Internalities and the effectiveness of taxing sugar-sweetened beverages*

Paul Gertler
UC Berkeley and NBER

Tadeja Gračner
University of Southern California and RAND

Ricardo Miranda
Duke University

Enrique Seira
Michigan State University and JPAL

This draft: November 6, 2026

Abstract

A number of governments have enacted "sin" taxes on sugar sweetened beverages (SSB) claiming —with weak evidence— that people overconsume them due to internalities. We estimate a structural model of SSB demand that incorporates both rational addiction and present bias. We use high-frequency household panel-data on SSB consumption straddling the 2014 SSB tax in Mexico and city-level heterogeneity in the tax-price pass through to estimate the model. We find evidence of rational addiction, forwarding looking behavior and present bias, providing economic justification for SSB taxes. Consistent with addiction, we show that there are adjacent complementarities in consumption, implying that long-run equilibrium price elasticities of demand are larger than short-run elasticities. Indeed, we estimate a short-run price elasticity of -1.0, and a long-run elasticity that is 50% higher. We benchmark the structural model against a reduced-form RDD estimate and find that the short-run structural elasticity estimate is almost identical to the RDD estimate.

Keywords: Rational Addiction, present bias, Sugary Drinks, Sin-food taxes.
JEL:

PRELIMINARY

*All errors are our own.

1 Introduction

Public health advocates have been promoting taxing sugar-sweetened beverages (SSBs) as a mechanism to reduce sugar consumption and thereby to help prevent and control associated metabolic disease.[1] So far, 7 cities in the U.S. and 39 countries worldwide have imposed a "sin tax" on SSBs (Busey, 2020; Cawley et al., 2019b). While the public health rationale for SSB taxes is to reduce sugar consumption and promote health, the economic rationale for such taxes rests on correcting for internalities such as addiction and present bias (Gruber and Köszegi, 2001; Bernheim and Rangel, 2004; O'Donoghue and Rabin, 2006; Allcott et al., 2019a,b).

Medical evidence supports an important role of addiction in sugar consumption. A number of studies have documented symptoms of sugar addiction, such as craving, binging, and withdrawal, and others find that the neurological reward for sugar intake could be similar to cocaine (Lustig, 2010; DiNicolantonio et al., 2018; Avena et al., 2009, 2008; Hoebel et al., 2009; Avena et al., 2012; Wiss et al., 2018). Moreover, while there is an innate human desire for sweetness, the degree of desire is modified by the amount of sugar one eats (Birch et al. (1990); Jayasinghe et al. (2017); Drewnowski et al. (2012); Sullivan and Brumfield (2016); Thurow (2016)). Gertler and Gracner (2021) show that excess sugar consumption early in life leads to elevated sugar intake across lifespan.

Few studies, however, have explicitly tested for internalities in SSB consumption, leaving policy makers with scant evidence to support government intervention. Moreover, few have accounted for the effect of these biases in the evaluation of the effect of taxes on SSB consumption, or on their influence on estimated price elasticities of demand. In the presence of addiction, long-term price elasticities of demand are larger than the short-term elasticities (Becker and Murphy, 1988; Gruber and Köszegi, 2001). This means that most of the existing estimates of the impact of a tax on SSB consumption likely understate the long-run effectiveness of the tax.

In this study we estimate a structural model of the demand for SSBs that explicitly incorporates both rational addiction and present bias, closely following the theoretical model in Gruber and Köszegi (2001) — GK henceforth.[2] The GK model generates a number of tests for addiction, forward looking behavior and time consistency based on the observed patterns of SSB purchases. We derive a tractable empirical model based on GK and an approach to estimating short-run and long-run price elasticities of demand.

We estimate the model using data from Mexico, an important place to study SSB taxes from a policy perspective.[3] Sugar is a large part of the Mexican diet. At 12.5% of total calories, Mexican sugar consumption is high and grossly exceeds the World Health Organization's nutritional guidelines of less than 5% of total calories. SSBs account for nearly 70% of those sugars and at more than 160 liters of SSBs per person per year, Mexico is the largest consumer of SSBs worldwide. To reduce SSB consumption, Mexico passed an excise

[1] https://tinyurl.com/y6ymx7ze, https://tinyurl.com/y4pa3hoh.

[2] Gruber and Köszegi (2001) incorporate present bias into the rational addiction model of Becker and Murphy (1988) and use that model to derive optimal tax policy for tobacco.

[3] Seventy percent of adults and 30% of children are overweight or obese (Rivera et al., 2014; Barquera et al., 2013; Rtveladze et al., 2014; Colchero et al., 2017).

tax of 1 peso (6 US cents) per liter on all non-alcoholic beverages with added sugars in the fall of 2013, and implemented it shortly thereafter starting on January 1, 2014 (Aguilar et al., 2021). The tax amounted to a 12% increase in the average pre-tax SSB price. The 2014 tax is relatively high – about three times that of the average soda tax in the US – and has a broad base applying to 46.3% of all beverage products.[4]

We use high quality household scanner panel data collected between January 2013 and December 2015 from the Mexican Kantar World Panel to estimate the model. We exploit city-level heterogeneity in the tax price pass-through across 89 cities; the tax pass through ranged from 6% to 19% across cities. We instrument for the tax price pass-through using city-level baseline local grocery store market structure excluding SSBs that reflects underlying supply elasticities and control for SSB demand preferences using household fixed effects.

We report two main results. First, there is strong evidence of addiction, forward-looking behavior, and rejection of time consistency in favor of present bias. Second, we estimate a short-run price elasticity of about -1.0 and a long-run price elasticity that is approximately 50% larger. These results have important policy implications, as they provide a rationale for taxing SSBs and imply that SSB taxes are more effective at reducing consumption than previously thought.

We benchmark the price elasticities derived from the structural model against price elasticities estimated using a regression discontinuity design (RDD), a common program evaluation method used in the SSB tax literature, and a naive demand function that ignores internatities. These approaches yields estimate almost identical to the short-run price elasticities estimated from the structural model. Consumption responses to SSBs taxes reported in the literature vary across settings, but are generally similar to our our short-run demand price elasticity estimate - ranging for instance from -0.6 in France to -1.0 in Philadelphia, PA (Zhong et al., 2018; Cawley et al., 2019a; Capacci et al., 2019; Silver et al., 2017; Lee et al., 2019; Colchero et al., 2016; Falbe et al., 2016).[5] There are also a number of observational studies that use cross-sectional variation in prices to estimate price elasticities, finding elasticities ranging from -0.8 to -1.0 (Andreyeva et al., 2010; Powell and Chaloupka, 2009; Smith, 2010; Powell et al., 2013; Gil, 2006; Barquera et al., 2008; Urzúa, 2008; Unar et al., 2013; Colchero et al., 2015a).

