sigma}^c} \underbrace{\mathbb{E} \left[ \frac{T'(z)}{1 - T'(z)} \zeta_s(z) \bar{\Sigma}(z) \right]}_{\text{correction}}
\]

The first term is the corrective benefit. The second term is the direct welfare effect of the commodity tax. The third term is the effect of the commodity tax on labor supply decisions. The intuition is that when the sin good is a normal good, a choice of higher earnings results in a higher total tax paid for consumption of the sin good; thus a higher tax on the sin good is equivalent to an increase in the marginal income tax rate. The converse obtains when the sin good is an inferior good. When the income tax is set optimally, the second and third terms combine to produce the formula in Proposition 1.

II.B.5 Equivalent Variation and Welfare Calculations

We can use the sufficient statistics discussed above to approximate the change in welfare from the introduction of a sin tax. Intuitively, the net change in welfare is equal to the equivalent variation (measured under normative utility) of the sin tax, weighted by each individual’s welfare weight, plus the increase in tax revenues and externality reductions, measured in terms of public funds. We consider each in turn.

Equivalent variation (EV) for a consumer of type \( \theta \) is defined as the change in wealth which would alter normative utility by the same amount as the introduction of a given tax \( t \). For a small tax, this is equal to decision utility EV (the wealth change such that the consumer would be indifferent to it or the introduction of the tax) plus the change in welfare due to internalities. For a small tax \( t \), decision utility EV is approximated by \( -t \cdot s(\theta) + \zeta_c(\theta) \cdot \frac{s(\theta)}{p} \cdot \frac{t^2}{2} \), and the welfare change due to internalities is equal to \( \gamma(\theta) \cdot \zeta_c(\theta) \cdot \frac{s(\theta)}{p} \cdot t \). Integrating across individuals, the total change in surplus among individuals earning \( z \) is equal to \( dEV(z) \approx -t \cdot \bar{s}(z) + \zeta_c(z) \cdot \frac{\bar{s}(z)}{p} \cdot t \left( \frac{t}{2} + \bar{\gamma}(z) \right) \).

If demand elasticities remain approximately constant under large price changes, then a better approximation is to substitute \( \Delta \bar{s}(z) \approx \bar{s}(z) \left( \exp \left[ \ln \left( \frac{p}{p + t} \right) \zeta_c(z) \right] - 1 \right) \), giving an approximate
change in consumer surplus of

\[ dEV(z) \approx -t \cdot \bar{s}(z) + \bar{s}(z) \left( 1 - \exp \left[ \ln \left( \frac{p}{p+t} \right) \tilde{\zeta}^c(z) \right] \right) \left( \frac{t}{2} + \tilde{\gamma}(z) \right). \]

By weighting the resulting EV by social marginal welfare weights (net of income effects, i.e., \( \hat{g}(z) \)), these welfare changes can be aggregated with the value resulting tax revenues and reduced externalities (both measured in units of public funds) to approximate the total change in welfare from the tax, accounting for distributional concerns.

The tax raises revenue approximated by

\[ t \cdot \bar{s} \cdot \exp \left[ \ln \left( \frac{p}{p+t} \right) \tilde{\zeta}^c \right] - t \cdot \mathbb{E} \left[ \frac{T'(z)}{1-T'(z)} \tilde{\zeta}_z \bar{s}(z) \tilde{\zeta}_{inc}(z) \right], \]

and further raises welfare through the reduction of externalities by approximately

\[ \bar{s} \cdot \left( 1 - \exp \left[ \ln \left( \frac{p}{p+t} \right) \tilde{\zeta}^c(z) \right] \right) \cdot e \] (measured in public funds). Combining these effects yields an approximation of the total change in social welfare from the introduction of the tax \( t \).

II.C Interpretation and Implications

Proposition 1 illustrates a number of important insights about the forces which affect the optimal sin tax. We highlight five key implications and special cases.

1. Special case: no redistributive concerns. If social marginal welfare weights are constant (implying the planner has no desire to distributive marginal resources from high to low income consumers) then \( t^* = \bar{\gamma} \). This special case could arise 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 individuals have the same marginal utility of consumption). In both cases, the optimal commodity must exactly offset the average marginal bias. The first case is a slight generalization of known sin tax results (e.g., O’Donoghue and Rabin 2006; Mullainathan et al. 2012; Allcott et al. 2014; Allcott and Taubinsky 2015) to the case of income effects, while the second case is a slight generalization of known sin tax results to situations in which the policymaker also has a nonlinear income tax at his disposal.

2. Special case: no corrective concerns. Another special case of our framework is the case in which there are no corrective concerns. Proposition 1 implies that for this case,

\[ \frac{t}{p+t} = -\frac{\text{Cov}[\hat{g}(z), \bar{s}_{pref}(z)]}{\bar{s} \tilde{\zeta}^c}, \]

and Proposition 2 can similarly be used to obtain a formula in terms of \( g(z) \) rather than \( \hat{g}(z) \).

Equation (10) generalizes the seminal Atkinson-Stiglitz theorem to the case of arbitrary preference heterogeneity. The Atkinson-Stiglitz theorem itself obtains as a special case of (10) when all variation in \( c_2 \) consumption is driven by income effects, which then implies that \( t = 0 \). Additionally, formula (10) generalizes the qualitative analyses about the usefulness of commodity taxes by Saez.
(2002b) and Edwards et al. (1994), the results for two-type models by Edwards et al. (1994) and Nava et al. (1996), and results for homogeneous preferences (but with consumption-labor complementarities) by ?.

More generally, this formula holds when \( \bar{\gamma} = 0 \), which is the case when the consumers who over-purchase (or under-purchase) the good are inelastic to the tax (or subsidy). A stark 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.23 Although the Bernheim and Rangel (2004) result that the sin good should be subsidized is seemingly in stark contrast to the sin tax results in, e.g., O’Donoghue and Rabin (2006), our general tax formula clarifies the economic forces that lead each result.

3. Special case: no preference heterogeneity. When differences in consumption are due purely to differences in income (i.e., the ratio \( \frac{U_s(c,s,\theta,z)}{U_c(c,s,\theta,z)} \) depends only on \( \theta \) but not on \( z ) \), then regressive consequences of a sin tax can be perfectly offset by modifications to the income tax, so that there are no fiscal externalities resulting from the full policy reform. In this case the optimal sin tax is

\[
t = \bar{\gamma} \left( \bar{\gamma} + \sigma \right)
\]

This formula generalizes the result in Proposition ???. Note that redistributive concerns still play a role in the size of the optimal commodity tax, reflected by the \( \sigma \) term. The more elastic the low-income consumers are, and the more biased they are relative to the high income consumers, the higher is \( \sigma \) and thus the higher is the benefit of bias correction.

4. A higher sin tax demand elasticity raises corrective benefits relative to regressivity costs. Proposition 1 shows that as that elasticity grows large, the optimal tax \( t \) approaches \( \bar{\gamma} \left( \bar{\gamma} + \sigma \right) \), which is the formula for the optimal tax when there are no regressivity costs due to heterogeneous preferences. If \( \sigma > 0 \) (as is the case if bias and elasticity are constant across incomes) then the size of the optimal tax will exceed the optimal Pigovian tax that prevails absent redistributive concerns. At the opposite extreme, as the elasticity grows small, corrective benefits become negligible compared to the regressivity costs, and the optimal tax approaches the formula in equation (10). If preference heterogeneity accounts for any share of the decrease in \( s \) consumption across incomes, then a sufficiently low elasticity implies the optimal tax becomes has to be negative (a subsidy) whenever it is regressive. Intuitively, if individuals do not respond to

\[23\] Because in Bernheim and Rangel (2004) over-consumption of the good is a consequence of cue-triggered neural processes that render the individual inelastic to prices, the average bias of consumers who are elastic to the tax is zero. Thus, their subsidy result is a direct consequence of the positive covariance between consumption of the addictive good and the marginal utility from income—analogous to our “regressivity costs” term.

19
commodity taxes, then such taxes become a powerful instrument to enact redistribution through targeted subsidies.

More generally, the results clarify that the role of consumer bias in shaping both the sign and magnitude of the optimal commodity tax is modulated by the price elasticity of demand. But perhaps most importantly, the elasticity also provides practical guidance on how sensitive the optimal tax is to different values of the bias \( \bar{\gamma} \). A lower elasticity dampens the responsiveness of the tax to the bias \( \bar{\gamma} \), since

As the two special cases we have examined so far show, whether the cross-sectional variation of \( c_2 \) is explained by preference heterogeneity or income effects plays a crucial role in whether concerns about regressivity are reflected in commodity taxes. In practice, both preference heterogeneity and income effects likely play a role in explaining the variation in consumption of the sin good. Importantly, however, if the sin good is a normal good (meaning that consumption rises with net income, conditional on type) our results imply that the regressivity would be even higher than those arising under pure preference heterogeneity, since the commodity both falls more heavily on the poor than the rich and decreases labor supply.\(^{24}\)

II.D Extensions

Our baseline results can be extended in a number of ways, which we do in Appendix B. In the appendix, we characterize results when the sin good is a composite good consisting of several different items, such as soft drinks of various sizes. Second, we derive a general condition for the optimal sin tax when there are more dimensions of consumption, which allows for substitution across goods with various degrees of marginal bias. In our companion paper, Allcott, Lockwood, and Taubinsky (2018), we also study the case in which individuals are may under-react to the implicit labor supply incentives generated by changes in the commodity tax on \( s \).

III Estimating Key Parameters for the Optimal Soda Tax

III.A Overview

In this section, we estimate the empirical parameters needed to calibrate the optimal tax on sugar-sweetened beverages. The theory generates for empirical questions. First, what is the price elasticity of SSB consumption \( \tilde{\zeta}_c \), and how does this vary by income? Second, how much of the cross-sectional relationship between SSB consumption and income documented in Figure 1 is due to causal income effects \( s'_{inc}(z) \) vs. between-income preference heterogeneity \( s'_{pref}(z) \)? Third, how large is bias \( \bar{\gamma} \), and how does this vary by income? Fourth, how large is the externality \( e \)? We begin this section by presenting the data. After that, we address each of the above empirical questions in turn.

\(^{24}\)This is because the term \( \phi(z) \), corresponding to how much variation is explained by income effects, is actually negative in this case.
III.B Data

III.B.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 2006-2015. RMS includes 53, 32, 55, 2, and 1 percent 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 city where each store is located (through newspaper or online ads and coupons) and whether the UPC was temporarily on “display” inside each store.

To measure household grocery purchases, we use the Nielsen Homescan Panel for 2006-2015. Homescan includes about 38,000 households in 2006, and about 61,000 households each year for 2007-2015.

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). Panel (a) of Table 1 presents descriptive statistics for Homescan households at the household-by-year level. If there are two household heads, we use the two heads’ mean age, 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 re-weight the sample for national representativeness.

Nielsen groups UPCs into product modules. We define sugar-sweetened beverages (SSBs) as the product modules that have typically been included in existing SSB taxes: “fruit drinks” (which

---

25 RMS reports two separate retail chains: the parent code, which is the source of the scanner data, and the retailer code, which is intended to reflect the retail banner. We are not allowed to disclose which retailers are in the RMS dataset. As a hypothetical example, Kroger has several store banners, including Kroger, City Market, Dillons, Harris Teeter, and others. If Kroger were in RMS and collected all scanner data from all of their banners before submission to Nielsen, all stores from all of those banners would appear in RMS under the same parent code but with different retailer codes. We define a retail chain by the RMS parent code, with the exception that we separate out individual retailer codes that have more than 100 store-by-year observations in the RMS stores data. There are 90 parent codes in the data, and after separating out these specific individual retailers, we define 106 retail chains.

26 The raw data include feature and display information for the same original set of audited stores, but it is safe to assume that if a UPC is featured at a given store, it is also featured at all of that retailer’s other stores in the same Designated Market Area (DMA). We thus impute the feature variable to all other stores in a retailer-DMA cell. After this imputation, we observe the feature and display variables, respectively, in approximately 8.6 and 81 percent of store-by-week observations.

27 See Einav et al. (2010) for a validation study of Homescan.

28 For analyses using the full sample, we use the “projection factors” provided by Nielsen. For analyses using the subsample of PanelViews survey respondents, we construct our own nationally representative sample weights using the first seven variables presented in Panel (a) of Table 1.
includes sports drinks and energy drinks), packaged liquid coffee and liquid tea (for example, bottled iced coffee and iced tea), carbonated soft drinks, and non-carbonated soft drinks (which includes cocktail mixes, breakfast drinks, ice pops, powdered drinks, etc.). Fruit and vegetable juice and artificially sweetened drinks such as diet soda are not included. The bottom two rows of Panel (a) of Table 1 show that the average Homescan household purchases 153 liters of SSBs per year, at an average price of $1.12 per liter.\textsuperscript{29} (Average price paid is undefined for the 3.3 percent of household-by-year observations with no SSB purchases.) We deflate all prices and incomes to real 2015 dollars.

Homescan households are asked to record the UPCs of all consumer packaged goods they purchase. While Homescan gives us a large and high-quality sample, there are three issues to keep in mind. First, Homescan does not include data on beverages purchased and consumed away from home, such as at restaurants. Second, purchases may differ from intake if people sometimes give soda to others or throw it out instead of drinking it themselves. Third, Homescan data are at the household level, not the individual level, which causes some mismatch between the consumption data and the individual-level bias proxies elicited in the PanelViews survey described below. To address these issues, we also carried out a beverage intake questionnaire as part of the PanelViews survey. To reduce the length of the paper, we analyze these self-reports only in Appendix H. Notwithstanding, the self-reports are an important complement to the Homescan data for our main policy analysis.

### III.B.2 Homescan PanelViews Survey

We measure bias proxies and preference controls using a survey of Homescan households that we designed for this project. Nielsen implemented the survey in October 2017 using their PanelViews survey platform. Panel (b) of Table 1 summarizes the respondent-level data. Appendix E gives the exact text of the survey questions described below.

The survey was fielded to all adult heads of the approximately 60,000 households that were in the 2015 or 2016 Homescan data. Because the 2016 Homescan data are not yet available, we temporarily drop several thousand responses from households that did not join the Homescan panel until 2016. We have complete and valid responses from 20,839 people at 18,159 households.

We attempt to 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).\textsuperscript{30} The GNKQ is widely used in the public health literature; see Kliemann

\textsuperscript{29}We measure quantity interchangeably in liters and ounces, under the standard approximation that all drinks have the density of water. For liquid drinks, the weight is reported by Nielsen directly from the package label. For powdered drinks, we transform to the weight when consumed as liquid, i.e., with water added.

\textsuperscript{30}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,”
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, 1/3, 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, pre-packaged tea or coffee, 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.

Finally, we asked gender, occupation, and whether the respondent is the “primary shopper.” The sample is so large that it happens to include 24 nutritionists and dietitians; we use this group’s average nutrition knowledge below.

III.B.3 County Mean Income

We use county-by-year mean per capita personal income from the Regional Economic Information System (BEA 2017).

III.C Price and Income Elasticities

III.C.1 Empirical Strategy Overview

In this section, we estimate the price and income elasticities of demand. Let $\ln s_{it}$ denote the natural log of Homescan SSB purchases (in liters per adult equivalent) by household $i$ in quarter $t$. Let $\ln p_{it}$ denote the natural log of average price paid per liter of SSB purchased, and let $\boldsymbol{f}_{it}$ denote a vector of feature and display variables, which we detail below. $z_{ct}$ is the mean per capita income in county $c$ in the calendar year that contains quarter $t$, $\omega_t$ is a vector of quarter of sample indicators, and $\mu_{ic}$ is a household-by-county fixed effect. We estimate the following regression:

and “Not sure.” A second example question 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).
\[
\ln s_{it} = -\zeta \ln p_{it} + \nu f_{it} + \xi \ln z_{ct} + \omega t + \mu_{ic} + \varepsilon_{it},
\]
(12)

with standard errors clustered by county.

Because SSBs are storable, previous purchases could affect current stockpiles and thus current purchases, and Hendel and Nevo (2006b) 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 to not include lags in Equation (12). See Appendix F for details.

The estimated price elasticity of demand is \( \hat{\zeta} \). In estimating price elasticity, it is important to control for merchandising conditions \( f_{it} \), because stores often reduce prices in concert with advertising and in-store displays. It is also important to instrument for price; we address this below. As discussed above, the Homescan data do not measure SSBs that are purchased and consumed away from home, such as at restaurants. If households substitute between scanned and non-scanned SSBs, for example by buying less soda at a restaurant when grocery store prices are lower, the estimated price elasticity of scanned purchases \( \hat{\zeta} \) will likely overstate the true SSB price elasticity \( \zeta \). We do not believe that this is a substantial confound, but we will explore a more inelastic alternative assumption in Section IV.

The estimated causal income elasticity is \( \hat{\xi} \). We think of this as the reduced form of an instrumental variable (IV) regression with time-varying household income as an endogenous variable and county mean income as an instrument, where the first stage should have a coefficient of one.\(^{31}\) In an alternative specification that we discuss below, we directly use Homescan panelists’ reported household income to identify \( \hat{\xi} \). \( \hat{\xi} \) represents the elasticity with respect to year-to-year income variation over the sample. One might also hypothesize that elasticity with respect to longer-run income variation could be quite different. While it is not possible to cleanly estimate such a long-run income elasticity, we will explore alternative assumptions in Section IV.\(^{32}\)

### III.C.2 Price Instrument Using Chain-Level Time-Varying Prices

A key challenge in demand estimation is addressing simultaneity bias: omitted variables bias generated by potential correlation between price and unobserved demand shifters. We address simultane-

---

\(^{31}\)This approach would be biased if changes in a county’s skill composition (i.e. migration of workers with different earning ability than the previous county average) are correlated with within-household SSB demand changes or income changes for continuing residents. However, these compositional effects are likely only a small share of year-to-year county income variation.

\(^{32}\)As we show in Appendix Table A7, employment status and weekly hours worked are not statistically significantly associated with SSB consumption when included in Equation (A7), so we cannot reject separability of SSB consumption and labor. Thus, it does not matter whether variation in \( z_{ct} \) results from non-labor 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. In practice, much of the within-county income variation results from differential effects of the Great Recession; see Appendix Figure A4 for illustrative maps.
ity bias using an instrument inspired by DellaVigna and Gentzkow (2017) and Hitsch, Hortacsu, and Lin (2017). Figure 3 presents the intuition, using data for an example UPC, the best-selling SSB UPC in the RMS data. Panel (a) presents the quarterly average price (unweighted across stores and weeks) for each retail chain that sold this UPC over the full sample, with darker shading indicating higher prices. Other than a nationwide price increase in early 2008, there is no clear pattern, illustrating that chains vary their prices independently of each other over time.

Panel (b), by contrast, shows a clear pattern. This figure looks within one example retail chain, presenting the quarterly average price (again unweighted across weeks) for each store at that chain. The vertical patterns illustrate that this chain varies prices in a coordinated way across all of its stores. In mid-2008, for example, prices at all stores were relatively high, whereas in early 2007, prices at all stores were relatively low. The figure echoes a similar figure in DellaVigna and Gentzkow (2017), who document that this within-chain, across-store price coordination is not limited to the example chain we chose for this figure.

This variation allows us to construct a powerful leave-out instrument for the price of each UPC sold at each store in the RMS data. Define $\ln p_{jkw}$ as the natural log of 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 and weeks. Then let $\ln p_{krt, -c}$, the leave-out price, denote the unweighted average of price deviation $\ln p_{jkw} - \ln p_{kw}$ at all of retail chain $r$’s stores outside of county $c$, across all weeks in quarter $t$. \cite{33}

Panel (a) of Figure 3 shows that retailer-specific price deviations $\ln p_{krt, -c}$ will have substantial variation, because different retailers charge different prices in a given quarter. This national differencing helps to remove responses to national-level demand shocks that might influence the price of the specific UPC $k$, which could still be a concern after we condition on time fixed effects $\omega_t$ that soak up shocks to overall SSB demand. Panel (b) of Figure 3 suggests that the leave-out price is highly predictive of the actual price experienced by customers of retailer $r$ at stores inside of county $c$. In principle, this local leave-out could be important because it guarantees that our instrument is not contaminated by store-specific responses to local demand shocks. In practice, we shall see that the local leave-out does not make much difference, because price variation is so uniform within chains.