Allcott et al. (2019a) is the only other study that we know of that tests for behavioral biases in the consumption of SSBs. They take a very different approach. They estimate the relationship between SSB consumption and the respondent's agreement with the following question: *"I drink soda pop or other sugar-sweetened beverages more often than I should"*. They calculate excessive consumption by comparing the consumption for those who agree vs those who disagree, while controlling for a rich set of observables. We instead use preferences revealed in inter-temporal consumption together with an actual SSB tax increase to conduct tests for forward looking behavior, rational addition, and present bias.

[4]Taxed drinks include sodas, nectars, concentrates with added sugar, flavored water, energy drinks, powdered drink mixes, and sweetened dairy products that are not milk. Sugar-free (diet) sodas, dairy products (even if they include added sugars), juices with no added sugar, and alcoholic beverages are exempt from this tax reform.

[5](Aguilar et al., 2021) examine the effect of 2014 Mexican SSB tax on nutrition including total calorie and macronutrient consumption from all sources as opposed to SSB consumption to account for substituion patterns in the response to the tax.

2 Theory and Identification

Here we describe a theoretical framework developed by GK who incorporate present bias into the rational addiction model of Becker and Murphy (1988). We then use this framework to derive an empirically tractable model of the demand for SSBs, and to develop strategies for identifying the presence of internalities as well as estimating price elasticities.

2.1 Model

The main contribution of GK is to incorporate Laibson (1997)'s representation of present bias in the form of quasi-hyperbolic discounting into Becker and Murphy (1988)'s model of rational addiction. In this model individuals consume two goods, an addictive one a and a non-addictive one c. Whereas utility from c is just a function of contemporaneous consumption, addiction enters the model by allowing the utility from a to depend on both on current consumption as well as past levels of consumption. This way, an agent who has consumed more of a in the past will have higher utility from consuming the good today.

Becker and Murphy (1988) parametarize these preferences with a utility function that is linear in c, and quadratic in a and the stock of the past consumption of a, S:

$$U_t(c_t, a_t, S_t) = \alpha_a a_t + \alpha_s S_t + \frac{\alpha_{aa}}{2} a_t^2 + \alpha_{as} a_t S_t + \frac{\alpha_{ss}}{2} S_t^2 + \alpha_c c_t \tag{1}$$

Addiction is captured by α_{as}, which describes how addictive stock S affects current enjoyment of a. While S depends on the full history of consumption, it depreciates from one period to the next at a rate of d:

$$S_{t+1} = (1 - d)(S_t + a_t) \tag{2}$$

To allow for present bias, GK introduce a more general inter-temporal discount function into the agents' present value of utility at time t:

$$U_t + \beta \sum_{i=1}^{T-1} \delta^i U_{t+i}, \tag{3}$$

where δ is the usual discount rate and β biases down the contribution of future utility – i.e. present bias.

GK show that the inter-temporal maximization problem has a solution for time-consistent, naifs and sophisticates agents,[6] and that the solutions converge quickly[7] to their steady state values. The solution can be expressed as a linear function of two terms: (1) an addiction mechanism that depends on past consumption but not on prices and (2) one that depends on prices but not on past consumption:

[6] Time-consistent agents are those whose consumption decisions chosen at the starting point of the model remain optimal from the viewpoint of all subsequent periods due to the fact that their beliefs about their own future preferences are correct. On the other hand, in every period, the preferences of naive agents for present consumption are stronger than they expected. This discrepancy causes a deviation from previously chosen consumption plans. Sophisticate agents have the same preferences than naifs, however, they choose consumption in a way that it is not optimal for their future selves to deviate from them. In practice, the optimization problem that sophisticates face is modeled as a subgame-perfect equilibrium between the agents' present and future selves.

[7] *i.e.* far before the time horizon of the model is reached. See GK for reference.

$$a_t = \lambda_t S_t + \mu_t \tag{4}$$

The first term is the addiction mechanism and λ can be interpreted as the marginal propensity to addiction. While both λ_t and μ_t depend on preferences (parameters of the utility function), current and future prices only enter through μ_t[8].

2.2 A Tractable Specification

To estimate this model we express equation 4 in terms of observable quantities through the following steps. First, we approximate λ_t by its steady state value λ[9]. Second, we apply the law of motion from equation 2 to replace λS_t with $\sum \lambda(1-d)^k a_{t-k}$. Third, GK show that μ_t can be expressed linearly in terms of a constant term η, current and future prices, plus a future value of its own sequence μ_{t+M}. Fourth, the series $(1-d)^k$ converges to zero as k trends to infinity, implying that the effect of S can be approximated with arbitrary precision using a finite number of lags of a. Similarly, the convergence of μ_t ensures that the coefficient for μ_{t+M} is decreasing in M and can also be approximated by a finite number of price leads.

Thus, equation 5 can be specified as:

$$a_t = \eta + \sum_{k=1}^{N} \lambda(1-d)^k a_{t-k} + \sum_{l=0}^{M} \gamma_l P_{t+l} + \kappa \mu_{t+M} \tag{5}$$

and the model we take to the data, then, is:

$$a_{it} = \eta_i + \sum_{k=1}^{N} \beta_k a_{i,t-k} + \sum_{l=0}^{M} \gamma_l P_{i,t+l} + \varepsilon_{it} \tag{6}$$

where a_{it} is the consumption of a by agent i at time t, N is the number of lags of consumption, P_{it} is the price of a expected by agent i at time t, M is the number of price leads, and $\beta_k = \lambda(1-d)^k a_{t-k}$. An individual fixed effect η_i captures time invariant preferences for SSBs and ε_{it} a random taste shock.[10]

2.3 Internalities

2.3.1 Addiction

The defining characteristic of an addictive good is that past levels of consumption increase its enjoyment today, i.e. adjacent complementary. Specifically, if a positive association between past and current

[8] This is true when utility from consumption of the non-addictive good is linear. GK assume this in the results section of their model. We will assume it throughout the whole paper.

[9] Note that the convergence results proved by GK allow us to assume that the parameter λ_t reaches its steady state values far beyond the time horizon of the model. Moreover, λ is a function of the parameters of the immediate and utility and discount factors but not of prices. This implies that price changes will not affect the convergence nor the limit of the series λ_t

[10] The effect of past consumption decreases exponentially over time at a rate $(1-d)$, implying that a finite number of lags of a can approximate with arbitrary precision the effect S on demand. If a sufficient number of lags of consumption of a are included, then the effect of the omitted elements will be statistically indistinguishable from zero. A similar argument applies to prices: convergence of the model to a steady state ensures that prices far into the future have small effects on current consumption. Furthermore, if consumers have limited forecasting ability, then, price changes far into the future should be irrelevant for contemporaneous consumption.

consumption (β_k>0) is observed in equation 6, we can reject the hypothesis that there is no evidence of addictive behavior.

Note that equation 5 shows that the link between past and current consumption depends on two parameters: the marginal propensity to addiction (λ) and the depreciation rate of the addictive stock, $(1-d)$. Because the coefficients β_1 and β_2 for the first and the second lag of consumption of the addictive good are estimates for $\lambda(1-d)$ and $\lambda(1-d)^2$ respectively, both λ and d can be recovered.

2.3.2 Forward Looking-Behavior

Forward-looking behavior means that consumption decisions today are affected by consumers' beliefs about future prices. The simplest test, then, for forward looking behavior is the significance of the γ_k (for $k > 1$) coefficients. If they are significant, we cannot reject that expectations about future prices affect present consumption.

Forward looking behavior can arise a number of different ways. If consumers know that SSBs are addictive, they may also be aware that an increase in consumption today affects their utility from consuming SSBs in the future, and they might take this into account by maximizing their discounted present value of utility instead of just current utility. Therefore, the cost of increasing consumption of the addictive good is not only immediate but also is incurred throughout all future periods which means the individual will incur in higher costs in every subsequent period. This mechanism induces a negative association between future prices and contemporaneous consumption.

Forward-looking behavior can also give rise to stockpiling in anticipation that prices will increase in the future. Stockpiling however would imply that current consumption would *rise* when prices are expected to increase. This effect goes in the opposite direction of what we would expect for consumption of an addictive good. Moreover, Aguilar et al. (2021) test and do not find evidence of stockpiling of SSBs as a result of the 2014 Mexican tax.

In our setting, we are dealing with a tax to SSBs which induced a permanent increase in prices, For this reason, testing how consumers react to permanent price changes might be a more appropriate test of forward looking behavior. In equation 6, the γ coefficients that correspond to prices are interpreted as the effects of current or future price changes on current consumption. However, the effect of a permanent price is the sum of all these individual coefficients. Thus, to measure and test how a permanent price change expected $n \geq 0$ periods into the future impacts a_t we need to add the effect of all price effects starting from p_{t+n}.

Formally, if $\Delta_m a_t$ is the change in a_t when prices are expected to change m periods into the future and our regression includes up to M leads of prices, we use the following approximation to obtain the direct effect of future expected and contemporaneous realized price changes on current consumption:

$$\Delta_m a_t \approx \sum_{l=m}^{M} \gamma_l \tag{7}$$

This yields estimates for the change in contemporaneous consumption when price changes are realized in the present or expected m periods into the future.

2.3.3 Price elasticities

If the addictive nature of sugar affects demand for SSBs, then current consumption affects future consumption through the accumulated addictive stock. This implies that price increases have two effects, one is an immediate reduction in demand that occurs through the direct effect of prices and remains throughout all subsequent periods. The second comes from the reduction in the addictive stock that the immediate effect causes. This initial reduction will start a convergence process to a new (lower) steady state level of the addictive stock. From adding these two effects (*i.e.* by measuring the total effect of the price change on steady state consumption) we obtain the long term effect of a price increase on consumption.

To do so, we start from the law of motion of S_t described by equation 2:

$$S_t = (1-d)(S_{t-1} + a_{t-1}) \tag{8}$$

This implies that in, in steady state:

$$S^* = \frac{(1-d)}{d} a^* \tag{9}$$

Plugging equation 9 into equation 4 yields:

$$a^* = \lambda \frac{(1-d)}{d} a^* + \mu \tag{10}$$

Which can be arranged as:

$$a^* = \frac{1}{1 - \lambda \frac{(1-d)}{d}} \mu \tag{11}$$

Recall from equation 4 that the effect of prices on current consumption corresponds to the change in μ with respect to prices. Therefore, if there is a permanent price increase today, the change in steady state consumption will be $\frac{1}{1-\lambda\frac{(1-d)}{d}}$ times the immediate impact of a contemporaneous permanent price increase on a_t. Hence, from our estimates for the immediate change in consumption we can recover the long term effect of a permanent price increase using the following equation:

$$\Delta a^* \approx \frac{\sum_{l=0}^{M} \gamma_l}{1 - \hat{\lambda}\frac{(1-d)}{d}} = \frac{\Delta_0 a_t}{1 - \hat{\lambda}\frac{(1-d)}{d}} \tag{12}$$

Equation 12 provides a test for the null hypothesis that price effects are equal in the short and long run. The term, $\frac{1}{1-\lambda\frac{(1-d)}{d}}$ can be expressed non-linearly as a function of the coefficients obtained from estimating equation 6 as $\frac{1}{1-\frac{\beta_1^2}{\beta_1-\beta_2}}$. Thus, using the delta method we can test if the effects of the tax were significantly larger in the long run (*i.e.* if $\theta := \frac{1}{1-\frac{\beta_1^2}{\beta_1-\beta_2}}$ is equal to one. If $\lambda = 0$ or if d=1 there is no addiction, and θ equals

one).

From the short and long term effects of the tax on SSBs consumption we can recover the short and long term elasticity of demand with respect to permanent price changes. To do so, we simply multiply our estimates for the immediate and steady-state changes in consumption by the average observed pre-tax prices and divide them by the average pre-tax consumption of SSBs. Formally:

$$\mathcal{E}_{p,s} = \Delta_0 a \frac{\bar{P}}{\bar{Q}} \tag{13}$$

$$\mathcal{E}_{p,l} = \Delta a^* \frac{\bar{P}}{\bar{Q}} \tag{14}$$

where $\mathcal{E}_{p,s}$ and $\mathcal{E}_{p,l}$ denote the short and long run elasticities respectiveley. The absolute immediate change in consumption is denoted $\Delta_0 a$, the change in steady-state consumption is denoted Δa^* and the average pre-tax prices and consumption of SSBs are denoted $\bar{P}$ and $\bar{Q}$. We follow this procedure both assuming that consumers are forward looking and assuming that they are myopic. Note that these elasticities can be recovered from the structural model's estimates.