We then fit these leave-out prices $\ln p_{krt, -c}$ to each individual household’s average purchasing patterns. Define $s_{ijkc}$ as household $i$’s 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.

\cite{33}In constructing $\ln p_{krt, -c}$, we include only the 81 percent of store-week observations where the feature variable is observed. This is important because the majority of features are associated with a price discount, and Appendix Table A7 shows that omitting the feature variable does change the estimated price elasticity. Therefore, we could introduce bias by including weekly price observations where feature is unobserved. By contrast, in-store displays are less frequently associated with a price decrease, and Appendix Table A7 shows that omitting display does not generate significant omitted variables bias. Thus, to increase power in the instrument, we construct $\ln p_{krt, -c}$ with weekly price observations regardless of whether display is observed.
Then $\pi_{ijkc} = s_{ijkc}/s_{ic}$ 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 predicted local price deviation in quarter $t$ is thus

$$Z_{it} = \sum_{k,j \in \text{RMS}} \pi_{ijkc} \ln p_{krt,c}. \tag{13}$$

Price deviation $\ln p_{krt,c}$ is only observed at RMS stores, so $Z_{it}$ sums only over purchases at RMS stores. The instrument is still powerful, because 34 percent of SSB purchases are at RMS chains. Because $\pi_{ijkc}$ is the purchase share across all SSB purchases (at both RMS and non-RMS stores), each household’s actual prices paid should move approximately one-for-one with $Z_{it}$.

The feature and display variables are constructed analogously. Let the 2-vector $f_{jkt}$ denote the number of weeks in which UPC $k$ was observed to be featured at 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. The feature and display variables we use in the household-by-quarter regressions are these UPC-by-store-by-quarter merchandising conditions weighted by the household’s purchase shares $\pi_{ijkc}$:

$$f_{it} = \sum_{k,j \in \text{RMS}} \pi_{ijkc} f_{jkt}. \tag{14}$$

Figure 4 presents binned scatterplots that illustrate the local price instrument $Z_{it}$. Panel (a) presents the contemporaneous first stage—that is, the relationship between $Z_{it}$ and $\ln p_{it}$—after conditioning on the other controls in Equation (12). We expect a slope close to one, and estimate is in fact 1.15. Panel (b) presents the reduced form—that is, the relationship between $Z_{it}$ and the natural log of liters of SSBs purchased—again after conditioning on the other controls. The slope is -1.66. Dividing the reduced form by the first stage gives a price elasticity of approximately $-1.65/1.15 \approx 1.44$.

The exclusion restriction is that the local price instrument $Z_{it}$ is uncorrelated with demand shifters $\varepsilon_{it}$, conditional on the set of controls in Equation (12): feature and display, county income, household-by-county fixed effects, and time controls. The key 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 natural threat to this assumption would be price cuts co-ordinated with retailer-specific advertising, but retailers do not do much advertising beyond the newspaper and online ads and coupons that are already captured by RMS feature variable (DellaVigna and Gentzkow, 2017). Furthermore, we show below that the estimates are largely unaffected by alternative instruments and fixed effect strategies that are designed to soak up other types of regional and city-specific demand shocks.

This instrument offers advantages over some instruments traditionally used in this type of demand estimation. Hausman (1996) and Nevo (2001), for example, construct instruments for
the price of product $k$ in period $t$ and market $c$ using the average price of product $k$ in all other markets in that time period. Those instruments require the assumption that time-varying demand shocks for product $k$ be uncorrelated across markets, which assumes away possibilities such as national advertising campaigns. Our instrument also delivers more power, because chains introduce substantial price variation that is independent of national prices, as illustrated in Panel (a) of Figure 3.

III.C.3 Estimation Results

Table 2 presents estimates of Equation (12). Column 1 presents OLS estimates, without the local price instrument, while column 2 presents the main IV estimates using $Z_{it}$ as defined above.

The estimated price elasticity $\hat{\zeta}$ is $-1$ times the price coefficient reported in the first row. The OLS estimate of $\zeta$ is less than half of the IV estimate. This is consistent with the usual effects of simultaneity bias: sellers respond to positive demand shocks by raising prices. This could also be driven by compositional effects in the types of SSBs that consumers purchase—for example, if people purchase more expensive SSBs in the same quarters when they purchase more SSBs. The IV estimates could also differ from OLS because the Local Average Treatment Effects are local to consumers that shop in RMS stores. Given that RMS covers such a large share of grocery purchases, however, it is unlikely that this could explain the entire difference.\footnote{Furthermore, the demographics of compliers predict that IV estimates would reflect less elasticity, not more. Higher-income households purchase a larger share of their beverages from RMS stores, and as we discuss below, these households are somewhat less price elastic.}

Column 3 presents estimates with an alternative IV construction that directly uses prices in the store where the consumer shops instead of “leave-out” prices from that retailer’s stores in other counties. Because price variation is so uniform within RMS chains, this makes little difference relative to column 2.

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 4 addresses these concerns by adding city-by-quarter fixed effects. Such 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 5 addresses this by interacting the city-by-quarter fixed effects with indicators for above- versus below-median household income. In both columns, the point estimates move slightly but are statistically indistinguishable.

While 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 \( Z_i \) that is correlated with that shock, even conditional on city-by-time fixed effects. Column 6 addresses this concern by using an instrument constructed with deviations from the census region average log price instead of the national 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.\(^{35}\)

The second row of coefficients in Table 2 are \( \hat{\xi} \), the estimated income elasticity of demand. Column 2’s primary estimate of \( \hat{\xi} \approx 0.25 \) changes little across the IV specifications. Appendix G presents alternative specifications substituting Homescan panelists’ self-reported income instead of county mean income \( z_{ct} \). The \( \hat{\xi} \) estimates are statistically zero or positive, and all are much smaller than those using county income, consistent with attenuation bias driven by measurement error in within-household income variation. Regardless of the income variable used, the estimates are consistent in their implication that at the steeply declining SSB consumption-income profile is driven not by causal income effects, but by the fact that lower-income households have stronger preferences for SSBs.\(^{36}\)

In addition to knowing the sample average price elasticity, we ideally would be able to estimate any differences in price elasticity between low- and high-income households. If low-income households are more price elastic, then an SSB tax generates more consumption reduction among the poor, and thus more social welfare gains if soda is overconsumed. This is reflected in the bias correction progressivity term \( \sigma \) from Section II. To capture this, we estimate a model in which we let \( \zeta \) and \( \nu \) vary linearly with the natural log of household \( i \)'s sample mean income. The quantitative estimates are in Appendix Table A8. The income-price interactions are not statistically significantly different from zero, although the price elasticity point estimates differ somewhat across incomes: fitted price elasticities are 1.50 and 1.41 at incomes of $5000 and $125,000, respectively.

III.D Causal Income Effects vs. Between-Income Preference Heterogeneity

The second key empirical question motivated by the theory is, how much of the cross-sectional relationship between SSB consumption and income documented in Figure 1 is due to causal income

\(^{35}\)The estimates include only observations with positive SSB consumption, as price paid \( p_{it} \) is undefined for the 15 percent of quarterly observations with no SSB purchases. In theory, this can bias our estimates, as high prices are more likely to cause zero-purchase observations. Appendix Table A6 addresses this by presenting Tobit estimates (thereby formally accounting for latent demand that is censored at zero) 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.

\(^{36}\)One additional concern with the identification of income elasticity \( \xi \) is that county income changes could be correlated with unobserved within-household preference changes, for example if high-SES households live in counties that grow faster during the sample and also experience faster declines in preferences for SSBs. Appendix Table A7 shows that the estimated \( \xi \) is unaffected by controlling for a linear time trend in county sample average income or household educational attainment.
effects $s'_{inc}(z)$ vs. between-income preference heterogeneity $s'_{pref}(z)$? The dark circles on Figure 5 repeat Figure 1, plotting average annual SSB purchases in liters per adult equivalent for households in nine different income bins. Although the y-axis scale is now compressed, recall that SSB purchases are sharply declining with income: households with income under $10,000 per year purchase 99 liters per adult equivalent per year, while households with income over $100,000 purchase only 50 liters.

The curve at the top of the figure uses the income elasticity $\hat{\xi} \approx 0.25$ from Table 2 to predict the causal effects of income increases on the lowest-income households' SSB consumption: $\bar{s}(z_{min}) + \hat{\xi} \ln z$. In Section II, we defined between-income preference heterogeneity $\bar{s}_{pref}(z)$ as the difference between actual consumption and consumption predicted by income effects. On the graph, $\bar{s}_{pref}(z)$ is thus the vertical difference between the dark circles and the curve, which is drawn in x’s.

The estimate of $\bar{s}_{pref}(z)$ illustrates very stark between-income preference heterogeneity. If all households were randomly re-assigned the same income, households currently making over $100,000 per year would consume 167 liters less SSBs than households currently making under $10,000 per year. This difference is about 2.5 times average consumption. When incorporated into the optimal tax simulations in Section IV, this means that SSB consumption is a highly predictive tag of low earning ability, so a policymaker will want to subsidize SSBs (or reduce SSB taxes) as a way of redistributing money to lower-income households.

III.E Measuring Bias

III.E.1 The Predicted Normative Consumer Estimation Strategy

Any empirical strategy that identifies behavioral bias in units of dollars could be used to calibrate the $\gamma$ term defined in Section II. In this paper, we use an approach inspired by Bronnenberg et al. (2015) and Handel and Kolstad (2015), which we call the “predicted normative consumer” estimation strategy. The process is to measure proxies for behavioral bias using surveys, estimate the relationship between bias and quantity consumed, predict the quantity change if consumers were unbiased (i.e. maximized normative utility), and transform that quantity into dollar units using the price elasticity.

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

$$\ln s_i = \ln s^V_i + \gamma_i \xi \zeta / p_i,$$

where $\ln s^V_i$ denotes the log of the amount that household $i$ would consume in the absence of bias, $p_i$ varies across households to match the data, and compensated elasticity $\xi \zeta$ is constructed using...
The Slutsky equation from uncompensated elasticity and income elasticity estimated earlier.\textsuperscript{37} One intuition for this equation is that bias as a proportion of quantity (i.e. $\ln s_i - \ln s_i^V$) equals bias as a proportion of price (i.e. $\gamma_i / p_i$) normalized by the elasticity of demand $\zeta_i^c$. We refer to $\gamma_i \zeta_i^c / p_i$ and $\ln s_i - \ln s_i^V$ as “the quantity effect of bias.”

To estimate bias $\gamma_i$, let $b_i$ denote the vector of household $i$’s biases, and let $b^V$ denote the value of $b$ for a “normative” consumer that maximizes $V$. Let $\bar{b}_i$ denote a vector of bias proxies measured in the PanelViews survey: nutrition knowledge $\hat{b}_{ki}$ and self-control $\hat{b}_{si}$. Let $a_i$ denote the vector of beverage tastes and health preferences measured in the PanelViews survey, let $x_i$ denote the vector of household characteristics introduced in Table 1, and let $\mu_c$ denote county fixed effects. In the 2,680 households with two valid PanelViews survey responses, we use the two respondents’ mean nutrition knowledge, beverage tastes, and health preferences. In those households, self-control $\bar{b}_{si}$ is the average of the primary shopper’s self-control ratings of herself and the other household head, and $\bar{b}_{2si}$ is an analogous variable using the ratings reported by the second respondent.

We now make four assumptions.

1. Unconfoundedness: $\gamma_i \zeta_i^c / p_i \perp (\ln s_i^V | a_i, x_i, \mu_c)$. In words, the quantity effect of bias is independent of normative consumption, after conditioning on county and measures of preferences and demographics. In the absence of additional survey measures designed to measure preferences, this assumption would be unreasonable. However, for this assumption to fail in our case, there must be some confounding preference that is not captured by our tailor-made survey measures of beverage tastes and health preferences.

2. Normative consumers: $b_k^V = E[b_k | Nutritionist, dietitian]$, $b_k^V = 1$. For nutrition knowledge, we set $b_k^V$ equal to the average nutrition knowledge score of the 24 nutritionists and dietitians in the PanelViews survey, which is 0.91. For self-control, we set $b_c = 1$: normative consumers are those that answer “not at all” in response to the statement, “I drink soda pop or other sugar-sweetened beverages more often than I should.”

3. Linearity and completeness: $\gamma_i \zeta_i^c / p_i = \tau (b^V - b_i)$, with $b_i := [b_{ki}, b_{si}]$ and $b^V := [b_k^V, b_c^V]$. In words, the quantity effect of bias is linear in nutrition knowledge and self-control. Linearity is not crucial to the approach and could be relaxed, for example by using higher order polynomials and interactions of bias proxies or by determining $s_i^V$ by matching to another household $j$ with the same covariates ($a_i, x_i, \mu_c$) as $i$ but with $b_j = b^V$.\textsuperscript{38} What is crucial is

\textsuperscript{37}To derive Equation (15), note that $\gamma$ by definition also satisfies $\ln s^V(\theta, y, p) = \ln s(\theta, y + s\gamma, p + \gamma)$. Thus, to a first-order approximation, $\ln s^V(\theta, y, p) \approx \ln s(\theta, y, p) + \frac{d \ln s(\theta, y, p)}{d \ln p} \cdot (\ln(p + \gamma(\theta)) - \ln p) \approx \ln s(\theta, y, p - \zeta_s^c(\theta) \gamma(\theta)/p.

Substituting the $i$ subscript for the $(\theta, y, p)$ triple and re-arranging gives $s_i = \ln s_i^V + \gamma_i \zeta_i^c / p_i$. Compensated elasticity is $\zeta_i^c = \zeta_i - \frac{\zeta_s}{1 - T^c(\zeta_i)} \frac{p_i a_i}{s_i}$, because $\zeta^c = \zeta - \frac{dz}{dp} p$ per the Slutsky equation, and $\frac{dz}{dp} p = \frac{dz}{dy} y p = \frac{dz}{dx} x p = \frac{dz}{dx} \frac{1}{1 - T^c(\zeta)} p = \frac{-\zeta}{1 - T^c(\zeta)} \frac{p a_i}{s_i}$. We use the $\zeta_i$ estimated at household $i$’s income from Appendix Table A8 and the $T^c(\zeta_i)$ implied by the U.S. income tax schedule, as described in Appendix J.B.

\textsuperscript{38}In our data, linearity is a realistic assumption: Figure 2 demonstrated this with consumption in levels, ad
that the bias measures are conceptually “complete”—that is, that the policymaker considers no biases other than nutrition knowledge and self-control.

4. Classical measurement error, construct validity, and content validity: 
\[ \tilde{b}_{ki} = b_{ki}, \tilde{b}_{si} = b_{si} + \nu_{si}, \]
and \[ \tilde{b}_{2si} = b_{si} + \nu_{2si}, \]
with \( \nu_{si}, \nu_{2si} \perp b_{si} \) and \( \nu_{si}, \nu_{2si} \perp b_{si} \). The General Nutrition Knowledge Questionnaire from which we derive nutrition knowledge \( \tilde{b}_{ki} \) is a many-question scale that is documented to have a high test-retest reliability of 0.89, as well as high sensitivity to change and construct validity (Kliemann et al., 2016). We thus assume that \( b_{ki} \) is measured without error. Our one-question self-control measure \( \tilde{b}_{si} \), however, could more plausibly suffer from measurement error. This is why we also gathered each household head’s self-control ratings for the other head, and we assume that the errors in these repeated observations are independent. By construct and content validity, we mean that this assumption embeds the requirement that our survey measures are valid and complete measures of the nutrition knowledge and self-control problems that might motivate soda taxes.

We can estimate \( \tau \) under assumptions 1 and 3 by regressing SSB consumption on bias proxies and controls:

\[
\ln(s_i + 1) = \tau \tilde{b}_i + \beta_0 a_i + \beta_x x_i + \mu_c + \varepsilon_i. \tag{16}
\]

For each household, we use its one most recent year in the data, which for about 92 percent of households is 2015.\(^{39}\) We add 1 to SSB consumption before taking the natural log so as to include observations with zero consumption. \( \hat{\tau} \) measures the extent to which bias is related to soda consumption. If \( \hat{\tau} = 0 \), we will infer that there is no bias affecting consumption that would need to be corrected by a tax.

If there were no measurement error, we could directly use the \( \hat{\tau} \) from Equation (16). However, measurement error in self-control could generate attenuation bias in \( \hat{\tau} \). If we had two spouses reporting self-control for all households, we could exploit the repeated observations in a standard instrumental variables procedure, instrumenting for \( b_{si} \) with \( b_{2si} \). Since we only have two spouses in a subset of the households, we instead implement an equivalent approach using a mechanical two-sample two-stage least squares procedure. In the first stage, we regress \( b_{2si} \) on \( b_{si} \) and the other covariates from Equation (16) in the subset of households with two survey respondents:

\[ \text{Appendix Figure A5 demonstrates this with consumption in natural logs. The linearity assumption can be derived as an approximation to any “structural” behavioral model in which } \tilde{b}_i \text{ scales the share of costs that are misperceived. For example, consider a model in which utility of soda consumption is } v(s) - ps - \beta h s, \text{ where } h \text{ is the future health cost per unit of soda consumption and } \beta \text{ is the self-control parameter in the standard } \beta, \delta \text{ model (Laibson, 1997). If the policymaker defines normative utility using the long-run criterion, then } \gamma_i = (1 - \beta_i) h. \text{ Thus, if } \beta_i = b_i / b^V, \text{ i.e. } b_i \text{ is proportional to self-control and } b^V \text{ represents the value for time-consistent consumers, then linearity holds if we set } \tau = \frac{B}{B_{\tau}}. \]

\(^{39}\)As we use more historical consumption data, the correlations with bias proxies measured in the 2017 survey become weaker, suggesting that biases may change over time.
\[
\tilde{b}_{2si} = \alpha \tilde{b}_i + \beta_a a_i + \beta_x x_i + \mu_c + \epsilon_i.
\] (17)

Under assumption 4, the fitted values \(\tilde{b}_{2si}\) equal the conditional expectation of true bias \(b_{si}\): \(\tilde{b}_{2si} = \mathbb{E}[b_{si} | (\tilde{b}_i, a_i, x_i, \mu_c)]\). We can construct these fitted values for the full sample where we have \(\tilde{b}_{si}\), even if we only have one PanelViews respondent with which to estimate the first-stage. We then use \(\hat{b}_{2si}\) instead of \(\tilde{b}_{si}\) when estimating Equation (16), with standard errors adjusted for uncertainty in the first stage estimates following Murphy and Topel (1986).

We now have all the ingredients to calculate average marginal bias. Re-arranging assumption 3, the empirical estimate of \(\gamma_i\) is

\[
\hat{\gamma}_i = \frac{\tau (b^V - \hat{b}_i) p_i}{\hat{\zeta}_i},
\] (18)

where \(\hat{b}_i = [\hat{b}_{ki}, \hat{b}_{si}]\), the bias proxies constructed with the fitted values of \(i\)'s self-control.

What are the strengths and weaknesses of this predicted normative consumer approach relative to alternative approaches to measuring money-metric bias? To fix ideas, one alternative approach to estimating \(\gamma\) caused by imperfect information and self-control problems would be to combine an information provision field experiment similar to Allcott and Taubinsky (2015) with a preference reversal field experiment similar to Sadoff, Samek, and Sprenger (2015), who offer grocery shoppers the chance to order groceries in advance and then measure whether shoppers systematically re-optimize (for example, toward less healthful items) when the groceries are delivered. The unconfoundedness assumption is the key weakness of our approach, although we take various steps to address this issue.

Notwithstanding, we chose this approach for our application because it has several relative strengths. First, measuring \(\gamma\) with experimental information provision requires the assumption that the information provided fully debiases the information treatment group, which could be unrealistic in the case of complex information around nutrition and health. Our nutrition knowledge scores capture information gathered over a very long period, which would be very hard to induce experimentally. Second, by asking people if they think they drink SSBs “more often that I should,” our self-control survey question measures what people want for themselves, so it is natural to use this to construct a normative benchmark without imposing the judgment of a third-party policymaker. By contrast, doing welfare analysis with present biased consumers requires an assumption about whether the policymaker should respect long-run or short-run preferences (Bernheim and Rangel, 2009). Third, one needs a price elasticity to transform bias estimates in quantity units into money-metric \(\gamma\) for policy analysis. In practice, estimating price elasticity in an experimental sample can require strong assumptions around whether experimental subjects have opportunities to buy SSBs outside of the experiment and can store SSBs for future consumption. Fourth, any experimental approach delivers estimates of specific biases that are inevitably specific to a small experimental
By contrast, the predicted normative consumer approach allows the researcher to estimate multiple biases in a nationwide sample.