2.3.4 Present Bias

If consumers are time-consistent exponential discounters, they have the ability to adhere to optimal consumption paths that internalize addiction. Furthermore, their inter temporal trade-offs are described by a single parameter, δ. This is also true for quasi-hyperbolic discounters in all periods except the first. To test for time consistency, we solve for the parameter δ using the ratio of the response of consumers to price changes that occur today and one period into the future. Then, we ask if the ratios observed between responses to consecutive future price changes are consistent with the obtained parameter.

To perform this test we follow GK who show that for time consistent agents, the ratio between the response of consumers to a price increase two versus one period into the future is equal to $(1-d)\delta(1+\lambda)$ whereas for both types of quasi-hyperbolic discounters this ratio is strictly larger than $(1-d)\delta(1+\lambda)$.

To implement the test we use our estimates for the response of consumers to contemporaneous and future price changes obtained according to section 2.2.3. We also use our estimates for λ and d described in section 2.2.1. In addition, we estimate the parameter δ under the null of time consistency using the following expression derived from dividing the response of consumers to an immediate price change by the the response to a price change one period into the future:

$$\delta = \frac{1}{\left(\frac{\Delta_0 a_t}{\Delta_1 a_t}\lambda + 1\right)(1-d)} \tag{15}$$

Note that our estimates for all three parameters d, λ and δ are non-linear transformations of our main regression coefficients obtained from equation 19 detailed below. Moreover, we can also express the ratio

between the response of consumers to a price change two versus one periods into the future as a non linear function of said coefficients. Simplifying the resulting expression for $(1-d)\delta(1+\lambda)$, our test of time-consistency can be performed using the delta-method in a one-sided test with null the hypothesis that:

$$\mathbf{H_0}: \frac{\Delta_2 a_t}{\Delta_1 a_t} = \frac{(\beta_2 + \beta_1^2)\Delta_1 a_t}{\beta_1^2 \Delta_0 a_t + \beta_2 \Delta_1 a_t} \tag{16}$$

Against the alternative hypothesis:

$$\mathbf{H_1}: \frac{\Delta_2 a_t}{\Delta_1 a_t} > \frac{(\beta_2 + \beta_1^2)\Delta_1 a_t}{\beta_1^2 \Delta_0 a_t + \beta_2 \Delta_1 a_t} \tag{17}$$

3 Data and measurement

We use data on household-level purchases of SSBs collected in the Kantar World Panel (KWP). Similar to the well-known Nielsen consumer panel, individuals in the KWP are asked to keep receipts for all weekly purchases made in establishments, including but not limited to super-centers, grocery or convenience stores. These receipts are then scanned weekly by surveyors who visit these households.[11] Receipt information includes each product's price and barcode, which provides information on beverage brand, volume (in liters), as well as classification into those that are carbonated vs not, or with vs without added sugars. KWP data also records information on date of purchase, number of units purchased, and store location, type and name.[12] Aguilar et al. (2014) show that the distribution of KWP data on SSB consumption closely matches SSB consumption data from Mexico's 2014 Income and Expenditure Survey (ENIGH) and the 2012 Health and Nutrition Survey (ENSANUT), Mexico's main reference economic and health surveys. KWP also records information on household socioeconomic and demographic characteristics.

Table 1 assesses the representatives of our sample by comparing consumption of SSBs and demographics of the households in our final sample with the national representative sample of ENSANUT. The pre-tax consumption levels and demographic of our sample matches the ones observed in urban households surveyed by ENSANUT.

Our data consist of bi-monthly purchases of SSBs between January 2012 and December 2015. SSBs include all tax-eligible beverages (2,560 barcodes), such as sodas, nectars, energy drinks, flavored waters, concentrates, and powdered drink mixes with added sugar. They represent 46.3% of all beverage products in KWP. Our sample consist of a balanced panel of households for whom SSBs purchase receipts are scanned at least once each month during this period. This amounts to 4,022 households in 74 cities and 48,264 household-bimonthly observations in total. Table 1 shows summary statistics of our sample.

Our primary outcome variable is the number of liters of taxable SSBs purchased by a household over a two-month period. We construct a SSB price index for each household in two steps. First, we calculate the

[11] The KWP records purchases in-home, which do not necessarily measure intake.
[12] See Aguilar et al. (2016) for more data details.

bi-monthly average "transaction" price per liter of every sugar sweetened beverage barcode b in each city c in period t, that is, $\bar{P_{bct}}$. Then, for each household, we compute a bi-monthly weighted average price of SSBs using weights that are proportional to household's volume of each purchased SSB barcode before tax was implemented in 2013. Formally, the SSBs price that household h faced in city c at time t is defined as:

$$P_{htc} = \frac{\sum_{b \in SSB} q_{hb} l_b \bar{P_{bct}}}{\sum_{b \in SSB} q_{hb} l_b} \tag{18}$$

where q_{hb} denotes the units of barcode $b \in SSB$ purchased by household h during 2013, l_b the liters of purchased barcode b and $\bar{P_{bct}}$ an average transaction price of barcode b across two months t in city c.

4 Empirical Model

We estimate the following version of equation 6:

$$a_{th} = \eta_h + \sum_{k=1}^{N} \beta_k a_{h,t-k} + \sum_{l=0}^{M} \gamma_l P_{h,t+l} + \omega_c t + \omega_n t + \varepsilon_{ht} \tag{19}$$

where a_{th} is consumption of SSBs in period t for household h, $a_{h,t-k}$ is consumption of SSBs k periods prior (for $k = 0, 1, 2$), $P_{h,t+l}$ is the price per liter that household h expects to pay l periods after t (with $l = 0, 1, 2, 3$), η_i is a household fixed effect that controls for time invariant tastes and preferences as well as for baseline addictive stocks, and ε_{ht} is random error interpreted as a shock to household h's preferences in period t. We also adjust for differential trends in consumption across cities ($\omega_c t$) and socio-economic status ($\omega_n t$). Our main specification sets $N = 2$ and $M = 2, 3$.

4.1 Identification

Our goal is to identify the parameters that determine how consumers respond to a permanent price increase from the moment it is announced until it actually happens. As in a standard static model, endogeneity between prices and consumption is a potential concern. Conditional on household fixed effects and time trends, the timing of the tax to SSBS is unlikely to be correlated with time varying unobserved determinants of SSBs consumption. Hence, the panel structure of our data allows us to use its announcement and enactment as sources of permanent variation in realized and forecast prices.