### III.E.2 Descriptive Facts

Figure 2, which we discussed in the introduction, shows that there is a strong unconditional relationship between bias proxies (nutrition knowledge and self-control) and SSB purchases. After conditioning on other controls, this is the variation that identifies $\tau$ in Equation (16).

Figure 6 shows that both nutrition knowledge and self-control have striking relationships with income. Households with income above $100,000 score 0.12 higher (0.85 standard deviations) than households with income below $10,000 on the nutrition knowledge questionnaire. Households with income above $100,000 also score about 0.11 higher (0.36 standard deviations) on self-control. In addition to being remarkable descriptive facts, these relationships matter for the optimal SSB tax because they imply larger bias correction benefits for low-income people, who have higher social welfare weights.

Figure 7 shows that preferences entering normative utility also differ systematically by income. Panel (a) shows that relative to households with income above $100,000, households with income below $10,000 are about 0.10 higher (0.28 standard deviations) in terms of how much they “like the taste and generally enjoy drinking” regular soft drinks. Panel (b) shows that relative to those highest-income households, the lowest-income households average about 0.07 points lower (0.38 standard deviations) in their reported importance of staying healthy. Both results imply that lower-income households have stronger normative preferences for SSBs. This corroborates the result illustrated in Figure 5 that the declining consumption-income relationship is driven by preference heterogeneity, not income effects.

### III.E.3 Regression Results

Table 3 presents estimates of Equation (16) using household-level Homescan purchase data. Column 1-4 are direct estimates of Equation (16), with various configurations of independent variables but without the measurement error correction. Column 1 presents the unconditional relationship between household $i$’s SSB purchases and bias proxies. The unconfoundedness assumption would be violated if unobserved preferences are correlated with bias. To provide a sense of the implications of observed preferences, we progressively add controls across columns 2-4. Column 2 adds the preference controls $a_i$. Since preferences are correlated with both consumption and bias, we expect that their inclusion will attenuate the $\hat{\tau}$ coefficients on bias proxies. Column 3 adds household income and education. The coefficient on self-control is stable. Not surprisingly, however, income and education are positively correlated with both nutrition knowledge and with consumption, and their inclusion further attenuates the nutrition knowledge coefficient. By the time we add the
remaining demographic controls (age, race, household size, etc.) and county indicators in column 4, including additional controls does not substantially change the estimated $\hat{\tau}$. While this is certainly not dispositive, it is consistent with the idea that additional unobserved factors would also not change the estimated $\hat{\tau}$.  

Column 5 incorporates the measurement error correction, substituting the fitted values of self-control $\tilde{b}_{si}$ in place of the direct survey measure $\tilde{b}_{si}$. Appendix Table A9 presents the “first stage” estimates used to generate these fitted values. Conditional on the other covariates, one household head’s self-control ratings predict the other head’s ratings with $\alpha$ coefficients of about 0.58, implying moderate measurement error. Furthermore, nutrition knowledge is positively correlated with self-control. Thus, the two-stage least squares measurement error correction procedure effectively scales down $\tilde{b}_{si}$ by 0.58 and allows the self-control variable to take credit for a component that would otherwise load onto nutrition knowledge. Relative to the uncorrected estimates in column 4, the $\hat{\tau}$ estimates in column 5 are therefore smaller for nutrition knowledge and substantially larger for self-control.

### III.E.4 Estimates of Bias

Plugging the estimates of $\tau$ from Table 3 into Equation (15) under the linearity assumption, we can estimate the quantity effects of bias $\ln s_i - \ln s_i^V$. On average, American households consume about 0.54 log points more SSBs than would be predicted if they had the nutrition knowledge of nutritionists and dietitians and no self-reported over-consumption problems. Put differently, about 42 percent of current U.S. SSB consumption is attributable to bias. Figure 8 shows that this predicted overconsumption is much larger for low-income households.

Finally, we can predict money-metric bias $\hat{\gamma}_i$ for each household using Equation (18) and the demand-slope weighted average marginal bias using the empirical analogue of Equation (5): $\hat{\gamma} = \frac{\sum_i \zeta_i \frac{\hat{\gamma}_i}{p_i}}{\sum_i \zeta_i}$. The average marginal bias across all American households is 1.13 cents per ounce. Figure 9 plots $\hat{\gamma}(z)$ by income. Since the quantity effect of bias declines in income and elasticities and prices do not differ much by income, money-metric bias will also decline in income. Average marginal biases are 1.38 and 1.01 cents per ounce, respectively, for households with income below $10,000 per year and above $100,000 per year, respectively.

Appendix H presents parallel estimates using PanelViews self-reported SSB intake instead of

---

40Our key unconfoundedness assumption is that after controlling for health preferences and taste for soda and other SSBs, there are no other unobserved components of preferences that might be correlated with nutrition knowledge and self-control. One way to relax this assumption could be to formally bound $\tau$ using coefficient movement, as in Oster (2016). Because the $R^2$ is not near one and because, as expected, controlling for preferences, income, and education moves the $\hat{\tau}$ coefficients, the bounds on the true $\tau$ using Oster’s (2016) approach would be wide under the assumption that unobservables are equally strongly correlated with $b$ as are observables. Because we have carefully measured preferences, however, we think that this assumption would be far too generous. Furthermore, there is only an $R^2 \approx 0.50$ relationship between current-year and two-year lagged Homescan SSB consumption, suggesting that idiosyncratic taste shocks substantially reduce the maximum achievable $R^2$.
Homescan purchases. The $\hat{\tau}$ coefficients are larger in PanelViews than Homescan, perhaps because measuring consumption and bias proxies for the same person at the same time reduces measurement error. As a result, the national average quantity effects of bias and $\hat{\gamma}$ are also larger in PanelViews than in Homescan: 0.90 log points and 2.65 cents per ounce, respectively. Using the PanelViews data, we estimate that 60 percent of current U.S. SSB consumption is attributable to bias.

III.F Externalities

We import a basic 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. The U.S. Department of Health and Human Services (Yong et al., 2011) estimates that on average, about 15 percent of health costs are borne by the individual, while 85 percent are covered by insurance. Thus, we use a health system externality of $e = 0.85$ cents per ounce.

IV Computing the Optimal Sugar-Sweetened Beverage Tax

We now 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 U.S., using Proposition 2. Second, we compute the optimal SSB tax assuming the income tax is also reformed to be optimal, using Corollary 1.

These computations require an assumption about redistributive preferences. We employ social marginal welfare weights proportional to $y_{US}^\nu$, where $y_{US}$ is post-tax income in the U.S., and $\nu$ is a parameter which 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. Calibrations of the status quo U.S. income distribution and income tax are drawn from Piketty, Saez, and Zucman (2018); see Appendix J.B for details. Since the sufficient statistics formulas used for these calculations are appropriate when income effects are small, here we do not distinguish between $g$ and $\hat{g}$.

The sufficient statistics formulas for the optimal tax depend on a number of population statistics, as well as key covariances involving welfare weights, for which we need estimates of sufficient statistics by income across the income distributions. The necessary statistics, based on our empirical estimates described in Section I, are reported in Table 4. Panel (a) presents estimates of population-level statistics, while Panel (b) presents estimates of statistics within each Homescan income bin, from which we can compute the welfare weight covariances.
With these statistics in hand, we are in a position to compute the optimal SSB tax. We calculate \( \sigma = tk^* \approx 0.41 \text{or whatever} \). The corrective motive in both sufficient statistics formulas is equal to \( \bar{\gamma}(1 + \sigma) + e \approx 0.97(1+0.41\text{or whatever})+0.85 \approx 2.23 \) cents per ounce. This highlights the importance of accounting for internalities when setting optimal corrective taxes. Less than half of the corrective motive (about 40%) is driven by externality correction. Indeed, the average marginal bias \( \bar{\gamma} \) is larger than the externality, and the internality correction is further inflated by about 30% due to bias correction progressivity, reflecting the fact that the benefits of bias correction accrue primarily to poorer consumers with high welfare weights.

Counteracting this corrective motive, the redistributive motive pushes toward a smaller optimal SSB tax, since consumption is concentrated among poorer consumers. In the presence of the optimal income tax, the redistributive motive is proportional to \(-Cov[g(z), s_{pref}(z)]\). This statistic can be computed directly from Table 4, 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 and observed consumption among the lowest earners (see table notes for details).\(^{41}\) In this case, the estimated covariance is \( Cov[g(z), s_{pref}(z)] = 19.85 \). Therefore, using Corollary 1 our sufficient statistics estimate for the optimal income SSB tax (in the presence of the optimal income tax) is \( t = 0.88 \) cents per ounce.

When the income tax is constrained to the status quo, the redistributive motive is instead proportional to \(-Cov[g(z), s(z)] - E\left[\frac{T'(z)}{1-T'(z)}\bar{\xi}_z \bar{s}(z) \bar{\xi}_{inc}(z)\right]\). The first term represents the mechanical effect of the tax based on actual SSB consumption, while the second represents the change in income tax revenues due to the effect of the SSB tax on labor supply. When the income tax is optimal, these two terms sum to exactly \(-Cov[g(z), s_{pref}(z)]\). If instead the status quo income tax is “too low” (i.e., not sufficiently redistributive—as is the case here) relative to the optimum, then the revenue costs of SSB taxation through the income tax are reduced, and as a result the optimal SSB tax is higher under the prevailing income tax than under the optimal one. Using Table 4, we can calculate \( Cov[g(z), s(z)] = 6.07 \) and \( E\left[\frac{T'(z)}{1-T'(z)}\bar{\xi}_z \bar{s}(z) \bar{\xi}_{inc}(z)\right] = 0.34 \).

These estimates of the optimal SSB tax are reported in Table 5, along with analogous calculations under several alternative assumptions. The Pigouvian specification reports the optimal tax in the absence of any redistributive motive, in which case the tax is simply equal to \( \bar{\gamma} + e \). Weak redistributive preferences lead to a higher tax than in the baseline, since the reduced corrective motive (due to smaller bias correction progressivity) is more than offset by a weaker redistributive motive. Correspondingly, stronger redistributive preferences lead to a lower tax. In both cases, the deviation from baseline is much greater under the optimal income tax than under the existing tax, for the same reason that the redistributive motive is stronger under the optimal tax in the baseline.

The optimal tax computed using self-reported SSB consumption is meaningfully higher than under our baseline calculation, driven primarily by higher bias estimates under that specification—see

\(^{41}\)Covariances are computed as \( Cov[a, b] = \sum_n f_n(a_n b_n) - \sum_n f_n a_n \sum_n f_n b_n, \) where \( n \) indexes rows of Table 4.
Appendix H. Although self-reported consumption may be subject to errors of recall or misrepresentation, we view this specification as an important alternative policy benchmark, since self-reports may represent a more accurate view of individual (rather than household) consumption, and may also better reflect outside-the-home consumption, e.g., from soda machines or at restaurants. As a result, our preferred estimate of the optimal tax lies in the range between our baseline estimate and the specification using self-reported consumption.

Finally, we report two specifications with alternate assumptions about the corrective motive. The first assumes internalities are zero, while the second assumes both internalities and externalities are zero. Note that the latter case generates a substantial subsidy at the optimum. These estimates emphasize the quantitative importance of accounting for both types of correction when setting policy.

The estimates in Table 5 are an approximation of the optimal tax, for two reasons. First, the formulas in Corollary 1 and Proposition 2 assume that some behavioral responses are quantitatively negligible. (Specifically, income effects on labor supply and the budget share of SSBs are assumed to be small, and the elasticity of labor supply, the SSB income elasticity, and SSB consumption are assumed to be mutually independent conditional on income.) Second, sufficient statistics in these formulas are computed in the status quo equilibrium, rather than under the optimal tax. To explore the importance of each potential source of error, Appendix J presents computations of the optimal tax using a structural model, with taxes computed 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. Appendix J also explores a fuller range of robustness checks and alternative specifications, which do not alter the main message of this section.

We can further use the sufficient statistics formulas from Section II.B.5 to approximate the equivalent variation and the resulting change in welfare from the optimal tax. Table 6 reports the equivalent variation from the tax within each income bin, as well as the resulting increase in tax revenues, and the welfare gain from reduced externalities. The final row of the table reports the total welfare change, equal to the sum of the equivalent variation in each bin (weighted by population share and social marginal welfare weight) and the revenue and externality effects. We compute these welfare changes holding fixed the status quo U.S. income tax, for three specifications from Table 5: our baseline optimal SSB tax, the optimum assuming no internalities, and the optimal tax assuming no internalities or externalities. Units are dollars per capita, per year.

Table 6 shows that equivalent variation generally increases in income, because tk see comment. Interestingly, however, although the lowest income group in the first row consumes the most SSBs, they do not suffer the largest reduction in consumer surplus when the optimal tax is imposed, because as we saw in Figures 6, 8, and 9, those households have a substantially higher bias than the $15,000 income group. In total, this computation indicates that the optimal tax raises welfare by approximately $9.49 per capita, or (aggregating across adults in the US, and ignoring consumption of children) about $2.4 billion annually. Repeating this calculation with the bias estimates and
optimal tax based on the PanelViews self-reports, we estimate that the optimal tax raises welfare by $tk per capita, or $tk in total, each year.

V Conclusion

This paper provides a tractable theoretical and empirical framework for optimal sin tax analysis that takes into account the following inputs: First, the extent to which people over- (or under-) consume the good because of externalities or internalities. Second, the extent to which the direct incidence of the sin tax is regressive; i.e., falls more heavily on the poor. Third, the extent to which over-consumption due to internalities is itself regressive. Fourth, the extent to which it is optimal to devote revenues from the sin tax to progressive policy initiatives or transfers to lower-income individuals to at least partially offset the regressivity of the tax. Prior work in behavioral economics has largely abstracted from issues of redistribution, and thus has only considered the first of the issues. Prior work in public economics has provided qualitative insights about the second and fourth issues, but has not produced implementable sufficient statistics formulas. We provide general sufficient statistics formula that addresses all four issues, and deploy a flexible empirical methodology to implement the formula in the domain of sugar-sweetened beverages. We find that at current income taxes, the optimal sugar-sweetened beverage tax is 1.5 to 2.5 cents per ounce.

In deriving our formula, we make contributions both to optimal tax theory and to behavioral public economics. In implementing our formula, we also make contributions to several empirical literatures. We provide the first (to our knowledge) attempt to quantify a sufficient statistic for preference heterogeneity. We also develop a new instrument for measuring demand elasticities that improves upon prior approaches such as those of Dube (2005), Hausman (1996), and Nevo (2001). We use this instrument to provide one of the first estimates of the demand elasticity of SSBs using plausibly exogenous variation, a question around which a large public health literature has developed. Finally, we provide new survey design and econometric techniques for obtaining unbiased measures of the associations between field behaviors and experimentally measured behavioral biases.

Beyond SSB taxes, our portable methodology could immediately be applied to study questions about taxes (or subsidies) on cigarettes, alcohol, various unhealthy foods, nutrients such as sugar, or consumer products such as energy-efficient appliances. 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.

Of course, as we discuss throughout the paper, our approach has its weaknesses, which naturally leads us to urge caution in taking an overly strong stance on a particular tax estimate—and to 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.
References


Table 1: **Descriptive Statistics**

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household income ($000s)</td>
<td>653,554</td>
<td>68.00</td>
<td>47.59</td>
<td>3</td>
</tr>
<tr>
<td>Years education</td>
<td>653,554</td>
<td>13.92</td>
<td>2.06</td>
<td>6</td>
</tr>
<tr>
<td>Age</td>
<td>653,554</td>
<td>52.24</td>
<td>14.41</td>
<td>18</td>
</tr>
<tr>
<td>1(White)</td>
<td>653,554</td>
<td>0.77</td>
<td>0.42</td>
<td>0</td>
</tr>
<tr>
<td>1(Black)</td>
<td>653,554</td>
<td>0.12</td>
<td>0.33</td>
<td>0</td>
</tr>
<tr>
<td>1(Have children)</td>
<td>653,554</td>
<td>0.33</td>
<td>0.47</td>
<td>0</td>
</tr>
<tr>
<td>Household size (adult equivalents)</td>
<td>653,554</td>
<td>2.48</td>
<td>1.36</td>
<td>1</td>
</tr>
<tr>
<td>1(Employed)</td>
<td>653,554</td>
<td>0.61</td>
<td>0.44</td>
<td>0</td>
</tr>
<tr>
<td>Weekly work hours</td>
<td>653,554</td>
<td>22.84</td>
<td>16.73</td>
<td>0</td>
</tr>
<tr>
<td>SSBs purchased (liters)</td>
<td>653,554</td>
<td>153.14</td>
<td>190.70</td>
<td>0</td>
</tr>
<tr>
<td>Average price ($/liter)</td>
<td>632,073</td>
<td>1.12</td>
<td>1.38</td>
<td>0</td>
</tr>
</tbody>
</table>

(a) **Homescan Household-by-Year Data**

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutrition knowledge</td>
<td>20,839</td>
<td>0.70</td>
<td>0.15</td>
<td>0</td>
</tr>
<tr>
<td>Self-control</td>
<td>20,839</td>
<td>0.77</td>
<td>0.34</td>
<td>0</td>
</tr>
<tr>
<td>Other head self-control</td>
<td>13,241</td>
<td>0.67</td>
<td>0.38</td>
<td>0</td>
</tr>
<tr>
<td>Taste for juice drinks</td>
<td>20,839</td>
<td>0.49</td>
<td>0.32</td>
<td>0</td>
</tr>
<tr>
<td>Taste for soda</td>
<td>20,839</td>
<td>0.52</td>
<td>0.36</td>
<td>0</td>
</tr>
<tr>
<td>Taste for tea/coffee</td>
<td>20,839</td>
<td>0.45</td>
<td>0.36</td>
<td>0</td>
</tr>
<tr>
<td>Taste for sports drinks</td>
<td>20,839</td>
<td>0.29</td>
<td>0.32</td>
<td>0</td>
</tr>
<tr>
<td>Taste for energy drinks</td>
<td>20,839</td>
<td>0.17</td>
<td>0.28</td>
<td>0</td>
</tr>
<tr>
<td>Taste for fruit juice</td>
<td>20,839</td>
<td>0.72</td>
<td>0.29</td>
<td>0</td>
</tr>
<tr>
<td>Taste for diet drinks</td>
<td>20,839</td>
<td>0.32</td>
<td>0.37</td>
<td>0</td>
</tr>
<tr>
<td>Health importance</td>
<td>20,839</td>
<td>0.84</td>
<td>0.18</td>
<td>0</td>
</tr>
<tr>
<td>SSB consumption (liters)</td>
<td>20,839</td>
<td>87.66</td>
<td>146.24</td>
<td>0</td>
</tr>
<tr>
<td>1(Male)</td>
<td>20,839</td>
<td>0.28</td>
<td>0.45</td>
<td>0</td>
</tr>
<tr>
<td>1(Primary shopper)</td>
<td>20,839</td>
<td>0.88</td>
<td>0.33</td>
<td>0</td>
</tr>
</tbody>
</table>

(b) **PanelViews Respondent-Level Data**

Notes: Panel (a) presents descriptive statistics on the Nielsen Homescan data, which are at the household-by-year level for 2006-2015. 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 zero 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 consumption of household heads. Prices and incomes are in real 2015 dollars. Panel (b) presents descriptive statistics on the Homescan PanelViews data, with one observation for each respondent. See Appendix E for the text of each question. Observations are weighted for national representativeness.
Table 2: Instrumental Variables Estimates of Price and Income Elasticities

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Average price/liter)</td>
<td>-0.744***</td>
<td>-1.438***</td>
<td>-1.466***</td>
<td>-1.565***</td>
<td>-1.581***</td>
<td>-1.395***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.096)</td>
<td>(0.089)</td>
<td>(0.106)</td>
<td>(0.107)</td>
<td>(0.100)</td>
</tr>
<tr>
<td>ln(County income)</td>
<td>0.159**</td>
<td>0.245***</td>
<td>0.249***</td>
<td>0.219**</td>
<td>0.183**</td>
<td>0.240***</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.071)</td>
<td>(0.071)</td>
<td>(0.086)</td>
<td>(0.086)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Feature</td>
<td>1.335***</td>
<td>1.074***</td>
<td>1.063***</td>
<td>1.056***</td>
<td>1.048***</td>
<td>1.090***</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.076)</td>
<td>(0.075)</td>
<td>(0.080)</td>
<td>(0.080)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>Display</td>
<td>0.563***</td>
<td>0.525***</td>
<td>0.523***</td>
<td>0.544***</td>
<td>0.537***</td>
<td>0.527***</td>
</tr>
<tr>
<td></td>
<td>(0.092)</td>
<td>(0.099)</td>
<td>(0.099)</td>
<td>(0.098)</td>
<td>(0.098)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>City-quarter fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>City-quarter-income fixed effects</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Kleibergen-Paap first stage F stat</td>
<td>180.5</td>
<td>257.0</td>
<td>177.3</td>
<td>178.0</td>
<td>205.7</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>1,998,407</td>
<td>1,998,451</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (12). All regressions include quarter of sample indicators and household-by-county fixed effects. Column 1 presents OLS estimates. Columns 2, 4, and 5 present instrumental variables estimates using the primary IV. Column 3 presents estimates using an IV based on the prices charged in the stores where each household shops, instead of the average prices charged at the same retail chain outside of the household’s county of residence. In columns 4 and 5, “city” is Nielsen’s Designated Market Area (DMA). Column 6 constructs the instrument using deviations from regional average prices instead of national average prices. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.
Table 3: *Regressions of Sugar-Sweetened Beverage Consumption on Bias Proxies*

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutrition knowledge</td>
<td>-1.295</td>
<td>-1.153</td>
<td>-0.879</td>
<td>-0.817</td>
<td>-0.608</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.073)</td>
<td>(0.075)</td>
<td>(0.081)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>Self-control</td>
<td>-1.237</td>
<td>-0.966</td>
<td>-0.928</td>
<td>-0.942</td>
<td>-1.594</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.039)</td>
<td>(0.039)</td>
<td>(0.041)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Taste for soda</td>
<td>0.504</td>
<td>0.483</td>
<td>0.496</td>
<td>0.348</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.036)</td>
<td>(0.037)</td>
<td>(0.040)</td>
<td></td>
</tr>
<tr>
<td>Health importance</td>
<td>-0.189</td>
<td>-0.152</td>
<td>-0.251</td>
<td>-0.121</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.065)</td>
<td>(0.065)</td>
<td>(0.068)</td>
<td>(0.069)</td>
<td></td>
</tr>
<tr>
<td>ln(Household income)</td>
<td>-0.087</td>
<td>-0.040</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.019)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Years education)</td>
<td>-1.019</td>
<td>-0.808</td>
<td>-0.761</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other beverage tastes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Other demographics</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>County indicators</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.108</td>
<td>0.136</td>
<td>0.153</td>
<td>0.285</td>
<td>0.285</td>
</tr>
<tr>
<td>N</td>
<td>18,568</td>
<td>18,568</td>
<td>18,568</td>
<td>18,568</td>
<td>18,568</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (16), with different configurations of control variables. 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. $b_{ki}$ and $a_i$ are the average nutrition knowledge and preference measures of all (one or two) PanelViews respondents in the household. If two household heads responded to the PanelViews survey, then $b_{si}$ is the average of head 1’s ratings of both heads’ self-control, and $b_{2si}$ is the average of head 2’s ratings of both heads’ self-control. In the more common case with only one respondent in a household, $b_{si}$ is head 1’s rating of her own self-control, and $b_{2si}$ is head 2’s rating of head 1’s self-control. 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 “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. Column 5 implements the measurement error correction, using two-sample two-stage least squares with $b_{2si}$ as the endogenous self-control measure and $b_{si}$ as the instrument, with the Murphy and Topel (1986) standard error correction. Observations are weighted for national representativeness. Observations are weighted for national representativeness. Robust standard errors are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.
Table 4: Baseline Sufficient Statistics By Income Bin

<table>
<thead>
<tr>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSB consumption (ounces per week): ( \bar{s} )</td>
</tr>
<tr>
<td>SSB price (cents per ounce): ( p )</td>
</tr>
<tr>
<td>SSB demand elasticity: ( \tilde{\zeta}^c )</td>
</tr>
<tr>
<td>Elasticity of taxable income: ( \tilde{\zeta}_z )</td>
</tr>
<tr>
<td>Average marginal bias (cents per ounce): ( \tilde{\gamma} )</td>
</tr>
<tr>
<td>Externality (cents per ounce): ( e )</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>41.21</td>
</tr>
<tr>
<td>3.17</td>
</tr>
<tr>
<td>1.45</td>
</tr>
<tr>
<td>0.33</td>
</tr>
<tr>
<td>1.13</td>
</tr>
<tr>
<td>0.85</td>
</tr>
<tr>
<td>(a) Population Sufficient Statistics</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>(b) Sufficient Statistics by Income Bin</td>
</tr>
<tr>
<td>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.F, and the elasticity of taxable income, which is drawn from Chetty (2012). Panel (b) reports sufficient statistics by income bin, which can be used to compute covariances in the optimal tax formula. Income bins (denoted ( z )) correspond to the bins reported in the HomeScan data discussed in III.B. The population share in each bin ( (f) ) represents the U.S. population share with pre-tax incomes in the ranges bracketed by the midpoints between each income bin, according to Piketty et al. (2018). The statistics ( \tilde{s}(z) ), ( \tilde{\zeta}^c(z) ), ( \tilde{\xi}(z) ), and ( \tilde{\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.C to III.E. The column ( g(z) ) reports our assumed marginal social welfare weights, equal to one over the level of post-tax consumption in each bin (from Piketty et al. 2018), while ( T'(z) ) represents estimated net marginal tax rates from the same source. The column ( \bar{s}<em>{pref}(z) ) is computed as ( \bar{s}(z) - \tilde{s}</em>{inc}(z) ), where ( s_{inc}(z_n) = \bar{s}(z_1) + \exp \left( \sum_{j=2}^{n} \left( \frac{x_{j+1} - x_{j-1}}{2} \right) \ln z_j - \ln z_{j-1} \right) ).</td>
</tr>
</tbody>
</table>
Table 5: **Optimal Sugar-Sweetened Beverage Tax (Cents Per Ounce)**

<table>
<thead>
<tr>
<th></th>
<th>Existing income tax</th>
<th>Optimal income tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1.71</td>
<td>0.88</td>
</tr>
<tr>
<td>Pigouvian (no redistributive motive)</td>
<td>1.98</td>
<td>-</td>
</tr>
<tr>
<td>Weak redistributive preferences</td>
<td>1.88</td>
<td>1.64</td>
</tr>
<tr>
<td>Strong redistributive preferences</td>
<td>1.54</td>
<td>0.08</td>
</tr>
<tr>
<td>Self-reported SSB consumption</td>
<td>2.64</td>
<td>1.64</td>
</tr>
<tr>
<td>No internality</td>
<td>0.46</td>
<td>-0.15</td>
</tr>
<tr>
<td>No corrective motive</td>
<td>-0.31</td>
<td>-0.79</td>
</tr>
</tbody>
</table>

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 Corollary 1) 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 4. The second row reports the Pigouvian optimal tax, equal to \( \bar{\gamma} + e \). The following two rows report the optimal tax under weaker and and stronger redistributive social preferences than the baseline. (Social marginal welfare weights are computed to be proportional to \( y_{US}^e \), where \( y_{US} \) is post-tax income in each bin—see Appendix J.B for details—with \( \nu = 1 \), \( \nu = 0.25 \), and \( \nu = 4 \) in the baseline, and under “weak” and “strong” redistributive preferences, respectively.) “Self-reported SSB consumption” reports results using SSB consumption data from our PanelViews survey, rather than from Homescan. “No internality” is computed assumes zero bias for all individuals, while “No corrective motive” assumes zero bias and zero externality.
Table 6: **Equivalent Variation and Welfare**

<table>
<thead>
<tr>
<th>$z$</th>
<th>$f$</th>
<th>Baseline $(t = 1.71)$</th>
<th>No internality $(t = 0.46)$</th>
<th>No corrective motive $(t = -0.31)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5000</td>
<td>0.12</td>
<td>-19.18</td>
<td>-12.07</td>
<td>9.63</td>
</tr>
<tr>
<td>15000</td>
<td>0.15</td>
<td>-20.67</td>
<td>-10.96</td>
<td>8.71</td>
</tr>
<tr>
<td>25000</td>
<td>0.13</td>
<td>-19.36</td>
<td>-10.35</td>
<td>8.21</td>
</tr>
<tr>
<td>35000</td>
<td>0.10</td>
<td>-17.23</td>
<td>-9.08</td>
<td>7.19</td>
</tr>
<tr>
<td>45000</td>
<td>0.08</td>
<td>-16.37</td>
<td>-8.65</td>
<td>6.85</td>
</tr>
<tr>
<td>55000</td>
<td>0.07</td>
<td>-15.87</td>
<td>-8.30</td>
<td>6.57</td>
</tr>
<tr>
<td>65000</td>
<td>0.09</td>
<td>-14.33</td>
<td>-7.51</td>
<td>5.94</td>
</tr>
<tr>
<td>85000</td>
<td>0.09</td>
<td>-12.91</td>
<td>-6.71</td>
<td>5.30</td>
</tr>
<tr>
<td>125000</td>
<td>0.14</td>
<td>-11.68</td>
<td>-5.76</td>
<td>4.55</td>
</tr>
</tbody>
</table>

**Revenue ($ per capita)**

<table>
<thead>
<tr>
<th></th>
<th>Baseline $(t = 1.71)$</th>
<th>No internality $(t = 0.46)$</th>
<th>No corrective motive $(t = -0.31)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.26</td>
<td>8.01</td>
<td>-7.59</td>
<td></td>
</tr>
</tbody>
</table>

**Externality reduction ($ per capita)**

<table>
<thead>
<tr>
<th></th>
<th>Baseline $(t = 1.71)$</th>
<th>No internality $(t = 0.46)$</th>
<th>No corrective motive $(t = -0.31)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.48</td>
<td>3.26</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

**Total ($)**

<table>
<thead>
<tr>
<th></th>
<th>Baseline $(t = 1.71)$</th>
<th>No internality $(t = 0.46)$</th>
<th>No corrective motive $(t = -0.31)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.49</td>
<td>0.98</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports the welfare effects of introducing the optimal tax, decomposed into equivalent variation by income group, revenue, and externality reduction. This computation is performed for three specifications—our baseline, the specification assuming no internalities, and the specification assuming no internalities or externalities. In each case, the total is computed by weighting the EV by population share and welfare weight (as reported in Table 4), and added to revenues and externality reductions.
Notes: This figure presents the average purchases of sugar-sweetened beverages in Nielsen Homescan data for 2006-2015, by income group. Purchases are measured in liters per “adult equivalent” in a household, 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.
Figure 2: Nutrition Knowledge and Self-Control vs. Sugar-Sweetened Beverage Consumption

Notes: These figures present average purchases of sugar-sweetened beverages for each household’s most recent year in the Nielsen Homescan data against the household’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 statements, “I drink soda pop or other sugar-sweetened beverages more often than I should,” and, if there is a second household head, “The other head of household in my house drinks soda pop or other sugar-sweetened beverages more often than they should.” Answers were coded as “Definitely”=0, “Mostly”=1/3, “Somewhat”= 2/3, and “Not at all”=1. Purchases are measured in liters per “adult equivalent” in a household, 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.
Figure 3: Intuition for the Price Instruments

(a) Relative Prices Vary Across Retail Chains

(b) Uniform Pricing Within an Example Retail Chain

Notes: This figure presents the prices for the best-selling UPC in the Nielsen RMS data. Panel (a) presents the quarterly average price (unweighted across stores and weeks) for each retail chain that sold this UPC over the full sample. Panel (b) presents the quarterly average price (again unweighted across weeks) for each store at an example retail chain.
Figure 4: Contemporaneous First Stage and Reduced Form of the Local Price Instrument

(a) First Stage

(b) Reduced Form

Notes: 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 (12). Both relationships are residual of the other variables in Equation (12): 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” in a household, 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.
Figure 5: Causal Income Effects vs. Between-Income Preference Heterogeneity

Notes: The dark circles plot average purchases of sugar-sweetened beverages in Nielsen Homescan data for 2006-2015, by income group. Purchases are measured in liters per “adult equivalent” in a household, 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 represents predicted average purchases at each income level that would result from applying the estimated income elasticity $\hat{\xi} \approx 0.25$ to the average consumption of households with annual income under $10,000: \bar{s}(z_{min}) + \hat{\xi} \ln z$. The x’s are $\bar{s}_{pref}(z)$, the difference between actual consumption and consumption predicted only using income elasticity. Observations are weighted for national representativeness.
Notes: These figures present average nutrition knowledge (in Panel (a)) and self-control (in Panel (b)) of Nielsen Homescan households, 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 statements, “I drink soda pop or other sugar-sweetened beverages more often than I should,” and, if there is a second household head, “The other head of household in my house drinks soda pop or other sugar-sweetened beverages more often than they should.” Answers were coded as “Definitely”=0, “Mostly”=1/3, “Somewhat”= 2/3, and “Not at all”=1. Observations are weighted for national representativeness.
Figure 7: Preferences vs. Household Income

(a) Taste for Soda

(b) Health Importance

Notes: These figures present average taste for soda (in Panel (a)) and health importance (in Panel (b)) for Nielsen Homescan households, 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.
Figure 8: **Quantity Effect of Bias by Household Income**

![Graph showing the quantity effect of bias by household income.](image)

Notes: This figure presents the average quantity effect of bias \( \hat{\tau}(b^V - b_i) \) by income, with \( \hat{\tau} \) from column 5 of Table 3, using Homescan purchase data. Observations are weighted for national representativeness.

Figure 9: **Average Marginal Bias by Household Income**

![Graph showing the average marginal bias by household income.](image)

Notes: This figure presents average marginal bias by income from Equation (??), with \( \hat{\tau} \) from column 5 of Table 3, using Homescan purchase data. Observations are weighted for national representativeness.
Online Appendix: Not for Publication

Regressive Sin Taxes, with Applications to the Optimal Soda Tax

*Hunt Allcott, Benjamin B. Lockwood, Dmitry Taubinsky*
A Income Elasticity Definitions

We define labor supply responses to include any “circularities” due to the curvature of the income tax function, which is assumed to be differentiable. Thus, following Jacquet and Lehmann (2014), we define a tax function $\hat{T}$ which has been locally perturbed around the income level $z_0$ by raising the marginal tax rate by $\tau$ and reducing the tax level by $\nu$:

$$\hat{T}(z; z_0, \tau, \nu) = T(z) + \tau(z - z_0) - \nu.$$  

Let $z^*(\theta)$ denote a type $\theta$’s choice of earnings under the status quo income tax $T$, and let $\hat{z}(\theta; \tau, \nu)$ denote $\theta$’s choice of earnings under the perturbed income tax $\hat{T}(z; z^*(\theta), \tau, \nu)$. Then the compensated elasticity of taxable income is defined in terms of the response of $\hat{z}$ to $\tau$, evaluated at $\tau = \nu = 0$:

$$\zeta_c^z(\theta) := \left( -\frac{\partial \hat{z}(\theta; \tau, 0)}{\partial \tau} \bigg|_{\tau=0} \right) \frac{1 - T'(z^*(\theta))}{z^*(\theta)}.$$  

The income tax is similarly defined in terms of the response of $\hat{z}$ to a tax credit $\nu$ (this statistic will be nonpositive if leisure is a non-inferior good):

$$\eta_z(\theta) := \left( \frac{\partial \hat{z}(\theta; 0, \nu)}{\partial \nu} \bigg|_{\nu=0} \right) (1 - T'(z^*(\theta))).$$  

These definitions are comparable to those in Saez (2001), except that they include circularities and thus permit a representation of the optimal income tax in terms of the actual earnings density, rather than the “virtual density” employed in that paper.

B Extensions

Our baseline results can be extended in a number of ways, which we do in Appendix TK. In the appendix, we characterize results when the sin good $s$ is a composite good consisting of several different items, such as soft drinks of various sizes. Second, we derive a general condition for the optimal sin tax when there are more dimensions of consumption, which allows for substitution across goods with various degrees of marginal bias. In our companion paper, TKTK, we also study the case in which individuals are may under-react to the implicit labor supply incentives generated by changes in the commodity tax on $s$.

B.A Composite Goods

We now allow for the possibility that the sin good $s$ is in fact a composite of several sin goods, $(s_1, s_2, \ldots, s_n)$ with pre-tax prices $(p_1, p_2, \ldots, p_n)$, and with marginal (money metric) bias of $\gamma_i(\theta) = \frac{U_i'}{U_i'} - \frac{V_i'}{V_i'}$ at the optimum. Assume all sin goods are subject to a common ad valorem tax $\tau$, so that
the post-tax price of each \( s_i \) is \( p_i(1 + \tau) \). Under this composite setup, the optimal sin tax \( \tau \) still satisfies the general expression from Proposition 1 in the case where units of \( s \) are normalized so that \( p = 1 \), provided the parameters therein are reinterpreted as follows:

- \( s(\theta) := \sum_i p_i s_i(\theta) \), representing total revenues (net of taxes) from sin good spending by \( \theta \)
- \( \gamma(\theta) := \sum_i \gamma_i(\theta) \frac{\partial s_i(q_i, \theta)}{\partial q_i} p_i / \sum_i \frac{\partial s_i(q_i, \theta)}{\partial q_i} p_i \), representing the response-weighted bias, where \( q_i := (1 + \tau) p_i \)
- \( \zeta(\theta) := \sum_i \frac{\partial s_i(q_i, \theta)}{\partial q_i} p_i (1 + \tau)/s(\theta) \), representing the aggregate elasticity of sin good spending

We adopt this specification for use in the empirical implementation of the next section, since there our data involve consumers spending on multiple soda and SSB products.

### B.B More Dimensions of Consumption

Our results easily generalize to \( N > 2 \) dimensions of consumption. Let utility be given by 

\[ U(c_1, c_2, \ldots, c_n, z; \theta) \]

with \( c_1 \) representing the numeraire, and assume that \( U_{c_1}/U_z = V_{c_1}/V_z \). Let \( S \) denote the Slutsky matrix of compensated demand responses in which the \( j,i \) entry is the compensated demand response of \( c_j \) with respect to \( t_i \).

Let \( \gamma_i(\theta) \), for \( i \geq 2 \), denote the bias in consumption dimension \( i \), and let \( \bar{\gamma}_i(z) \) denote the statistic for \( z \)-earners. Let \( \bar{\gamma}_{ij} \) denote the average marginal bias from consumption of \( c_j \) with respect to tax \( i \):

\[
\bar{\gamma}_{ij} = \int_\Theta \gamma_j(\theta, t, T) \left( \frac{dH_i(\theta,t,T)}{dt_i} \right) d\mu(\theta)
\]

where \( H_j \) is the compensated (Hicksian) demand for good \( j \). We also define \( \sigma_{ij} := Cov_H \left[ g(z), \frac{\gamma_j(\theta, t, T)}{\gamma_j(\theta, t, T)} \right] \).

We define \( R \) to be a column vector in which the \( j \)th entry is \( R_i = -\sum \frac{dH_j}{dt_i} \bar{\gamma}_{ij}(\bar{g} + \sigma_{ij}) \). That is, \( R_i \) is the total corrective benefit from increasing tax \( t_i \), keeping income constant.

As before, we define \( \phi_j(z) \) to be the portion of \( c_j \) consumption explained by income effects, and we define \( \bar{\phi}_j(z) := (c_j - \phi_j) / C_j \) to be the variation due to preference heterogeneity. We define \( \rho \) to be the column vector in which the \( j \)th entry is \( Cov[g(z), \bar{\phi}_j(z)] \).