Our identification strategy exploits city-level heterogeneity in the SSB tax-price pass-through. While the tax led to a 9% average price increase starting in January 2014 that persisted thereafter,[13] the price tax pass-through varied greatly across cities. Figure 1 pictures the coefficients obtained from regressing the log-transformed SSBs prices on the interactions between city indicators and an indicator for when the tax was

[13]Several other studies have observed immediate and persistent increase in SSBs prices soon after the tax in Mexico ranging from 8-13 percent on average (Aguilar et al. (2016), Capacci et al. (2019), Cawley (2015), Silver et al. (2017), Colchero et al. (2017), Zhong et al. (2018)), Aguilar et al. (2021).

implemented.[14] SSBs price changes ranged from 6% to 19%. This observed variation in tax pass-through is consistent with theory, which suggests that a tax will under-shift or over-shift prices depending on local market primitives, i.e. demand and supply elasticities (Stern (1987), Delipalla and Keen (1992), Anderson et al. (2001), Weyl and Fabinger (2013)). In competitive markets, for instance, the shift in prices is bounded between zero and the amount of the tax depending on how inelastic demand is relative to supply. Evidence suggests that the lower the elasticity of demand in each market, the higher the level of shifting. In less or non-competitive markets, price adjustments may rise by any amount relative to the tax (Hamilton, 1999; Fullerton and Metcalf, 2002; Delipalla and Keen, 1992; Weyl and Fabinger, 2013).

We consider two sets of instruments that both exploit geographic variation in the tax-price pass-through. The first set is comprised of interactions between the indicator for tax implementation in 2014 and city indicators. Household fixed effects control for time-invariant SSBs determinants of demand such as tastes that result from location or demographics as well as for unobserved supply characteristics like distribution costs. The tax-city interactions capture the average local differences in prices before and after enactment. Our identifying assumption is that these price differences are uncorrelated with time variant idiosyncratic city-level demand shocks. This is true if the local tax effect is determined only by static or stationary features of the market and there is no remaining differential post-tax variation at the city level.

Using this set of instruments, the first stage model for current prices at time t is:

$$p_{htcrn} = \eta_h + \tau_c \mathbb{I}_{y(t)>2013,c} + \omega_c t + \omega_n t + \upsilon_{ht} \tag{20}$$

where p_{htcrn} is the price per liter faced by household h of socio-economic status n located at city c in region r during period t that falls in year $y(t)$. Time invariant household characteristics are represented by η_h. The function $\mathbb{I}_{y(t)>2013,c}$ indicates tax enactment at the city-level. We include linear trends that are allowed to vary by city and socio-economic status (SES) $\omega_c t$ and $\omega_n t$. [15] $(\omega_n t)$. The term υ_{ht} represents the remaining unexplained variation in prices.

The second set of instruments is comprised of multiple pre-tax measures of local grocery market competition or concentration. Our choice of instruments is informed by prior work studying the pass-through of SSBs and junk food taxes in Mexico, which found empirical evidence that with greater competition there was less overall tax pass-through (Campos-Vázquez and Medina-Cortina, 2019), and that price increases by leading firms or stores were larger than increases at the market level or at the smaller retailers (Salgado and Ng, 2019; Gracner et al., 2021).

Using Kantar dataset, we first calculate market shares for ten largest retailers based on their food sales data at both the city and regional level. We then use that information to calculate the HH indices at both the

[14] The specification is as follows: $lnP_{hct} = \eta_h + \tau_c \mathbb{I}_{y(t)>2013} \mathbb{I}_c + \omega_c t + \upsilon_{ht}$. Where P_{hcit} is the price faced by household h in city c at time t. $\mathbb{I}_{y(t)>2013}$ indicates the tax enactment and $\mathbb{I}_c$ indicates that household h is located in city c. $\tau_c t$ denotes the average price increase in prices observed in city c. $\omega_c t$ denotes a city trend and υ_{ht} is the error term. The coefficients of interest are each of the 74 ω_c that correspond to the 74 cities in our sample.

[15] Socio-economic status is calculated by Kantar according to an index that takes into account households' assets and socio-economic characteristics.

city and regional level. We construct all measures of local market concentration at baseline (i.e., pre-2014), and exclude sales of SSBs from our calculations so that they are independent of shocks to SSBs demand. Identification under this set of intruments only requires our pre-tax measures of market concentration to be uncorrelated with time varying unobserved demand shocks conditional on linear local time trends.

Under this approach, the first stage is specified as:

$$p_{htcrn} = \eta_h + \theta_c \mathbb{I}_{c,y(t)>2013} M_c + \omega_c t + \omega_n t + \upsilon_{ht} \tag{21}$$

Where the dependent variable p_{htcrn} is the price per liter of SSBs faced by household h of socio-economic status n located at city c in region r during period t that falls in year $y(t)$. The tax enactment in city c is indicated by $\mathbb{I}_{c,y(t)>2013}$, and is interacted with M_c, which is a vector of statistics that depend on the composition of the retail market (M) in city c. Time invariant household and city characteristics are captured by η_h. We include city-specific ($\omega_c t$) and SES-specific ($\omega_n t$) trends. The remaining variation in prices is represented by ν_{ht}.

4.1.1 Instrument selection

We use a data-driven approach to obtain a sparse set of instruments that approximates the optimal instrument by capturing the best links between market concentration and the pass-through of the tax. Specifically, we employ a machine learning post-regularization procedure for linear models with multiple controls and instruments proposed by Chernozhukov et al. (2015) and Ahrens et al. (2018) (often referred to as post-double selection) using a LASSO regression approach.

The post-regularization method is ideal for our setting for several reasons. The interactions of the tax yield a large set of variables that correlate with prices. Similarly, we have a wide array of potential controls (e.g., city and socio-economic status trends). It is likely that some of the instruments and controls are collinear or redundant. Furthermore, we ignore *a priori* the identities of the instruments that better predict how the tax was passed through and allow for efficient second stage estimators. For instance, for the fist set of instruments (i.e., the interactions of the tax enactment with city indicators), we might want to drop the coefficients that correspond to cities in which the observed effect was not different from the national average. Additionally, it is not obvious that the observed cross-sectional variation in the tax pass-through is due to market characteristics varying at the city level instead of at a broader level of aggregation such as region. To this end, this data-driven approach is helpful because there is no theoretical or institutional way to identify the most appropriate market concentration measure. [16] Note that we do not penalize household fixed effects in any of these regressions, because they play a crucial role in our identification assumptions.

For our estimation based on city indicators, we estimate equation 20 using LASSO regression with data driven penalties to select cities where the tax shifted prices the most as well as cities where prices exhibit linear trends. The objective function of the problem is:

[16] We implement this method using the Stata package *ivlasso* developed by Ahrens et al. (2018)

$$\sum_{h\in H}\sum_{t\in T}\Big(p_{htcrn}-\eta_h-\tau_c\mathbb{I}_{y(t)>2013,c}-\omega_c t-\omega_n t\Big)^2+||\Lambda_1(\tau_c,\omega_x,\omega_n)|| \tag{22}$$

Where H is the set of all households in our sample and T all our observed time periods. The dependent variable p_{htcrn} is the price per liter faced by household h of socio-economic status n located at city c in region r during period t that falls in year $y(t)$. Time invariant household characteristics are represented by η_h. The function $\mathbb{I}_{y(t)>2013,c}$ indicates tax enactment at the city-level. We include linear trends that are allowed to vary by city and socio-economic status (SES) $\omega_c t$ and $\omega_n t$ ($\omega_n t$). The last term penalizes large coefficients, Λ_1 is a diagonal matrix with data-driven parameter specific penalty loadings.

In the context of equation 23, post-double selection approach helps us identify a valid set of instruments (i.e., out of Herfindahl-Hirschmann indices and market shares of the 10 largest retailers, interacted with an indicator for tax enactment) that are able to explain the effect of the market structure M_c on the price change observed in city c denoted τ_c. We seek to minimize the following expression:

$$\sum_{h\in H}\sum_{t\in T}\Big(p_{htcrn}-\eta_h-\theta_r\mathbb{I}_{r,y(t)>2013}h(M_c)-\theta_c\mathbb{I}_{c,y(t)>2013}h(M_c)-\omega_c t-\omega_n t\Big)^2+||\Lambda_2(\theta_c,\theta_r)|| \tag{23}$$

Where H is the set of all households in our sample and T all our observed time periods. The dependent variable p_{htcrn} is the price per liter of SSBs faced by household h of socio-economic status n located at city c in region r during period t that falls in year $y(t)$. The tax enactment in city c is indicated by $\mathbb{I}_{c,y(t)>2013}$ and in region r is indicated by $\mathbb{I}_{r,y(t)>2013}$. Indicators are interacted with M_c and M_r, which are vectors of statistics that depend on the composition of the retail market (M) at the city and region level. Time invariant household and city characteristics are captured by η_h. We include city-specific ($\omega_c t$) and SES-specific trends ($\omega_n t$). The LASSO regression penalizes large coefficients on the components of both M_c and M_r (*i.e.* $theta_r$ *and* $theta_c$). The penalties are variable specific and represented by the diagonal matrix Λ_2. Note that equations 20 and 23 share the fixed effects, trends and the main effect of the tax, but they do not share any of the tax indicator interactions.

4.1.2 Future prices

In the context of addiction, current consumption affects marginal utility in the future, making current consumption a function of future prices. In the model of Gruber and Köszegi (2001), consumers can perfectly predict future prices, under this assumption, we can include leads of realized prices as forecasts for future prices in equation 19.

Ideally, we would like to address the potential endogeneity between future prices and current demand shocks using a source of variation that is statistically independent from our instruments for contemporaneous prices. Unfortunately, we do not have another source of variation in expected prices. However, since all instruments and controls are included in every first stage regression equation, we can asses *ex-post* if we

have independent variation between current and future prices. This can be done by looking at how different instruments interact with the enactment indicator and its leads; if current prices are shifted by interactions of the enactment indicator and future prices by its leads then we have at least one period in which consumers face prices that differ from those they will face in the future. This reveals the dynamic behavior of consumers and allows us to test if consumers respond to price changes in the future and if this response is different when current prices change.

Furthermore, if the supply-side instruments we use to predict prices differ from one city to another, then, we have cross-sectional variation in the pricing schedule that consumers face from the moment the tax was announce to the moment it is enacted. On the extremum, this would allow us to compare the response of consumers who face an immediate constant change in prices to that of consumers who face a gradual increase over multiple periods.

Regarding the *exp-post* asses of our estimates, aggregating our data over relatively long periods of time mechanically reduces the collinearity between prices in adjacent periods. Moreover, aggregating over longer periods allows the LASSO regression method to potentially select different interactions across first stage equations which induces differences between the the predictions for contemporaneous and future prices.

From the standpoint of any period t, we use the leads of the tax indicator to predict prices l periods into the future. If P_{th} is the price faced by household h at time t, then, the $n-th$ lead of price at time t for household h is $l^n(P_{th}) = P_{t+n}$. The same applies to the variable that indicates the tax enactment. At time t, its $n-th$ lead is given by $l^n(\mathbb{I}_{y(t)>2013}) = \mathbb{I}_{y(t+n)>2013}$. To predict future prices, we use the specifications described with equations 20 and 23, but with the $l-th$ price lead as dependent variable and substituting $\mathbb{I}_{y>2013}$ with its $l-th$ lead. When we use the first set of instruments (i.e., interacting the tax enactment indicator with city indicators), our first stage for the $l-th$ lead of prices is given by:

$$p_{h(t+l)crn} = \eta_h + \tau_c \mathbb{I}_{y(t+l)>2013,c} + \omega_c t + \omega_n t + \upsilon_{ht} \tag{24}$$

Similarly, in the case of local and regional market structure measures, the first stage equation for the $l-th$ lead of prices is:

$$p_{h(t+l)crn} = \eta_h + \theta_r \mathbb{I}_{r,y(t+l)>2013} M_r + \theta_c \mathbb{I}_{c,y(t+l)>2013} M_c + \omega_c t + \omega_n t + \upsilon_{ht} \tag{25}$$

We group all of the instruments derived from equations 20 (*i.e.* the specification that uses city indicators as instruments) to 25 into a final set of first stage models that we estimate using data driven LASSO regressions selecting controls and instruments for SSBs consumption, prices and the leads of prices:

$$p_{h(t+i)ycrn} = \eta_h + \sum_{l=0}^{M} \left(\tau_c \mathbb{I}_{y(t+l)>2013,c} \right) + \omega_c t + \omega_n t + \upsilon_{ht} \tag{26}$$

For the second set of our instruments (i.e., tax enactment indicator interacted with city- or region-level

measures of market concentration), the stage models to be estimated using data driven LASSO regression are as follows:

$$p_{h(t+i)ycrn} = \eta_h + \sum_{l=0}^{M}\Big(\theta_r \mathbb{I}_{r,y(t+l)>2013} M_r + \theta_c \mathbb{I}_{c,y(t+l)>2013} M_c\Big) + \omega_c t + \omega_n t + \upsilon_{ht}$$

In both equations, 26 and 27, the dependent variable $p_{h(t+i)ycrn}$ is the $i-th$ lead of the price variable with $i = 0, 1, 2, 3$. Depending on the specification, M equals 2 or 3. Household fixed effects are again not penalized since they are crucial in our identification strategy.