**Proposition 3.** The optimal commodity taxes \( \mathbf{t} = (t_1, \ldots, t_N) \) satisfy

\[ \mathbf{t}S = \mathbf{\rho} - \mathbf{R} \tag{19} \]

and the optimal income tax satisfies
\[
\frac{T'(z)}{1 - T'(z)} = \sum_i \frac{g(z)\gamma_i(z) - t_i}{p_i + t_i} \eta_i(z) + \frac{1}{\zeta(z) h(z)} \int_z^\infty (1 - \hat{g}(x)) h(x) dH(x)
\]

(20)

\section{Proofs of propositions in the paper}

\subsection{Proof of Proposition 1, including additional derivations without assumption 4}

\textbf{Income tax:}

Following Saez (2001), we have the following effect from increasing the marginal tax rate between \(z\) and \(z + dz\):

1. Direct effect (fiscal and welfare): \(\int_{x=z}^\infty (1 - g(x)) h(x) dH(x)\)

2. Compensated elasticity effect (including both fiscal and welfare components):

\[
- \zeta'(z) \cdot z \cdot \frac{T'}{1 - T'} h(z) - \mathbb{E} \left[ \frac{\zeta'(z)}{\zeta(z)} \cdot z \cdot \frac{t - g(z)\gamma(\theta) - e}{p + t} \cdot \eta(\theta) h(z) \mid z(\theta) = z \right] =

- \zeta'(z) \cdot z \cdot \frac{T'}{1 - T'} h(z) - \zeta'(z) \cdot z \cdot \mathbb{E} \left[ \frac{\zeta'(z)}{\zeta(z)} \cdot \frac{t - g(z)\gamma(\theta) - e}{p + t} \cdot \eta(\theta) \mid z(\theta) = z \right]
\]

3. Income effect (fiscal and welfare): \(- \int_{x\geq z} \eta_s(x) \cdot \frac{T'}{1 - T'} h(x) - \int_{x\geq z} \mathbb{E} \left[ (1 + \eta_s(\theta)) \cdot \frac{t - g(x)\gamma(\theta) - e}{p + t} \cdot \eta(\theta) \mid z(\theta) = x \right] h(x)\)

where the second term combines the income effect on \(s\) from the change in after-tax income with the income effect on \(s\) from the adjustment in earnings earnings \(z\).

These three terms must sum to zero, as the net welfare effect of this reform should be zero. Using the definition of \(\hat{g}(z)\) from footnote 22, the direct effect from 1 combines with the income effects from 3 to become \(\int_{x=z}^\infty (1 - \hat{g}(x)) h(x) dH(x)\). Therefore, the optimal income tax satisfies

\[
\frac{T'(z)}{1 - T'(z)} = \frac{1}{\zeta'(z) \cdot z \cdot h(z)} \int_{x=z}^\infty (1 - \hat{g}(x)) h(x) dH(x) - \mathbb{E} \left[ \frac{\zeta'(z)}{\zeta(z)} \cdot \frac{t - g(z)\gamma(\theta) - e}{p + t} \cdot \eta(\theta) \mid z(\theta) = z \right]
\]

Using Assumption 4, the expectation simplifies to \(\mathbb{E} \left[ \frac{g(z)\gamma(\theta) + e - t}{p + t} \cdot \eta(\theta) \mid z(\theta) = z \right]\).

\textbf{Commodity tax:}

First, consider the total effect on welfare of a marginal increase in the commodity tax \(dt\). The total welfare effect, written in terms of the marginal value of public funds, can be decomposed into the following components:
• **Mechanical revenue effect:** the reform mechanically raises revenue from each consumer by $dt \cdot s(\theta)$, for a total of $dt \cdot S$.

• **Mechanical welfare effect:** the reform mechanically reduces each consumer’s net income by $dt \cdot s(\theta)$. To isolate the mechanical effect, we compute the loss in welfare as if this reduction all comes from composite consumption $c$ for a welfare loss of $dt \cdot s(\theta)g(\theta)$, and we account for adjustments in $s$ and $z$ in the behavioral effects below. Thus the total mechanical welfare effect is $-dt \int_{\Theta} s(\theta)g(\theta)d\mu(\theta)$.

• **Direct effect on sin good consumption:** the reform causes an increase in the commodity tax alters the utility that type $\theta$ would choose to consume conditional on earning income $z$. This generates a total fiscal externality through the income tax equal to $-dt \cdot \int_{\Theta} g(\theta)\gamma(\theta)\zeta(\theta)s(\theta)_{\mu+1}d\mu(\theta)$ and an externality effect equal to $dt \int_{\Theta} c\zeta(\theta)s(\theta)_{\mu+1}d\mu(\theta)$.

• **Effect on earnings:** The reform causes a change in income tax revenue collected from type $\theta$ equal to $\frac{d\theta}{dt}T'(z(\theta))$. To compute $\frac{d\theta}{dt}$, we use Lemma 1 from Saez (2002a), which carries through in this context: the change in earnings for type $\theta$ due to the increase in the commodity tax $dt$ is equal to the change in earnings which would be induced by imposing a type-specific income tax reform, raising the income tax by $dT(z) = dt \cdot \hat{s}(z|\theta)$, where $\hat{s}(z|\theta)$ is defined as the amount of $s$ that type $\theta$ would choose to consume conditional on earning income $z$. Intuitively, an increase in the commodity tax alters the utility that type $\theta$ would gain from selecting each possible level of earnings, in proportion to the amount of $s$ that $\theta$ would consume at each income. This alternative income tax reform reduces the utility $\theta$ would realize at each possible income by instead altering the income tax. The resulting adjustment in earnings can be decomposed into a compensated elasticity effect and an income effect. The former depends on the change in the marginal income tax rate, which equals $dt \cdot s(\theta) = \frac{\partial s(\theta)}{\partial z} = \frac{\eta(\theta)(1-T'(z(\theta)))}{p+\tau}$; the latter depends on the change in the tax level, which equals $dt \cdot s(\theta)$. Combining these effects, we have $\frac{d\theta}{dt} = -\zeta^c(\theta)\left(\frac{z(\theta)}{1-T'(z(\theta))}\right)\frac{\partial s(\theta)}{\partial z} - \frac{\eta(s(\theta))}{1-T'(z(\theta))}$. This generates a total fiscal externality through the income tax equal to $-dt \cdot \int_{\Theta} T'(z(\theta))\frac{\partial s(\theta)}{\partial z} + \zeta(\theta)s(\theta)\mu(\theta)$.

• **Indirect effects on sin good consumption:** The change in earnings affects consumption indirectly. This generates additional fiscal externalities and additional welfare effects. These total to $dt \cdot \int_{\Theta} \frac{d\theta}{dt} \cdot \frac{\partial s(\theta)}{\partial z}(g(\theta)\gamma(\theta)) + e - t)\mu(\theta) = -dt \cdot \int_{\Theta} \left(zs\gamma+\zeta(\theta)s\theta\right) \frac{n(\gamma+c-e-t)}{p+\tau}d\mu(\theta)$.

Combining these components, the total welfare effect of the tax reform $dt$ is equal to
Since the income tax is optimal by assumption, we can simplify \( \frac{dW}{dt} \) by subtracting \( \frac{dW}{dT} = 0 \). The resulting expression can be simplified by writing the integral over incomes and over types conditional on income. Then employing the assumption that \( g(\theta) \) is constant conditional on income,
the mechanical effects cancel out, leaving

\[
\frac{dW}{dt} = \int_0^\infty \left( \frac{s(z)z^c(z)}{p + t} \left( t - g(z)\gamma(z) - e \right) - \frac{1}{p + t} \left( g(z)\Sigma_{\eta\gamma,s}^{(z)} - t\Sigma_{\eta,s}^{(z)} \right) \right) dH(z)
\]

\[
+ \int_0^\infty \left( \int_{x \geq z} (1 - \hat{g}(x)) dx \left( \frac{\partial \hat{s}(z)}{\partial z} - \hat{s}'(z) \right) + \frac{\Sigma_{\gamma,s}^{(z)}}{z\zeta^c(z)} \right) + T'(z)\Sigma_{\eta,s}^{(z)} dx dH(z)
\]

(Recall that \( \Sigma_{a,b}^{(z)} = Cov_{a|z}[a(\theta), b(\theta)] z(\theta) = z^* \) denotes the income-conditional covariance between two variables \( a(\theta) \) and \( b(\theta) \) at the optimum.) Under the optimal commodity tax, (21) is equal to zero. This expression is similar to the derivation in Saez (2002a) of the effect of introducing a small commodity tax (expression (5) in that paper) with three important differences. First, (21) depends on the bias \( \gamma \), which is not present in that paper. Second, since Saez (2002a) considers the introduction of a small commodity tax (rather than the perturbation of a possibly non-zero tax) there is no first-order fiscal externality from the change in \( s \) consumption. Since we are interested in characterizing the optimal tax in situations where the optimal tax is non-zero, (21) includes that fiscal externality. Third, the income responses \( \zeta^c \) and \( \eta_s \) in this expression represent the full income response, accounting for nonlinearities in \( T(z) \), which simplifies the expression.

We now compute the covariance terms, employing assumption 4 from Section II.B.1 and the fact that \( \eta = \frac{\xi - \frac{(p + t)\zeta}{z}}{1 - T'} \)

\[
\bullet \Sigma_{\eta\gamma,s}^{(z)} = \frac{\xi(p + t)}{z(1 - T')} \Sigma_{\gamma,s}^{(z)} = \frac{\eta(z)}{s(z)} \gamma(z)
\]

\[
\Sigma_{\eta,s}^{(z)} = \frac{\xi(p + t)}{z(1 - T')} \Sigma_{s,s}^{(z)} = \frac{\eta(z)}{s(z)} \gamma(z)
\]

\[
\Sigma_{\zeta^c,\eta}^{(z)} = 0
\]

\[
\Sigma_{\zeta^c,s}^{(z)} = 0
\]

Thus the necessary condition for optimality reduces to

\[
\frac{dW}{dt} = \int_0^\infty \left( \frac{s(z)z^c(z)}{p + t} \left( t - g(z)\gamma(z) - e \right) - \frac{1}{p + t} \left( g(z)\Sigma_{\eta\gamma,s}^{(z)} - t\Sigma_{\eta,s}^{(z)} \right) \right) dH(z)
\]

\[
+ \int_0^\infty \left( \int_{x \geq z} (1 - \hat{g}(x)) dx \left( \frac{\partial \hat{s}(z)}{\partial z} - \hat{s}'(z) \right) + \frac{\Sigma_{\gamma,s}^{(z)}}{z\zeta^c(z)} \right) + T'(z)\Sigma_{\eta,s}^{(z)} dx dH(z)
\]

The second term can be integrated by parts and rewritten as follows:

\[
\int_0^\infty \int_{x = z}^\infty (1 - \hat{g}(x)) dH(x) \left( s_{inc}(z(\theta)) - \hat{s}'(z) \right) dx dz = \int_0^\infty (1 - \hat{g}(z))h(z) \left( \int_0^z s_{inc}(x) dx - \Delta \hat{s}(z) \right) dz
\]

\[
= \int_0^\infty (1 - \hat{g}(z)) \left( \hat{s}_{inc}(z) - \hat{s}(z) \right) dH(z)
\]

\[
= Cov[\hat{g}(z), \hat{s}_{pref}(z)]
\]
Substituting this into 22 and rearranging yields

\[
\begin{align*}
t & = \int \left( \bar{s}(z) \bar{c}(z) (g(z) \bar{\gamma}(z) + e) + \bar{\gamma}^{(z)}(z) g(z) \Sigma_{s,s}^{(z)} \right) dH(z) - (p + t) \text{Cov}[\bar{g}(z), \bar{s}_{\text{pref}}(z)] \\
& = \mathbb{E}[g(z)] \mathbb{E}[\bar{s}(z) \bar{c}(z) \bar{\gamma}(z)] + \mathbb{E}[\bar{s}(z) \bar{c}(z) e] + \text{Cov}[g(z), \bar{s}(z) \bar{c}(z) \bar{\gamma}(z)] + \mathbb{E} \left[ \frac{\bar{\gamma}^{(z)}(z) g(z) \Sigma_{s,s}^{(z)}}{\bar{s}(z)} \right] \\
& \quad - (p + t) \text{Cov}[\bar{g}(z), \bar{s}_{\text{pref}}(z)] \\
& = \frac{\mathbb{E}[\bar{s}(z) \bar{c}(z) \bar{\gamma} + \sigma] + \mathbb{E} \left[ \frac{\bar{\gamma}^{(z)}(z) g(z) \Sigma_{s,s}^{(z)}}{\bar{s}(z)} \right]}{\bar{s} \bar{c}^2} + \mathbb{E} \left[ \frac{\bar{\gamma}^{(z)}(z) g(z) \Sigma_{s,s}^{(z)}}{\bar{s}(z)} \right] - (p + t) \text{Cov}[\bar{g}(z), \bar{s}_{\text{pref}}(z)] \\
& = \frac{\mathbb{E}[\bar{s}(z) \bar{c}(z) \bar{\gamma} + \sigma] + \mathbb{E} \left[ \frac{\bar{\gamma}^{(z)}(z) g(z) \Sigma_{s,s}^{(z)}}{\bar{s}(z)} \right]}{\bar{s} \bar{c}^2} + \mathbb{E} \left[ \frac{\bar{\gamma}^{(z)}(z) g(z) \Sigma_{s,s}^{(z)}}{\bar{s}(z)} \right]
\end{align*}
\]

C.B

C.C  Extension to a Composite Sin Good

A composite sin good can be represented in the utility function by assuming that individuals maximize decision utility \( U(c, s_1, s_2, \ldots, s_n, z; \theta) \) subject to their budget constraint \( c + (1+\tau) \sum_{i=1}^n p_i s_i \leq z - T(z) \), while the policymaker seeks to maximize aggregate normative utility \( V(c, s_1, s_2, \ldots, s_n, z; \theta) \).

As in the proof of Proposition 1 above, consider the implications of a small joint perturbation to the commodity tax \( d\tau \), combined with an offsetting compensation through the income tax which preserves labor supply choices for all individuals. This reform has the following effects:

- **Mechanical revenue effect:** \( d\tau \cdot \sum_i p_i s_i(\theta) \).
- **Mechanical welfare effect:** the reform mechanically reduces each consumer’s net income by \( d\tau \cdot \sum_i p_i s_i(\theta) \), for a mechanical welfare loss of \( d\tau \cdot \sum_i p_i s_i(\theta) g(\theta) \). Thus the total mechanical welfare effect is \( -d\tau \int_{\Theta} \sum_i p_i s_i(\theta) g(\theta) d\mu(\theta) \). Thus by defining \( s(\theta) := \sum_i p_i s_i(\theta) \) to represent total revenues (net of taxes) from sin good spending by \( \theta \), the mechanical expressions from C.A remain unchanged.
- **Direct effect on sin good consumption:** each consumer alters their consumption of each sin good \( i \) by \( d\tau \cdot \frac{\partial s_i(q_i, \theta)}{\partial q_i} \cdot p_i \), where \( q_i := (1+\tau) p_i \), generating a fiscal externality of \( d\tau \int_{\Theta} \tau p_i \frac{\partial s_i(q_i, \theta)}{\partial q_i} d\mu(\theta) \) and a behavioral effect (in terms of public funds of the allocation) of \( \frac{\alpha(\theta)}{\lambda} \sum_i \left( V'_i(\theta) - V'_i(\theta)p_i(1+\tau) \right) \left( d\tau \frac{\partial s_i(q_i, \theta)}{\partial q_i} \right) p_i \), or \( d\tau \int_{\Theta} g(\theta) \sum_i \gamma_i(\theta) \frac{\partial s_i(q_i, \theta)}{\partial q_i} p_i d\mu(\theta) \). Therefore we define \( \gamma_i(\theta) := \sum_i \gamma_i(\theta) \frac{\partial s_i(q_i, \theta)}{\partial q_i} p_i / \sum_i \frac{\partial s_i(q_i, \theta)}{\partial q_i} p_i \),
representing the response-weighted bias, and \( \zeta(\theta) := \sum_i \frac{\partial s_i(q_i, \theta)}{\partial q_i} p_i \cdot (1 + \tau) / s(\theta) \), representing the aggregate elasticity of sin good spending. Then the direct effect on sin good consumption can be written \( d\tau \int g(\theta) \gamma(\theta) \zeta(\theta) \frac{s(\theta)}{1 + \tau} d\mu(\theta) \), equivalent to the expression from C.A when the sin good units are normalized so that \( p = 1 \).

\[ \]
\[
\frac{dW}{dT} = \sum_j \int \left( \bar{c}_i(z)(1 - g(z)) - \frac{\eta_j \tilde{c}_i}{p + t_j} (t_j - g(z)\gamma_j) \right) dH(z) \\
- \int_0^{T'} \frac{T'(z)}{1 - T'} \left( \zeta_t z \tilde{c}_i + \eta_z \tilde{c}_i(z) \right) dH(z) \\
+ \sum_j \int \left( \zeta_t z \tilde{c}_i + \eta_z \tilde{c}_i \right) dH(z) \\
= \sum_j \int \left( \bar{c}_i(z)(1 - g(z)) - \frac{\eta_j \tilde{c}_i}{p + t_j} (t_j - g(z)\gamma_j) \right) dH(z) \\
- \sum_j \int \int_{s \geq z} (1 - \hat{g}(s)) ds \left( \bar{c}_j(z) + \frac{\eta_z \tilde{c}_i}{\tilde{c}_j} \right) dH(z)
\]

Setting \( \frac{dW}{dT} = 0 \) and subtracting that from \( \frac{dW}{dt} \) shows that

\[
\frac{dW}{dt_i} = \sum_j \int \left( \bar{c}_i(z)(1 - g(z)) + \frac{\partial}{\partial t_i} \mathcal{H}_j(t_j - g(z)\gamma_j) \right) dH(z) \\
- \int \int_{s \geq z} (1 - \hat{g}(s)) ds \left( \frac{\partial \tilde{c}_i(z)}{\partial z} - \bar{c}_j(z) \right) dH(z)
\]

We can transform equation (26) using integration by parts as in the proof of Proposition 1, exactly as we did in equation (23). This yields the desired result.

**D Additional Theoretical Results**

**D.A Optimal commodity tax in terms of social marginal welfare weights**

The key variable needed for this solution is the following: “When we raise the commodity tax by \( dt \), how much money can we give back to each person to offset the tax so as to keep the average labor supply choices of each \( z \)-earner constant?” Roughly, we seek to characterize the extent to which we can offset the regressivity costs of the commodity tax by making the income tax more progressive. We begin with a lemma providing this characterization:

**Lemma 1.** Let \( \chi(z) := \phi(z) - \int_0^z w(x, z) \frac{\eta_z}{\tilde{c}_z} (s(x) - \phi(x)) dx \), where \( w(x, z) = e^{\int_{x'=x}^{x'=z} \frac{\eta_z}{\tilde{c}_z} dx'} \). Then increasing the commodity tax by \( dt \) and decreasing the income tax by \( \chi(z)dt \) leaves the average labor supply of \( z \)-earners unchanged.

**Proof.** We can instead ask the following intuitive question: “When we raise the commodity tax by \( dt \), how much money can we give back to each person to offset the tax so as to keep the average
labor supply choices of each z-earner constant?” Call this quantity $\chi(z)$. Under Assumption A, this term satisfies the following differential equation:

$$\ddot{\zeta}_z c z \chi'(z) + \eta_z \chi(z) = \ddot{\zeta}_z c z \psi(z) + \eta_z \bar{s}(z)$$

(27)

where $\psi(z) = E \left[ \frac{ds(z, \theta)}{dz} | z^*(\theta) = z \right]$. The right-hand side is the impact of the commodity-tax, as shown in Appendix C.A. The left-hand side follows straightforwardly. The first term is just compensated effect of increasing the marginal tax rate, while the second term is the income effect. Rearranging yields

$$\chi'(z) + \frac{\eta_z}{\dot{\zeta}_z} \chi(z) = \psi(z) + \frac{\eta_z}{\dot{\zeta}_x} \bar{s}(z)$$

The solution to this first-order differential equation is

$$\chi(z) = \int_{x=0}^{z} e^{\int_{x}^{x'} \frac{\eta_z}{\dot{\zeta}_x} dx'} \left( \psi(x) + \frac{\eta_z}{\dot{\zeta}_x} \bar{s}(x) \right) dx + K e^{\int_{z}^{0} \frac{\eta_z}{\dot{\zeta}_x} dx'}$$

where $K$ is some integration constant.