4.1.3 Lagged consumption

Equation 19 includes lags of consumption, the estimation of panel data models with lagged dependent variables as regressors can lead to biased estimates. Flannery and Hankins (2013) show that, in long panels, the coefficients on exogenous and endogenous regressors are more accurately estimated (in the Root Mean Squared Error sense) by fixed effect (FE) models than existing methods of bias correction, long differencing, or available GMM estimators in panels with 12 or more time periods. In our case, we have twelve two-months periods distributed across two years. Therefore, this is not a concern.

There is a concern however that lagged SSBs consumption is correlated with determinants of contemporaneous consumption other than prices. This is plausible if for example, consumers' adjust their SSBs consumption according to expectations about income, health or occupation. While we cannot reject this possibility, we argue that over longer periods of time (such as two months) these effects are less likely to be an issue since it would require anticipation and persistence to span over long periods of time..

5 Results

First, we we present the results from estimating the structural model on a balanced panel of 4,022 households whose consumption of SSBs we track each bi-semester over two years.

We estimate equation 19 to obtain the direct effect of lagged consumption and prices (contemporaneous and future) using a post-regularization instrumental variables approach. The obtained estimates are reported in panel A of table 2. Columns (1) to (3) show results using the interactions of the tax with city indicators as instruments, and columns (4) to (6) use interactions of the tax with local and regional measures of SSB market concentration. For each first-stage model, the columns sequentially include an additional lead of prices in order to asses how our results are affected by the inclusion of further leads of the price variable.

Panel B reports a number of identification statistics. From the post-lasso model we get weak identification statistics that range between 12 and 23, consistent with no weak identification issues. Moreover, we report the individual F-statistics for the test of significance of the instruments for each first stage regression.

From the magnitude of the F-statistics we conclude that the final set of instruments is strongly correlated with prices. A more detailed discussion of the first stage results can be found in appendix A.

We then use these estimates to test for addiction, forward looking behavior and present bias exploiting the structural model. Results are reported in table 3 and discussed below.

5.1 Addiction

We find a strong association between past consumption and current purchases of SSBs. As panel A of table 2 shows, both the first and second lag of SSB consumption have positive and significant effects on current purchases. Because we include household fixed effects, this result is not driven by unobserved fixed household characteristics. This result is consistent with addiction.

The rate at which the effect of past consumption decreases from one period to another is directly related to the depreciation rate of the addictive stock. We find that the second lag of consumption is significant, however it decreases in magnitude by almost an order of 10. Formally, the coefficient for the $n-th$ lag of SSB consumption is an estimator for $\lambda(1-d)^n$. Solving for d and λ, we find that the addictive stock of sugar sweetened beverages depreciates at a rate of 0.90 (*i.e.* if consumption was reduced to zero, over the span of two months, the addictive effect of SSBs would be just a 10% of its initial magnitude). This interacts with the marginal propensity to addiction λ, which according to our estimates lies between 2.9 and 3. This implies that the fraction of the addictive stock that remains from one two-months period to another has a large and significant effect on contemporaneous SSBs consumption. These estimates are reported in panel B of table 3.

5.2 Forward looking behavior

We find that future prices affect current consumption (table 2). This is consistent with a model for an addictive good as opposed to stock piling. The individual coefficients for temporary price effects can be added as described in section 6.1.3 to recover the response to a permanent price increase at different points in time. The obtained responses to permanent price changes are reported in panel A of table 3. We find that both realized and expected changes in prices significantly affect consumption in the present.

Note that as price changes are expected further into the future, the response of consumers is smaller and its statistical significance tends to decrease. For example, according to columns 3 and 6 of table 3, three periods before the enactment, just the coefficient for P_{t+3} would affect the consumption decisions of households today. The effect of this coefficient alone is significant at the 10% level. However, two periods before the enactment, consumption is affected by both P_{t+3} and P_{t+2} and therefore, the total effect of future expected prices on consumption of SSBs is larger and significant at the .1% level.

5.3 Short and long term price elasticities

In addition to providing evidence of addiction and forward looking behavior, we use the structural model to incorporate information about the depreciation rate and the marginal propensity to addiction

which allows us to estimate the effect of price changes on steady state consumption. With this approach, we can compare elasticities of demand in the short and long run.

We find that for a forward looking agent, the short term elasticity ranges between -1.0, and -1.2. However, in the long run it is about 50% larger and lies between -1.5% and -1.8%. This results illustrate how not accounting for effect of price changes on steady state addictive stocks can lead to significantly understating the long term effect of policies aimed at reducing consumption of addictive goods. Estimates are reported in panel A of 3.

5.4 Present bias

To understand the welfare effects of a tax on an addictive good it is crucial to know if consumers are able to account for the internalities in consumption. Consumers of SSBs behave like forward looking addicts, however, their ability to internalize the health costs of consumption depends on how capable they are to choose an optimal consumption path and follow it. We test for time consistency and find p-values that range between 0.08 and 0.10.

Recall from section 6.1.5 that the test for time-consistency resolves around the fact that, if consumers are quasi-hyperbolic discounters, assuming time consistency leads to underestimations of the parameter δ by the confounding effect of β in comparisons between the present and the period that comes immediately after. To implement this test we need estimates for the immediate response of consumers as well as for their response to price changes expected one and two periods into the future.

Our results for the time-consistency test are reported in panel B of table 3. We are not able to reject the null hypothesis of time-consistency and find evidence in favor of quasi-hyperbolic discounting. The test does not allow to distinguish between naive or sophisticate agents.

6 Comparison with other approaches

In this section, we validate our structural model by benchmarking our short-term price elasticities against price elasticity estimates published in prior work that mostly uses reduced-form approaches, such as a regression discontinuity design (RDD), or a naive demand function that ignores internalities. We use these approaches on our data as well, and compare these reduced-form estimates to the literature and our structural short-term elasticities. We find that these reduced-form approaches yield estimates almost identical to the short-run price elasticities obtained from our structural model, increasing confidence in our empirical approach.

6.1 RDD

First, we apply the RDD approach of Calonico et al. (2014) and use local regression with a second order local polynomial and data-driven selection of bandwidths to identify the period around the threshold that we include observations for. We opt for a data-driven bandwidth selector that chooses the median band-

width from three mean squared error-optimal methods for the RD treatment effect estimator, and assign triangular weights to the observations as a function of the distance to the cutoff (i.e., December 2013). [17] We perform this estimation after adjusting for seasonality and household-specific time-invariant unobservables by including calendar bi-monthly and household fixed effects.