Now set $\phi(z) = \int_{x=0}^{z} \psi(x) dx + \bar{s}(0)$. Then

$$\frac{d}{dx} \phi(x) e^{\int_{x}^{x'} \frac{\eta_z}{\dot{\zeta}_x} dx'} = e^{\int_{x}^{x'} \frac{\eta_z}{\dot{\zeta}_x} dx'} \left( \psi(x) + \frac{\eta_z}{\dot{\zeta}_x} \phi(x) \right)$$

and thus

$$\chi(z) = \phi(z) + \frac{K - \phi(0)}{e^{\int_{z}^{0} \frac{\eta_z}{\dot{\zeta}_x} dx'}} + \int_{0}^{z} w(x, z) \frac{\eta_z}{\dot{\zeta}_x} (\bar{s}(x) - \phi(x)) dx$$

where $w(x, z) = e^{\int_{x'}^{x} \frac{\eta_z}{\dot{\zeta}_x} dx'}$. And to get the initial conditions right $\chi(0) = \phi(0)$, we must have $K = \phi(0)$, so that

$$\chi(z) = \phi(z) + \int_{0}^{z} w(x, z) \frac{\eta_z}{\dot{\zeta}_x} (\bar{s}(x) - \phi(x)) dx$$

(28)

The intuition for the “income effect adjustment” term is that the higher the income effect, the more the commodity tax increases labor supply through the income effect, and thus the higher $\phi$ can be.

Note that $\chi(z) = \phi(z)$ when $\eta_z = 0$. That is, in the absence of labor supply income effects, the term $\phi(z)$—which we defined as the portion of $s$ consumption explained by consumption income effects—is the extent to which the income tax should be optimally lowered for $z$-earners. When $\eta_z \neq 0$ and $\phi(z) < \bar{s}(z)$, $\chi(z)$ will be higher than $\phi(z)$. Intuitively, this is because giving back $\phi(z) dt$
to consumers offsets the effective impact on the marginal keep-rate from before-tax earnings, but does not fully offset the income effect as it leaves consumers poorer by an amount \((s - \phi)dt\). Analogous to \(\tilde{\phi}(z)\), we define the index \(\tilde{\chi}(z) := \frac{\tilde{s}(z) - \chi(z)}{s}\).

**Proposition 4.** The optimal commodity tax \(t\) satisfies

\[
t = \tilde{\gamma}(\tilde{g} + \sigma) + \frac{p + t}{\zeta c} E[(1 - g(z))\tilde{\chi}(z)] - \frac{1}{\zeta c} \int \tilde{\chi}(z)\eta(z)(t - g(z)\tilde{\gamma}(z))
\]

**Proof.** Consider now the welfare impacts of a reform that increase the commodity tax by \(dt\) and the income tax \(\chi(z)\) at each point \(z\). Because by Lemma has the following impact on welfare, under Assumption A:

\[
dW = \int_z \left(\frac{(1 - g(z))(\tilde{s}(z) - \chi(z))}{p + t} - \tilde{\chi}(z)(t - \gamma)\frac{\zeta c(z)}{p + t} - \eta(z)(t - g(z)\tilde{\gamma}(z))(\tilde{s}(z) - \chi(z))\right) dH(\theta)
\]

Equation (30) follows from the following effects. The first effect is the direct revenue and welfare effect of decreasing each individual’s income by \(\tilde{s}(z) - \chi(z)\). The second effect corresponds to the compensated demand response, which generates both a welfare effect and a fiscal externality from substitution. The third effect comes from the effect that a wealth decrease of \(\tilde{s}(z) - \chi(z)\) has on \(s\) consumption; again, this generates both a fiscal externality and a welfare effect.

\[
t = \tilde{\gamma}(\tilde{g} + \sigma) - \frac{1}{\zeta c} \int \tilde{\chi}(z)\eta(z)(t - g(z)\tilde{\gamma}(z)) + \frac{p + t}{\zeta c} E[(1 - g(z))\tilde{\chi}(z)].
\]

\[
= \tilde{\gamma}(\tilde{g} + \sigma) + \frac{1}{\zeta c} E[\tilde{\chi}(z)\eta(z)g(z)\tilde{\gamma}(z)] - \frac{t}{\zeta c} E[\tilde{\chi}(z)\eta(z)] + \frac{p + t}{\zeta c} E[(1 - g(z))\tilde{\chi}(z)]
\]

Thus

\[
t = (\zeta c - E[\tilde{\chi}(z)(1 - g(z) - \eta(z))])
\]

\[
= \zeta c \tilde{\gamma}(\tilde{g} + \sigma) + pE[(1 - g(z))\tilde{\chi}(z)] + E[\tilde{\chi}(z)\eta(z)g(z)\tilde{\gamma}(z)]
\]

from which the result follows.

Proposition 4 provides a commodity tax formula similar to Corollary 1, with two differences. First, the covariance term \(\text{Cov}[g(z), \tilde{\phi}(z)]\) corresponding to regressivity costs is written more generally as \(E[(g(z) - 1)\tilde{\chi}(z)]\). These two terms are equivalent when \(E[g(z)] = 1\), but differ when
$E[g(z)] > 1$ in the presence of labor supply income effects ($\eta_z \neq 0$). Second, the presence of income effects $\eta$ means that a tax reform that decreases consumers’ after-tax income by $s(z) - \chi(z)$ also impacts consumption of $s$, beyond the compensated elasticity with respect to prices. This generates fiscal externalities proportional to $t\eta(s(z) - \chi(z))$ and additional corrective benefits proportional to $g\gamma(s(z) - \chi(z))$. Combined, this generates the third term in equation (29) above. In the absence of income effects, the formula reduces to

$$t = \gamma(g + \sigma) - \frac{p + t}{\zeta}Cov\left[g(z), \tilde{\phi}(z)\right]$$

which can be solved for $t$ to get the formula in Corollary 1.

**D.B For discussion of term with conditional covariances**

As is apparent from Equation (21), this derivation of the optimal commodity tax using the perturbation approach can accommodate heterogeneity conditional on income. Although Assumption A simplifies the form by assuming that the income-conditional covariances which appear there are equal to zero, that restriction is not necessary for the basic approach, and here we relax it by writing the necessary conditional for the optimal tax with flexible multidimensional heterogeneity:

$$t = \gamma(g + \sigma) - \frac{p + t}{\zeta}Cov\left[g(z), \tilde{\phi}(z)\right] - \frac{p + t}{\zeta} \int_{0}^{\infty} \left( \int_{z \in C_2} ds \left( \Sigma^{(z)}_{\chi, \zeta} \right) \right) dH(z)$$

Equation 31 resembles Equation (6) in Proposition ??, with additional terms corresponding to the income conditional covariances between income effects, elasticities, and consumption concentration ($c_2(\theta)/C_2$). (Note that although this expression allows for more general multidimensional heterogeneity, it still requires that the first-order approach characterizes the optimal tax system—i.e., heterogeneity must not lead to discontinuous jumping of earnings or consumption bundles in response to small tax perturbations.) As with Proposition ??, this equation is implicit, and thus it does not represent a closed form solution for the optimal commodity tax—however it could be solved numerically via fixed point iteration, like the simulations in Section ?? Additionally, it provides a heuristic guide to the likely comparative static effects of the income-conditional covariance terms on the optimal tax. For example, focusing on the final term, suppose $\Sigma^{(z)}_{\chi, \zeta} > 0$ for all $z$, indicating that labor supply income effects are smallest in magnitude (least negative) among those who consume the most $c_2$ at each income—then the optimal commodity tax will tend to be reduced. Intuitively, the joint reform of a higher $t$ and a regressivity-offsetting income tax reform confers a
greater benefit on those who consume little $c_2$ at each income. If those individuals exhibit larger labor supply responses to such benefits, then the negative fiscal externality (through the income tax) of this joint reform is larger, suggesting a lower optimal income tax. Similar reasoning can be applied to the other covariance terms.

D.C Optimal taxes when $\eta_z = 0$ but $\eta$ is not negligible

Corollary 3. Suppose that labor supply income effects are negligible $\eta_z \approx 0$. Then

$$t \approx \frac{\tilde{c} \gamma(g + \sigma) + \text{Cov}[\tilde{g}(z), \eta g(z) \tilde{\gamma}(z)]}{\tilde{c} + \text{Cov}[g(z) + \eta \tilde{\phi}(z)]} - p \frac{\text{Cov}[g(z), \tilde{\phi}(z)]}{\tilde{c} + \text{Cov}[g(z) + \eta \tilde{\phi}(z)]}$$

$$\frac{T'(z)}{1 - T'(z)} \approx \frac{g(z) \gamma(z) - t}{p + t} \eta(z) + \frac{1}{\tilde{c} z h(z)} \left[ \int_{\gamma(z)}^{\infty} \frac{g(x) \gamma(z) - t}{p + t} \eta(x) (1 + \eta(z)) dH(x) + \int_{\gamma(z)}^{\infty} (1 - g(x)) dH(x) \right]$$

Proof. We set $\eta_z = 0$ in equation (6), with

$$\hat{g}(z) = g(z) - (1 + \eta_z) \frac{g(z) \gamma(z) - t}{p + t} \eta(z) + \eta_z \frac{T'}{1 - T'}$$

to get

$$t = \gamma(g + \sigma) - \frac{p + t}{\tilde{c}} \text{Cov} \left[ g(z) - \frac{g(z) \gamma(z) - t}{p + t} \eta(z), \tilde{\phi}(z) \right]$$

$$= \gamma(g + \sigma) - \frac{p + t}{\tilde{c}} \text{Cov} \left[ g(z), \tilde{\phi}(z) \right] + \frac{1}{\tilde{c}} \text{Cov} \left[ g(z) \tilde{\gamma}(z) - t, \eta(z), \tilde{\phi}(z) \right]$$

Solving equation (32) for $t$ yields the expression in the Corollary above.

D.D Solving the differential equation for the optimal income tax

We now solve the differential equation for the optimal income tax. To that end, define $Y(z) = \int_{z}^{\infty} - \eta_z T'(x) h(x) dx$. Then the equation comes down to

$$\frac{1}{\eta_z h(z)} Y' - \frac{1}{z h(z) \tilde{c}} Y = \frac{1}{z h(z) \tilde{c}} \int_{\gamma(z)}^{\infty} (1 - \hat{g}(x)) h(x) dx + \frac{g \gamma - t}{p + t}$$

or

$$Y'(z) - B(z) Y(z) = B(z) A(z) + D(z)$$

where $A(z) = \int_{z}^{\infty} \left( 1 - g(x) + \frac{g(x) \gamma(x) - t}{p + t} \eta(x) (\eta_z + 1) \right) h(x) dx$ and $B(z) = \frac{\eta_z}{\tilde{c}_z}, D(z) = \frac{g(z) \gamma(z) - t}{p + t} \eta(z) \eta_z h(z)$.

This gives us a first order differential equation, which can be solved for $Y$.
Labor supply misoptimization

It is also possible to relax the assumption that \( U_s/U_c = V_s/V_c \). Labor supply misoptimization modifies the optimal income tax formula to allow to correct for labor supply internalities, as shown in Farhi and Gabaix (2015), Lockwood (2016), and \(? \). Labor supply misoptimization does not alter the optimal commodity tax in two special cases: (1) s consumption differences come from...
preference heterogeneity, or (2) the commodity tax \( t \) is not salient when labor supply is chosen (in this case, the formula from Corollary 5, employing uncompensated demand elasticities, should be used). More generally, the same formulas could be applied under labor supply misoptimization in the presence of income effects, with the exception that \( \hat{g}(z) \)—the welfare impact of giving \( z \)-earners one more unit of post-tax income—must be modified to incorporate the corrective benefit (or cost) from any change in labor supply behavior from this income shock, in addition to the usual fiscal externalities.

### E PanelViews Survey Questions

This appendix presents the text of the PanelViews survey questions used for this project.

#### E.A Self-Reported Beverage Intake

For each of the following types of drinks, please tell us how many 12-ounce servings you drink in an average week. (A normal can of soda is 12 ounces.)

- 100% fruit juice
- Sweetened juice drinks (for example, fruit ades, lemonade, punch, and orange drinks)
- Regular soft drinks (soda pop)
- Diet soft drinks and all other artificially sweetened drinks
- Pre-packaged (i.e. canned or bottled) tea or coffee (for example, iced tea, iced coffee, and flavored tea)
- Sports drinks
- Caffeinated energy drinks

#### E.B Self-control

Please indicate how much each of the following statements reflects how you typically are: I drink soda pop or other sugar-sweetened beverages more often than I should.

- Not at all
- Somewhat
- Mostly
- Definitely
Please indicate how much each of the following statements describes the other head of household: The other head of household in my house drinks soda pop or other sugar-sweetened beverages more often than they should.

- Not at all
- Somewhat
- Mostly
- Definitely
- I am the only head of household

### E.C Preferences: Taste for beverages and health importance

Imagine for a moment that you could drink whatever beverages you want without any health or nutritional considerations. Leaving aside any health or nutrition considerations, how much would you say you like the taste and generally enjoy drinking the following? Please indicate your liking by selecting a number on a scale of 0-10, with 0 being “not at all” and 10 being “very much.”

- 100% fruit juice
- Sweetened juice drinks (for example, fruit ades, lemonade, punch, and orange drinks)
- Regular soft drinks (soda pop)
- Diet soft drinks and all other artificially sweetened drinks
- Pre-made tea or coffee
- Sports drinks
- Caffeinated energy drinks

In general, how important is it to you to stay healthy, for example by maintaining a healthy weight, avoiding diabetes and heart disease, etc.? Please indicate the importance by selecting a number on a scale of 0-10, with 0 being “not at all important” and 10 being “extremely important.”

### E.D Nutrition Knowledge

Now we’d like to ask about nutrition knowledge. This is a survey, not a test. If you don’t know the answer, mark “not sure” rather than guess. Your answers will help identify which dietary advice people find confusing.
Do health experts recommend that people should be eating more, the same amount, or less of the following foods? Please select one response for each. Possible responses: More, Same, Less, Not sure.

- Fruit
- Food and drinks with added sugar
- Vegetables
- Fatty foods
- Processed red meat
- Whole grains
- Salty foods
- Water

Which of these types of fats do experts recommend that people should eat less of? Please select one response for each. Possible responses: Eat less, Not eat less, Not sure.

- Unsaturated fats
- Trans fats
- Saturated fats

Do you think these foods and drinks are typically high or low in added sugar? Please select one response for each. Possible responses: High in added sugar, Low in added sugar, Not sure.

- Diet cola drinks
- Plain yogurt
- Ice cream
- Tomato ketchup
- Melon

Do you think these foods are typically high or low in salt? Please select one response for each. Possible responses: High in salt, Low in salt, Not sure.

- Breakfast cereals
• Frozen vegetables
• Bread
• Baked beans
• Red meat
• Canned soup

Do you think these foods are typically high or low in fiber? Please select one response for each. Possible responses: High in fiber, Low in fiber, Not sure.

• Oats
• Bananas
• White rice
• Eggs
• Potatoes with skin
• Pasta

Do you think these foods are a good source of protein? Please select one response for each. Possible responses: Good source of protein, Not a good source of protein, Not sure.

• Poultry
• Cheese
• Fruit
• Baked beans
• Butter
• Nuts

Which of the following foods do experts count as starchy foods? Please select one response for each. Possible responses: Starchy food, Not a starchy food, Not sure.

• Cheese
• Pasta
• Potatoes
• Nuts
• Plantains

Which is the main type of fat present in each of these foods? Please select one response for each. Possible responses: Polyunsaturated fat, Monounsaturated fat, Saturated fat, Cholesterol, Not sure.

• Olive oil
• Butter
• Sunflower oil
• Eggs

Which of these foods has the most trans-fat? Please select one.

• Biscuits, cakes and pastries
• Fish
• Rapeseed oil
• Eggs
• Not sure

The amount of calcium in a glass of whole milk compared to a glass of skimmed milk is...? Please select one.

• Much higher
• About the same
• Much lower
• Not sure

Which one of the following nutrients has the most calories for the same weight of food? Please select one.

• Sugar
• Starchy
• Fiber/roughage
• Fat
• Not sure

If a person wanted to buy a yogurt at the supermarket, which would have the least sugar/sweetener? Please select one.

• 0% fat cherry yogurt
• Plain yogurt
• Creamy fruit yogurt
• Not sure

If a person wanted soup in a restaurant or cafe, which one would be the lowest fat option? Please select one.

• Mushroom risotto soup (field mushrooms, porcini mushrooms, arborio rice, butter, cream, parsley and cracked black pepper)
• Carrot butternut and spice soup (carrot, butternut squash, sweet potato, cumin, red chilies, coriander seeds and lemon)
• Cream of chicken soup (British chicken, onions, carrots, celery, potatoes, garlic, sage, wheat flour, double cream)
• Not sure

Which of these combinations of vegetables in a salad would give the greatest variety of vitamins and antioxidants? Please select one.

• Lettuce, green peppers and cabbage
• Broccoli, carrot and tomatoes
• Red peppers, tomatoes and lettuce
• Not sure

One healthy way to add flavor to food without adding extra fat or salt is to add...? Please select one.

• Coconut milk
• Herbs
• Soy sauce
Online Appendix

Not sure

Which of these diseases is related to how much sugar people eat? Please select one.

- High blood pressure
- Tooth decay
- Anemia
- Not sure

Which of these diseases is related to how much salt (or sodium) people eat? Please select one.

- Hypothyroidism
- Diabetes
- High blood pressure
- Not sure

Which of these options do experts recommend to prevent heart disease? Please select one.

- Taking nutritional supplements
- Eating less oily fish
- Eating less trans-fats
- Not sure

Which of these options do experts recommend to prevent diabetes? Please select one.

- Eating less refined foods
- Drinking more fruit juice
- Eating more processed meat
- Not sure

Which one of these foods is more likely to raise people’s blood cholesterol? Please select one.

- Eggs
- Vegetable oils
- Animal fat
• Not sure

Which one of these foods is classified as having a high Glycemic Index (Glycemic Index is a measure of the impact of a food on blood sugar levels, thus a high Glycemic Index means a greater rise in blood sugar after eating)? Please select one.

• Wholegrain cereals
• White bread
• Fruit and vegetables
• Not sure

Fiber can decrease the chances of gaining weight. Please select one.

• Agree
• Disagree
• Not sure

If someone has a Body Mass Index (BMI) of 23, what would their weight status be? (BMI is conventionally measured in kg/m²) Please select one.

• Underweight
• Normal weight
• Overweight
• Obese
• Not sure

If someone has a Body Mass Index (BMI) of 31, what would their weight status be? (BMI is conventionally measured in kg/m²) Please select one.

• Underweight
• Normal weight
• Overweight
• Obese
• Not sure
E.E  Other questions

Are you the primary shopper? By this we mean, the one household member who makes the majority of your household’s grocery purchase decisions.

Are you . . .

- Male
- Female

What is your primary occupation?

F  Stockpiling and Optimal Lag Length for Price Elasticity Estimates

Because SSBs are storable, past prices and merchandising conditions that affect past purchases could in theory affect current stockpiles and thus current purchases. To address this, we estimate a version of Equation (12) that includes lags of the price and merchandising condition variables. Letting \( l \) index quarterly lags, we estimate the following regression:

\[
\ln s_{it} = \sum_{l=0}^{L} \zeta_l \ln p_{i,t-l} + \sum_{l=0}^{L} \nu_l f_{i,t-l} + \xi \ln z_{ct} + \omega_t + \mu_{ic} + \varepsilon_{it},
\]

with standard errors clustered by county.

In Appendix Table A1, we find that the local price IV \( Z_{it} \) is a powerful predictor of price paid in quarter \( t \) but is not strongly conditionally correlated with prices paid in the quarters before and after. This implies that we have strong first stages and do not have large serial autocorrelation problems.

A standard approach to determining the optimal number of lags in a distributed lag model is to choose the specification that minimizes the Aikake Information Criterion (AIC) or Bayesian Information Criteria (BIC). In essence, these criteria find the specification that best predicts purchases in period \( t \). As shown in Appendix Table A2, the specification that minimizes AIC and BIC is to set \( L = 0 \), i.e. to include only contemporaneous quarter prices. Furthermore, the coefficients on lagged prices in the table are all statistically zero, implying no statistically detectable stockpiling of SSBs from quarter to quarter.\(^{42}\) For these reasons, we set \( L = 0 \) when estimating demand in the body of the paper, thereby including only contemporaneous prices and merchandising conditions.

---

\(^{42}\)This result is consistent with DellaVigna and Gentzkow (2017), who find that stores offering lower prices in a given week see little decrease in sales in future weeks and months, even for relatively storable goods. Prior work highlighting stockpiling by households (e.g. Hendel and Nevo (2006a; 2006b; 2010), Wang (2015)) typically uses weekly purchase data, and stockpiling is naturally more relevant across weeks than across quarters.
### Table A1: Regressions of Price Paid on the Local Price IV

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price (t)</strong></td>
<td><strong>Price (t-1)</strong></td>
<td><strong>Price (t-2)</strong></td>
<td><strong>Price (t-3)</strong></td>
<td><strong>Price (t-4)</strong></td>
</tr>
<tr>
<td>Local price IV (t)</td>
<td>1.153***</td>
<td>-0.059</td>
<td>-0.251***</td>
<td>0.103</td>
</tr>
<tr>
<td>(0.078)</td>
<td>(0.066)</td>
<td>(0.060)</td>
<td>(0.075)</td>
<td>(0.090)</td>
</tr>
<tr>
<td>Local price IV (t-1)</td>
<td>-0.179***</td>
<td>1.274***</td>
<td>0.037</td>
<td>-0.252***</td>
</tr>
<tr>
<td>(0.060)</td>
<td>(0.082)</td>
<td>(0.068)</td>
<td>(0.070)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Local price IV (t-2)</td>
<td>-0.182**</td>
<td>-0.148**</td>
<td>1.264***</td>
<td>0.045</td>
</tr>
<tr>
<td>(0.073)</td>
<td>(0.061)</td>
<td>(0.094)</td>
<td>(0.077)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>Local price IV (t-3)</td>
<td>-0.002</td>
<td>-0.180***</td>
<td>-0.143***</td>
<td>1.151***</td>
</tr>
<tr>
<td>(0.061)</td>
<td>(0.068)</td>
<td>(0.065)</td>
<td>(0.098)</td>
<td>(0.081)</td>
</tr>
<tr>
<td>Local price IV (t-4)</td>
<td>0.174**</td>
<td>0.048</td>
<td>-0.088</td>
<td>-0.120</td>
</tr>
<tr>
<td>(0.076)</td>
<td>(0.061)</td>
<td>(0.069)</td>
<td>(0.071)</td>
<td>(0.103)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>1,998,451</td>
<td>1,868,165</td>
<td>1,724,091</td>
<td>1,547,098</td>
</tr>
</tbody>
</table>

Notes: This table presents regressions of price paid $\ln p_{i,t} - l$ on the local price instrumental variable $Z_{it}$ and four additional quarterly lags. All regressions include the additional control variables in Equation (12): feature and display (and four lags thereof), natural log of county average per capita income, quarter of sample indicators, and a household-by-county fixed effect. Columns 1-5, respectively, use the contemporaneous price paid and then the first through fourth lags of prices paid as the dependent variable. Sample sizes vary across columns because using a longer lag of price paid $p_{i,t-l}$ reduces the number of observations. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

### Table A2: Determining Optimal Number of Lags

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ln(Average price/liter)</strong></td>
<td>-1.363***</td>
<td>-1.361***</td>
<td>-1.358***</td>
<td>-1.326***</td>
</tr>
<tr>
<td>(0.138)</td>
<td>(0.116)</td>
<td>(0.117)</td>
<td>(0.110)</td>
<td>(0.123)</td>
</tr>
<tr>
<td><strong>ln(Average price/liter) (t-1)</strong></td>
<td>0.047</td>
<td>0.066</td>
<td>0.071</td>
<td>0.043</td>
</tr>
<tr>
<td>(0.119)</td>
<td>(0.118)</td>
<td>(0.107)</td>
<td>(0.116)</td>
<td></td>
</tr>
<tr>
<td><strong>ln(Average price/liter) (t-2)</strong></td>
<td>-0.152</td>
<td>-0.163</td>
<td>-0.151</td>
<td></td>
</tr>
<tr>
<td>(0.113)</td>
<td>(0.106)</td>
<td>(0.117)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ln(Average price/liter) (t-3)</strong></td>
<td>-0.054</td>
<td>-0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.130)</td>
<td>(0.134)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ln(Average price/liter) (t-4)</strong></td>
<td>0.075</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.169)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ln(County income)</strong></td>
<td>0.183**</td>
<td>0.184**</td>
<td>0.181**</td>
<td>0.163*</td>
</tr>
<tr>
<td>(0.090)</td>
<td>(0.089)</td>
<td>(0.088)</td>
<td>(0.087)</td>
<td>(0.087)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>1,103,144</td>
<td>1,103,144</td>
<td>1,103,144</td>
<td>1,103,144</td>
</tr>
<tr>
<td><strong>Akaike Information Criterion</strong></td>
<td>2,528,934</td>
<td>2,526,310</td>
<td>2,523,294</td>
<td>2,499,815</td>
</tr>
<tr>
<td><strong>Bayesian Information Criterion</strong></td>
<td>2,529,542</td>
<td>2,526,882</td>
<td>2,523,830</td>
<td>2,500,315</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (12), with $L = 4, 3, 2, 1, 0$ in columns 1-5, respectively. All regressions include feature and display (including $L$ lags), quarter of sample indicators, and household-by-county fixed effects. All regressions use a common sample: the set of observations in column 1. Observations are weighted for national representativeness. Robust standard errors in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.


G Income Elasticity Estimated with Household-Level Variation

When estimating the income elasticity in Equation (12), our specifications in the body of the paper use county mean income. In this appendix, we present alternative specifications using household income reported by Homescan households on annual surveys. One benefit of the household-level data is that there is substantially more variation to exploit, so the estimates can be more precise.

There are three problems, however. First, one might naturally expect more measurement error in the household-level survey data, especially because we are exploiting within-household variation. Second, using household-level income can generated additional endogeneity problems. In particular, changes in household composition (for example, a divorce or a adult moving in) can affect household income and the quantity of SSBs (and other groceries) purchased. Third, there is some uncertainty over the year for which Homescan panelists are reporting their income. The surveys reported for year $t$ are taken in the fall of year $t - 1$. In the early years of the sample, panelists were asked to report total annual income as of year-end of the previous calendar year, i.e. year $t - 2$. Nielsen believes that panelists are actually reporting their annualized income at the time of the survey, and as of 2011, the instructions for income no longer mention the previous calendar year, but rather annualized income.

Appendix Table A3 explores this third issue by regressing the natural log of household income in year $t$ on the natural log of current and lagged county mean income. The second lag of county income is most predictive, suggesting that the modal Homescan household is reporting their income for that year. However, other lags also have predictive power, suggesting that there will be some measurement error. The low t-statistics on county income also highlight that county income variation is not very predictive of reported household-level income variation.

Appendix Table A4 presents estimates of Equation (12). Column 1 reproduces the primary estimate from column 2 of Table (2). Columns 2-5 substitute household income and the first and second lead of household income in place of county income. Across the various columns, the estimated income elasticities $\hat{\xi}$ range from 0 to about 0.04. This is substantially smaller than the estimate of $\hat{\xi} \approx 0.25$ in the primary estimates reproduced in column 1. One natural explanation is that the difference is attenuation bias driven by measurement error in household income. Even though the magnitudes differ, we emphasize that the income elasticity estimates $\hat{\xi}$ are consistent in their key qualitative implication: they all imply that the steeply declining consumption-income profile is not driven by year-to-year causal income effects.

To address the endogeneity concern introduced above, columns 2-5 all include the full set of demographic controls used elsewhere in the paper: natural logs of education and age, race, an indicator for the presence of children, household size in adult equivalents, employment status, and weekly work hours. Column 6 demonstrates the endogeneity concern by omitting these variables. The estimated income elasticity $\hat{\xi}$ becomes slightly negative.
Table A3: **Regressions of Household Income on County Income**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(County income)</td>
<td>0.060</td>
<td>0.124***</td>
<td>(0.045)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(County income) (t-1)</td>
<td>0.025</td>
<td>0.169***</td>
<td>(0.039)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(County income) (t-2)</td>
<td>0.113***</td>
<td></td>
<td>(0.036)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(County income) (t-3)</td>
<td>0.103**</td>
<td></td>
<td>(0.043)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>590,404</td>
<td>590,404</td>
<td>590,404</td>
<td>590,404</td>
<td>590,404</td>
</tr>
</tbody>
</table>

Notes: This table presents regressions of the natural log of household income on county income and its lags, using household-by-year Homescan data for 2006-2015. All regressions include year indicators and household-by-county fixed effects. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A4: **Estimating Income Elasticity Using Household Income Instead of County Income**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Average price/liter)</td>
<td>-1.438***</td>
<td>-1.253***</td>
<td>-1.402***</td>
<td>-1.296***</td>
<td>-1.250***</td>
<td>-1.425***</td>
</tr>
<tr>
<td>ln(County income)</td>
<td>0.245***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Household income)</td>
<td></td>
<td>0.004</td>
<td>0.015**</td>
<td>-0.056***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Household income) (year+1)</td>
<td></td>
<td>0.041***</td>
<td></td>
<td>0.039***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Household income) (year+2)</td>
<td></td>
<td>0.002</td>
<td></td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>1.074***</td>
<td>1.091***</td>
<td>1.116***</td>
<td>1.119***</td>
<td>1.093***</td>
<td>1.074***</td>
</tr>
<tr>
<td>Display</td>
<td>0.525***</td>
<td>0.485***</td>
<td>0.534***</td>
<td>0.449***</td>
<td>0.486***</td>
<td>0.524***</td>
</tr>
<tr>
<td>Other demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Kleibergen-Paap first stage F stat</td>
<td>180</td>
<td>73</td>
<td>182</td>
<td>95</td>
<td>73</td>
<td>181</td>
</tr>
<tr>
<td>N</td>
<td>1,998,451</td>
<td>1,085,133</td>
<td>1,998,451</td>
<td>1,451,357</td>
<td>1,085,133</td>
<td>1,998,451</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (12). Column 1 presents base IV estimates, while columns 2-6 present estimates using natural log of household income instead of county mean income. All regressions include quarter of sample indicators and household-by-county fixed effects. “Other demographics” are natural logs of education and 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, clustered by county, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.
H Bias Estimation Using PanelViews SSB Consumption Data

As introduced in Section III, the Homescan grocery purchase data are imperfect measures of SSB consumption, both because they do not measure away-from-home SSB consumption that panelists do not scan into the data and because the Homescan data are at the household level, while the bias proxies from PanelViews are at the individual level. For a comprehensive measure of individual-level SSB consumption, we therefore delivered a beverage intake frequency questionnaire as part of the PanelViews survey. In this appendix, we describe these SSB consumption data and use them to estimate average marginal bias.

We used a modified version of the BEVQ-15, a validated questionnaire that is standard in the public health literature (Hedrick et al., 2012). We asked, “For each of the following types of drinks, please tell us how many 12-ounce servings you drink in an average week,” for five types of SSBs (sweetened juice drinks, regular soft drinks, pre-packaged tea or coffee, sports drinks, and caffeinated energy drinks) and two non-SSBs (100% fruit juice and diet soft drinks).

How does Homescan consumption compare to the self-reports? Because average SSB consumption is declining over time in the U.S., we make this comparison using only the most recent year of Homescan data (2015) for closest comparability with the 2017 PanelViews survey. Homescan 2015 purchases average 58 liters per adult equivalent, while annualized PanelViews self-reported intake averages 88 liters, or 58 percent larger. As a benchmark, Kit et al. (2013, Table 2) report that away from home SSB intake is 86 percent larger than intake at home for adults over our sample period, although away from home vs. at home intake is not the only difference between our two consumption measures. Up to some sampling error in the smaller PanelViews sample, Appendix Figure A1 shows that the two data sources are both consistent in showing a similar slope of consumption with respect to income.

At the individual household level, however, the two sources are not always consistent. The relationship between natural log of SSB purchases per adult equivalent in the most recent year of Homescan purchases and natural log of average self-reported SSB consumption per adult in the PanelViews survey has only a 0.16 $R^2$. This relatively low value underscores the importance of having both consumption measures. Because each measure has different strengths (the PanelViews surveys are at the individual level, measure intake, and include consumption away from home, while scanner data are not self-reports), readers may disagree over which measure should be prioritized.

Appendix Table A5 presents estimates of Equation (16) using PanelViews respondent-level data. This parallels Table 3, except that $i$ indexes PanelViews respondents instead of Homescan Households, $s_i$, $b_{ki}$, $b_{si}$, and $a_i$ are the SSB consumption, nutrition knowledge, self-control, and

\[43\text{Only some of this low } R^2 \text{ is from not having self-reports from the full household: when limiting to one-person households, the relationship has a 0.21 } R^2. \text{ Much of the low } R^2 \text{ could be from variation within households over time. The PanelViews self-reports are from October 2017, whereas the most recent year of Homescan data is 2015, and a regression of household-level Homescan SSB purchases on its two year lag has only a 0.50 } R^2.\]
preference measures for respondent $i$, $b_{2si}$ is the self-control of respondent $i$ as rated by the other household head, and $x_i$ are the household-level characteristics for respondent $i$’s household. We cluster standard errors at the household level. Appendix Figures A2 and A3 present the predicted quantity effect of bias and average marginal bias by income, paralleling Figures 8 and 9 in the body of the paper. The patterns of results are similar to in the Homescan data, except that the $\hat{\tau}$ estimates, and the resulting bias estimates, are materially larger.

Figure A1: Homescan Purchases vs. Self-Reported Intake

Notes: This figure presents the average purchases of sugar-sweetened beverages for each household’s most recent year in the Nielsen Homescan data and the average annualized self-reported SSB consumption from the PanelViews survey, by income group. Homescan purchases are measured in liters per “adult equivalent” in a household, 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 A5: **Regressions of Sugar-Sweetened Beverage Consumption on Bias Proxies Using PanelViews Self-Reported Consumption**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutrition knowledge</td>
<td>-1.643</td>
<td>-1.233</td>
<td>-1.064</td>
<td>-0.942</td>
<td>-0.441</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
<td>(0.115)</td>
<td>(0.117)</td>
<td>(0.122)</td>
<td>(0.123)</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.054)</td>
<td>(0.054)</td>
<td>(0.058)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Taste for soda</td>
<td>1.111</td>
<td>1.099</td>
<td>1.112</td>
<td>0.718</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.058)</td>
<td>(0.059)</td>
<td>(0.065)</td>
<td></td>
</tr>
<tr>
<td>Health importance</td>
<td>-0.303</td>
<td>-0.277</td>
<td>-0.407</td>
<td>-0.058</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.099)</td>
<td>(0.104)</td>
<td>(0.105)</td>
<td></td>
</tr>
<tr>
<td>ln(Household income)</td>
<td>-0.092</td>
<td>-0.098</td>
<td>-0.103</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.028)</td>
<td>(0.028)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Years education)</td>
<td>-0.477</td>
<td>-0.383</td>
<td>-0.248</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.149)</td>
<td>(0.150)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other beverage tastes: No, Yes
Other demographics: No, Yes
County indicators: No, Yes

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other beverage tastes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Other demographics</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>County indicators</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R²</td>
<td>0.191</td>
<td>0.316</td>
<td>0.318</td>
<td>0.413</td>
<td>0.413</td>
</tr>
<tr>
<td>N</td>
<td>20,839</td>
<td>20,839</td>
<td>20,839</td>
<td>20,839</td>
<td>20,839</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (16), with different configurations of control variables. Data are at the PanelViews respondent level, and the dependent variable is the natural log of PanelViews self-reported SSB consumption. $b_{ki}$, $b_{si}$, and $a_i$ are the knowledge, self-control, and preference measures for respondent $i$, $b_{2si}$ is the self-control of respondent $i$ as rated by the other household head, and $x_i$ are the household-level characteristics for respondent $i$’s household. “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. 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. Column 5 implements the measurement error correction, using two-sample two-stage least squares with $b_{2si}$ as the endogenous self-control measure and $b_{si}$ as the instrument, with the Murphy and Topel (1986) standard error correction. Observations are weighted for national representativeness. Robust standard errors, clustered by household, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.
Figure A2: **Quantity Effect of Bias by Household Income Using PanelViews Self-Reported Consumption**

Notes: This figure presents the quantity effect of bias $\hat{\tau}(\mathbf{b}^V - \mathbf{b})$, with $\hat{\tau}$ from column 5 of Table 3, using PanelViews self-reported consumption. Observations are weighted for national representativeness.
Figure A3: **Average Marginal Bias by Household Income Using PanelViews Self-Reported Consumption**

Notes: This figure presents average marginal bias by income from Equation (??), with $\hat{\tau}$ from column 5 of Table 3, using PanelViews self-reported consumption. Observations are weighted for national representativeness.
I  Additional Empirical Results

I.A  Price and Income Elasticities

Figure A4: County Per-Capita Income Changes

(a) 2007-2009

(b) 2009-2015

Notes: This figure presents the percent changes in county mean real personal income per capita for 2007-2009 (Panel (a)) and 2009-2015 (Panel (b)). Data are from the Regional Economic Information System (BEA 2017).
### Table A6: Estimates of Price and Income Elasticities to Address Censoring at Zero Consumption

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Average price/liter)</td>
<td>-21.72*** (2.22)</td>
<td>-24.97*** (3.11)</td>
<td>-20.06*** (2.41)</td>
<td>-24.77*** (2.69)</td>
</tr>
<tr>
<td>Local price IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(County income)</td>
<td>2.62 (1.62)</td>
<td>0.07 (1.44)</td>
<td>0.06 (1.28)</td>
<td>-0.07 (1.43)</td>
</tr>
<tr>
<td>Feature</td>
<td>13.63*** (2.34)</td>
<td>19.61*** (2.08)</td>
<td>17.87*** (1.65)</td>
<td>24.32*** (1.91)</td>
</tr>
<tr>
<td>Display</td>
<td>8.29*** (2.30)</td>
<td>8.72*** (2.37)</td>
<td>7.88*** (2.01)</td>
<td>10.24*** (2.31)</td>
</tr>
<tr>
<td>Price elasticity</td>
<td>1.37</td>
<td>1.58</td>
<td>1.27</td>
<td>1.56</td>
</tr>
<tr>
<td>SE(Price elasticity)</td>
<td>0.140</td>
<td>0.196</td>
<td>0.152</td>
<td>0.170</td>
</tr>
<tr>
<td>N</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>2,361,616</td>
<td>2,361,616</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (12). All regressions include quarter of sample indicators and a household-by-county fixed effect. Column 1 presents base IV estimates. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.
Table A7: **Alternative Estimates of Price and Income Elasticities**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Average price/liter)</td>
<td>-1.438***</td>
<td>-1.437***</td>
<td>-1.700***</td>
<td>-1.457***</td>
<td>-1.439***</td>
<td>-1.438***</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.096)</td>
<td>(0.096)</td>
<td>(0.096)</td>
<td>(0.095)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>ln(County income)</td>
<td>0.245***</td>
<td>0.247***</td>
<td>0.262***</td>
<td>0.247***</td>
<td>0.239***</td>
<td>0.245***</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.071)</td>
<td>(0.076)</td>
<td>(0.071)</td>
<td>(0.074)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>Feature</td>
<td>1.074***</td>
<td>1.074***</td>
<td>1.100***</td>
<td>1.078***</td>
<td>1.073***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.076)</td>
<td>(0.076)</td>
<td>(0.076)</td>
<td>(0.076)</td>
<td></td>
</tr>
<tr>
<td>Display</td>
<td>0.525***</td>
<td>0.526***</td>
<td>0.671***</td>
<td>0.521***</td>
<td>0.525***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.099)</td>
<td>(0.099)</td>
<td>(0.103)</td>
<td>(0.099)</td>
<td>(0.099)</td>
<td></td>
</tr>
<tr>
<td>1(Employed)</td>
<td>-0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekly work hours</td>
<td>-0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year × ln(County mean income)</td>
<td></td>
<td></td>
<td></td>
<td>-0.012*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year × ln(Education)</td>
<td></td>
<td></td>
<td></td>
<td>0.000***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kleibergen-Paap first stage F stat</td>
<td>180.5</td>
<td>180.3</td>
<td>206.3</td>
<td>180.6</td>
<td>180.7</td>
<td>180.4</td>
</tr>
<tr>
<td>N</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>1,998,451</td>
<td>1,998,451</td>
</tr>
</tbody>
</table>

Notes: Column 1 reproduces the primary IV estimates of Equation (12) from column 2 of Table 2, while the other columns present alternative estimates. All regressions include quarter of sample indicators and household-by-county fixed effects. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.
Table A8: **Heterogeneous Price Elasticity by Income**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Average price/liter)</td>
<td>-1.717</td>
</tr>
<tr>
<td></td>
<td>(1.331)</td>
</tr>
<tr>
<td>ln(Income) × ln(Average price/liter)</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
</tr>
<tr>
<td>ln(County income)</td>
<td>0.246***</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
</tr>
<tr>
<td>Feature</td>
<td>1.396</td>
</tr>
<tr>
<td></td>
<td>(1.086)</td>
</tr>
<tr>
<td>ln(Income) × Feature</td>
<td>-0.030</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
</tr>
<tr>
<td>Display</td>
<td>0.736</td>
</tr>
<tr>
<td></td>
<td>(1.356)</td>
</tr>
<tr>
<td>ln(Income) × Display</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>(0.127)</td>
</tr>
<tr>
<td>Kleibergen-Paap first stage F stat</td>
<td>30.5</td>
</tr>
<tr>
<td>N</td>
<td>1,998,451</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (12), including additional interactions between natural log of household income ln $w_{it}$ with price paid, county income, feature, and display. The instruments are the local price IV $Z_{it}$ and $Z_{it} \cdot \ln w_{it}$. Observations are weighted for national representativeness. Robust standard errors, clustered by county, are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.
I.B Measuring Bias

Figure A5: **Linearity of Relationship Between Consumption and Bias Proxies**

Notes: This figure presents binned scatterplots of the relationship between natural log of SSB consumption and bias proxies, residual of the other variables in Equation (16). In the top two panels, the dependent variable is natural log of PanelViews self-reported consumption. In the bottom two panels, the dependent variable is the natural log of purchases per adult equivalent in the most recent year that the household was in Homescan.
### Table A9: “First Stage” Regressions to Predict Bias Proxies

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PanelViews Homescan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nutrition knowledge</td>
<td>0.153***</td>
<td>0.131***</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Self-control</td>
<td>0.603***</td>
<td>0.591***</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Taste for soda</td>
<td>-0.120***</td>
<td>-0.093***</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Health importance</td>
<td>0.107***</td>
<td>0.082**</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>ln(Household income)</td>
<td>-0.001</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>ln(Years education)</td>
<td>0.041</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Other beverage tastes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Other demographics</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>County indicators</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.609</td>
<td>0.763</td>
</tr>
<tr>
<td>N</td>
<td>5,588</td>
<td>3,040</td>
</tr>
</tbody>
</table>

Notes: This table presents estimates of Equation (17). Observations are weighted for national representativeness. Robust standard errors are in parentheses. *, **, ***: statistically significant with 90, 95, and 99 percent confidence, respectively.

## J Structural Simulations of Optimal SSB Tax

This appendix presents the optimal estimated SSB tax using a structural model calibrated to the parameters estimated in Section III. This also provides a robustness test of the stability of the sufficient statistic policy estimates reported in Table 5.

### J.A Functional Form Specifications

For the simulations to follow, we employ the following functional forms:

$$U(c, s, z; \theta) = c + a(c)b(\theta) \left( \frac{s^{1-1/\zeta^c(\theta)}}{1 - 1/\zeta^c(\theta)} \right) - \Psi \left( \frac{z}{w(\theta)} \right) \left( \frac{s^{1-1/\zeta^c(\theta)}}{1 - 1/\zeta^c(\theta)} \right)$$

$$V(c, s, z; \theta) = c + a(c)b(\theta) \left( \frac{s^{1-1/\zeta^c(\theta)}}{1 - 1/\zeta^c(\theta)} \right) - \Psi \left( \frac{z}{w(\theta)} \right) - \tilde{\gamma}(\theta)s$$

As in the theory notation, $\zeta^c(\theta)$ denotes the compensated elasticity of SSB demand for type $\theta$, which may vary across individuals. The product $a(c)b(\theta)$ controls the level of soda consumption, and can be calibrated nonparametrically using the observed cross-income profile of SSB consumption. The components $a(c)$ and $b(\theta)$ control whether that cross-sectional variation is driven by income effects.
or preference heterogeneity, respectively. We can nonparametrically calibrate \( a(c) \) to generate the income elasticity estimated from our data, and we attribute the residual cross-sectional SSB variation to preference heterogeneity.

The term \( \tilde{\gamma}(\theta) \) controls the marginal internality from SSB consumption. If some SSB variation is due to income effects (non-constant \( a(c) \)), then experienced utility in equation 35 is not quasilinear in numeraire consumption, and so \( \tilde{\gamma}(\theta) \) is not exactly equal to \( \frac{U_s'}{U_c'} - \frac{V_s'}{V_c'} \). Therefore we use the “tilde” notation to distinguish this structural internality parameter from the equilibrium money metric internality, which we will continue to denote \( \gamma \).

We assume isoelastic disutility of labor effort, \( \Psi(\ell) = \frac{1}{1+1/\zeta_c} \ell^{(1+1/\zeta_c)} \), where \( \zeta_c \) is the Frisch elasticity of labor supply, which is assumed to be homogeneous conditional on income.

To compute the optimal SSB tax, we first calibrate the parameters in equations 34 and 35 to match the patterns of SSB consumption documented in Section III, and other data from the U.S. economy described below. We then use a numerical solver to compute the optimal tax policies under those parameter values.

### J.B Calibration Procedure and Data Sources

We draw from Piketty, Saez, and Zucman (2018) to calibrate the pre- and post-tax income distributions for the United States. We use their reported distribution for 2014, which includes percentiles 1 to 99, and much finer partitions within the 99th percentile. Each of these points in the income distribution is treated as an individual point in a discretized ability grid for the purposes of simulation.\(^\text{44}\) Thus in the simulations there is a one-to-one mapping between type \( \theta \) and ability \( w(\theta) \). We also use these distributions to calibrate effective marginal tax rates in the U.S. status quo, computed as \( 1 - \frac{dy_{us}}{dz_{us}} \), where \( z_{us} \) and \( y_{us} \) represent the reported pre- and post-tax income distributions.\(^\text{45}\)

To encode a preference for redistribution, we assume type-specific Pareto weights given by \( \alpha(\theta) = y_{us}(\theta)^{-\nu} \), where \( \nu \) controls the degree of inequality aversion. Following Saez (2002b), we use a baseline value of \( \nu = 1 \), approximately corresponding to the weights which would arise under logarithmic utility from consumption, and we report alternative specifications with \( \nu = 0.25 \) and \( \nu = 4 \). We compute these weights using the status quo U.S. net income distribution, and they are held fixed during simulations in order to isolate the effects of other model parameters on optimal taxes without endogenously changing distributional motives.

\(^{44}\) We drop the bottom three percentiles of the distribution, which have slightly negative reported incomes due to reported business or farm losses.

\(^{45}\) An alternative approach would be to use estimated effective marginal tax reported, for example from the NBER TAXSIM model or as computed by CBO. However since these estimates often omit some types of taxes, and fail to include many types of benefits, we have chosen to use the more comprehensive notion of taxes and benefits from Piketty et al. (2018).
We calibrate the status quo level of SSB consumption using the consumption estimates across incomes, as shown in Figure A1. To interpolate consumption across the full ability distribution in our simulations, we find the percentile in our income grid which corresponds to each income point in Figure A1, and we then interpolate (and extrapolate) SSB consumption linearly across income percentiles in our ability grid.

The price elasticity of SSB demand, $\zeta^c(\theta)$, is calibrated to match the estimates described in Section III.C. We predict the uncompensated SSB elasticity as a linear function of (log) net income, using the interaction term on elasticity and log income reported in Table A8. We then convert these estimates into the compensated elasticity $\zeta^c(\theta)$ using the Slutsky equation and the estimated SSB income elasticity reported in Table 2.

We jointly calibrate the functions $a(c)$ and $b(\theta)$ nonparametrically to match both the level of SSB consumption (across the income distribution) and the estimated SSB income elasticity, using the following procedure. We first assume $a'(c) = 0$, and we calibrate the product $a(c)b(\theta)$ at each point in the income distribution—without yet worrying about the decomposition into separate terms—to match the observed schedule of SSB consumption via the first-order condition for SSB consumption. We then compute a path of $a'(c)$ which is consistent with our estimated SSB income elasticities, and we numerically integrate to find $a(c)$, which in turn identifies $b(\theta)$ at each ability gridpoint. Finally, we compute the implied ability distribution $w(\theta)$ which generates the observed income distribution. We then repeat this procedure with the new $a'(c)$ (which affects the FOC for SSB demand) and we iterate to convergence. In this manner, we find a non-parametric schedule of $a(c)$ and $b(\theta)$ which is consistent with both the level of estimated soda consumption at each income, and with the estimated SSB income elasticities in our data. During simulations, we compute the nonparametric function $a(c)$ using linear interpolation (extrapolation) over $\log(c)$.

For estimates which assume an externality cost from SSB consumption, we use a value of 0.85 cents per ounce, as calculated in Section III.F.

### J.C Simulation Results

Our estimates for the optimal SSB tax, computed under a variety of specifications, are presented in Table A10. As in Table 5, for each specification we report the optimal SSB tax under two income tax regimes. First, we compute the optimal SSB tax holding fixed the income tax at the current status quo, under the assumption that SSB tax revenues are distributed equally across all consumers as a lump-sum payment. Second, we jointly solve for the optimal income tax, which may be quite different from the U.S. status quo. The former exercise is most relevant if the SSB tax is viewed as an isolated policy over which policymakers much make choices, independent of overall income tax reforms. The latter is more relevant as an illustration of the theoretical results

---

46 We use Homescan consumption as our baseline, though we also report results using self-reported consumption from our PanelViews survey below.
in Section II, which follow the tradition in the optimal taxation literature of jointly characterizing optimal income and commodity taxes. To illustrate the difference between these two regimes, Figure A6 plots the mapping between pre- vs. post-tax income in the U.S. status quo and under the simulated optimal income tax. Finally, we consider the possibility that SSB tax revenues might be distributed in a more limited income-targeted manner—see subsection J.C.2 below.

J.C.1 The Optimal Sweetened Beverage Tax

Estimates of the optimal SSB tax across a variety of specifications are reported in Table A10. The first row presents our baseline estimate of the optimal tax on sugar sweetened beverages under the prevailing U.S. income tax. In this specification, both bias and SSB elasticities are allowed to vary across incomes. (We describe the effect of imposing greater homogeneity along these dimensions below.) If we jointly solve for the optimal income tax, the optimal SSB tax decreases, as shown in the second column. Intuitively, because our estimates reflect positive SSB income elasticities, a higher sin tax reduces the appeal of earning more (since consumers rightly predict that they would then purchase more SSBs and therefore pay more in total taxes). Under the status quo income tax, which is insufficiently progressive relative to our assumed welfare weights, this labor supply distortion is not very costly, since marginal income tax rates are suboptimally low. In contrast, the optimal income tax is more progressive (with much higher marginal tax rates) and as a result the labor supply distortion from a sin tax is more costly, thereby reducing the optimal SSB tax.

Rows 2–5 of Table A10 report the optimal SSB tax under alternative assumptions about redistributive preferences. “Pigouvian” reports the optimal tax computed with no redistributive preferences whatsoever—in which case the optimal SSB tax is simply equal to the average bias of marginal SSB consumption, plus the externality.47 The fact that the Pigouvian optimal tax is higher than the baseline optimal tax from Row 1 indicates that distributional concerns tend to generate regressivity costs which outweigh their amplifying effect on corrective benefits. The specifications with “weak” and “strong” redistributive preferences are computed using Pareto weights of \( \alpha(\theta) = y_{us}(\theta)^{-\nu} \) with \( \nu = 0.25 \) and \( \nu = 4 \), respectively. Row 5 reports the optimal tax under distributional motives computed via the “inverse optimum” approach, wherein welfare weights are calculated as those which rationalize the prevailing U.S. income tax. Since that income tax corresponds to quite modest redistributive preferences, the optimal SSB tax in this case is close to the Pigouvian tax.

Row 6 presents results using self-reported SSB consumption from the PanelViews survey to compute all results, rather than observed SSB consumption measured in the Nielsen HomeScan data.

Rows 7 and 8 report simulations under assumptions of either homogenous bias (row 7) or

---

47Since this specification is computed with no redistributive motive, it is not clear what the optimal income tax would be in such a setting, so we report the optimal tax only in the first column.
homogeneous SSB demand elasticity (Row 8) across incomes. Both are lower than the optimal taxes in the baseline specification, reflecting the extra corrective benefit from a sin tax when bias and elasticity are more extreme among lower income consumers. Row 9 imposes constant bias and elasticity across incomes, and as a result the optimal tax is lower still.

Finally, Rows 10 and 11 report the optimal tax under alternative assumptions about income effects vs. between-income preference heterogeneity. One can understand these two cases as special cases on a spectrum where the (downward-sloping) SSB profile across incomes is decomposed into the sum of (causal) income effects and between-income preference heterogeneity—see Subsection II.B.1 for a discussion of this decomposition. (If SSBs are a normal good, as our empirical estimates suggest, then the true economy does not lie “between” these special cases—rather, the taste for soda declines so quickly with income that it more-than-offsets the rise due to income effects.) Like Row 9, these specifications are computed assuming constant values of bias and elasticity across the income distribution, in order to ensure that in the “Pure income effects” case, consumption differences are driven only by income effects and not (for example) by type-dependent heterogeneity in elasticities. Consistent with the theoretical results in II, under “Pure preference heterogeneity” (Row 10) the optimal SSB tax is the same whether or not the income tax is optimized—the two tax instruments are determined independently. On the other hand, under “Pure income effects” (Row 11) the optimal SSB tax in the presence of the optimal income tax is larger than the Pigouvian benchmark, reflecting that all sin tax regressivity is offset through the income tax, and the tax is equal to corrective benefits, which are amplified by the redistributive motive relative to the Pigouvian benchmark.

J.C.2 How the Optimal Internality-Correcting Tax Varies with Elasticity and Bias

The rows in Table A10 are intended to represent realistic alternative specifications for the purpose of understanding the robustness of our estimates. However, one might also wish to understand how more extreme variations in parameters affect the optimal sin tax. In particular, the analytic formulas in Section II emphasize the role of bias and elasticity of demand in determining the optimal tax. In order to illustrate those insights quantitatively, Figure A7 plots the optimal sin tax across a range of values of bias and elasticities of demand. To isolate the role of the bias and elasticity of demand, the simulations in this figure assume constant values of each parameter across the income distribution (akin to row 10 in Table A10), and zero externality cost. As a result, the optimal tax in the absence of distributional concerns (i.e., the optimal Pigouvian tax) lies on the 45-degree line, and departures from that line illustrate the effect of redistributive concerns on the optimal sin tax. The baseline (average) value of bias is represented in the figure by the vertical dashed line, and the baseline (average) demand elasticity is plotted by the red line. All values are computed holding fixed the status quo U.S. income tax.

Figure A7 illustrates an important theoretical insight from Section II. In particular, the regres-
sivity costs term (which reduces the optimal SSB tax) rises with the demand elasticity. Quantitatively, these costs are modest for elasticities in the range estimated in our data, and the optimal SSB tax lies close to (although distinctly below) the 45 degree line representing the optimal Pigouvian tax. However at lower elasticity values, regressivity costs become large. Indeed at a low elasticity value of 0.25, the optimal SSB tax is in fact a subsidy for the average bias computed in our data.

The Optimal SSB Tax Under Constrained Targeted Revenue Recycling

The simulations in the first column of Table A10, which hold fixed the prevailing U.S. income tax, reflect the possibility that the SSB policy makers may not have control over the nonlinear income tax when setting the SSB tax. In those simulations assume SSB tax revenues are distributed in lump-sum fashion. If instead the SSB policy maker can distribute and collect taxes differentially across incomes, then they can jointly solve for the optimal SSB tax and nonlinear income tax, as in the second column of Table A10. (This could represent the case where local policy makers also set a local income tax, allowing them to correct for any perceived suboptimality of the federal income tax.)

We now consider a third possibility—which may be the most realistic in some settings—that lies between these extremes: SSB policy makers may distribute SSB tax revenues in an income-targeted manner with a non-negativity constraint. That is, they cannot collect more in taxes from any income level than the prevailing income tax. In this case, the marginal social value of SSB tax revenues is equal to the marginal welfare weight of the individuals on whom the targeted revenues are spent (net of any fiscal externalities due to behavioral adjustments).

To capture this possibility succinctly, we recompute the baseline optimal SSB tax across a range of marginal social values of SSB tax revenues. This relationship is plotted in Figure A8, where the horizontal axis represents the social value of SSB tax revenues, relative to a lump sum grant. Thus a value of 1 corresponds to the baseline specification in which SSB tax revenues are distributed evenly across the population. Values to the left of that benchmark imply a lower social value of SSB tax revenues, e.g., if SSB policy makers tend to allocate revenues for political purposes rather than socially valuable projects. Points to the right imply a higher social value of SSB tax revenues—for example, Figure A8 plots a vertical line at the average welfare weight (net of behavioral responses) on the bottom 10% of the population. As the social value increases, the optimal tax levels off, with an asymptote at the revenue-maximizing tax corresponding to the peak of the Laffer curve for SSBs.
Table A10: **Simulation Results: Optimal SSB Tax (cents per ounce)**

<table>
<thead>
<tr>
<th>Case</th>
<th>Existing income tax</th>
<th>Optimal income tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Baseline</td>
<td>1.69</td>
<td>1.30</td>
</tr>
<tr>
<td>2. Pigouvian (no redistributive motive)</td>
<td>1.95</td>
<td>-</td>
</tr>
<tr>
<td>3. Weak redistributive preferences</td>
<td>1.87</td>
<td>1.78</td>
</tr>
<tr>
<td>4. Strong redistributive preferences</td>
<td>1.51</td>
<td>0.40</td>
</tr>
<tr>
<td>5. Inverse optimum redist. prefs</td>
<td>1.92</td>
<td>1.93</td>
</tr>
<tr>
<td>6. Self-reported SSB consumption</td>
<td>2.62</td>
<td>2.28</td>
</tr>
<tr>
<td>7. Constant bias</td>
<td>1.63</td>
<td>1.22</td>
</tr>
<tr>
<td>8. Constant elasticity</td>
<td>1.65</td>
<td>1.26</td>
</tr>
<tr>
<td>9. Constant bias and elasticity</td>
<td>1.59</td>
<td>1.19</td>
</tr>
<tr>
<td>10. Pure preference heterogeneity</td>
<td>1.61</td>
<td>1.61</td>
</tr>
<tr>
<td>11. Pure income effects</td>
<td>1.66</td>
<td>2.04</td>
</tr>
</tbody>
</table>

Notes: The first column displays the optimal SSB tax when the income tax is held fixed at the U.S. status quo; the second column displays the optimal SSB tax when the simulations also solve for the optimal income tax.
Figure A6: **Simulations: Status Quo Income Tax and Optimal Income Tax in Baseline Specification**

![Graph showing simulations of status quo income tax and optimal income tax in baseline specification.](image)

**Notes:** Figure notes in small font. tk Ben.
Figure A7: **Optimal Internality-Correcting Tax Across Values of Bias and Elasticities**

Notes: This figure plots the simulated optimal SSB tax in cents per ounce across values of internality bias (in cents per ounce) for different values of the SSB price elasticity of demand. To illustrate the effects parameters on the optimal internality-correcting tax, these simulations assume zero externality.
Figure A8: **Optimal SSB Tax for Different Marginal Social Values of SSB Tax Revenues**

Notes: This figure shows how the optimal SSB tax varies depending on the social usefulness of SSB tax revenues. A value of one on the horizontal axis implies that SSB tax revenues are valued the same as marginal funds raised via the income tax.