The RDD estimates are visually presented in Figure 1, with the effects on prices in panel A and the effects on SSBs purchases in panel B. Both figures show large and significant jumps in SSBs prices at the cut-off. Panel A of table 5 presents the RDD point estimates. The first column shows the effect of the 2014 tax enactment on the price per liter implied by households' 2013 SSBs' consumption basket. Before the tax enactment, the average price faced by households was 9 Mexican pesos per liter. The tax increased prices to 9.8 Mexican pesos, which translate into a permanent 8.9% increase on average. The second column presents the effect of the tax on total purchases measured in liters. On average, after the tax, households' reduced their purchases by 6.5 liters per bi-mester, which is equivalent to an 8.9% reduction in SSBs purchases relative to the 2013 baseline. Thus, we find that the tax, on average, increased SSB prices by 8.9% and decreased consumption by 9.5%, implying that the short-term price elasticity is about -1. These results are consistent with the results from studies evaluating the 2014 SSB tax in Mexico (Gil, 2006; Barquera et al., 2008; Urzúa, 2008; Unar et al., 2013; Colchero et al., 2015a; Powell et al., 2013; Aguilar et al., 2019) but also from other settings with estimates ranging for instance from -0.6 in France to -1.0 in Philadelphia, PA (Zhong et al., 2018; Cawley et al., 2019a; Capacci et al., 2019; Silver et al., 2017; Lee et al., 2019; Falbe et al., 2016).

6.2 Naive demand model

In this section we estimate a naive model of the demand for SSBs, ignoring internalities. We specify the demand model as follows:

$$a_{htycrn} = \eta_h + \gamma P_{htycrn} + \tau_c t + \tau_n t + \varepsilon_{ht} \tag{27}$$

Where a_{htycrn} represents the liters of sugar sweetened beverages purchased by household h during time t. P_{htycrn} is the price per liter faced by the household during the corresponding period, η_h is a household specific fixed effect, $\tau_c t$ and $\tau_n t$ represent city and socio-economic status time trends and ε_{ht} is an error term. The household fixed effect controls for time-invariant income, prices of other goods and preferences. The error term allows for preferences shocks.

To estimate equation 27, we address the potential endogeneity between prices and purchases of SSBs by using an instrumental variables approach similar to the one used in our structural model. Specifically, we again exploit two different sets of instruments: the tax enactment indicator interacted with city indicators, and the tax enactment indicator interacted with city and regional measures of market concentration.

Panel B of table 5 shows the naive elasticity estimates. The first column shows the results using the first set of instruments (i.e., the tax enactment interacted with city indicators), and the second shows the same results using the second set of instruments (i.e., interactions of tax enactment with local and regional retail

[17] We use the option msecomb2 within the STATA rdrobust command Calonico et al. (2017). Results are not sensitive to using another bandwidth selection procedure or kernel functions.

market concentration variables). We find that an increase of one Mexican peso per litter reduces purchases of sugar sweetened beverages by 8 liters per bi-mester; implying that short-term price elasticity of SSBs demand is about -1. This is consistent with our reduced-form RDD results shown in panel A, but also with other work studying demand responses to price change. Though largely observational, a number of studies using cross-sectional variation in prices found price elasticities ranging from -0.8 to -1.0 (Andreyeva et al., 2010; Powell and Chaloupka, 2009; Smith, 2010; Powell et al., 2013; Gil, 2006; Barquera et al., 2008; Urzúa, 2008; Unar et al., 2013; Colchero et al., 2015a).

6.3 Comparison of studies

We find that the short-term elasticity estimates from our structural model lie within the range of what previous studies have found and of our own elasticity estimates using reduced-form approaches ignoring internalities. Comparison of our reduced form short-term price-elasticity estimates with prior work is illustrated in figure 3. However, our long-term elasticity structural estimates are larger than the short-term ones by about 50%. Thus, while our results show that the 2014 SSBs tax decreased SSBs demand, we claim that if added sugar - one of the main SSBs' ingredients - is addictive, commonly used demand estimation techniques might fail to capture long-term effects of the tax, as well as the mechanisms through which the tax operates in reducing consumption.

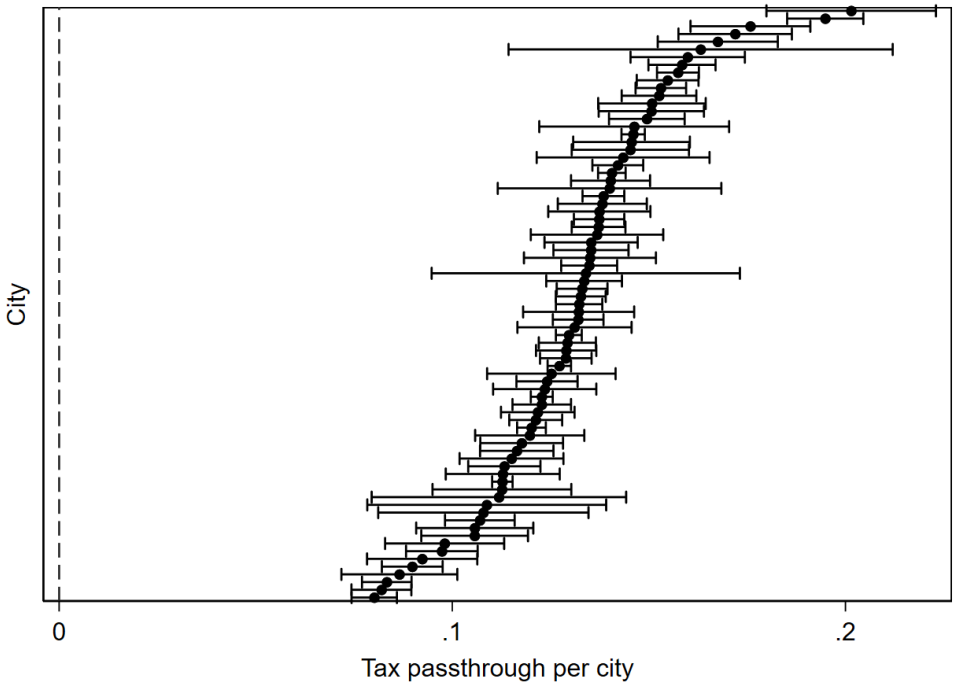
.

7 Conclusion

The economic justification for taxes on SSBs rests on correcting for internalities such as rational addiction and present bias. We provide new evidence of such biases by specifying and estimating a structural model of SSB demand that explicitly incorporates both rational addiction and present bias. We estimate the model using city-level variation in the Mexican 2014 SSB tax price pass-through and panel data on household SSBs purchases between 2013-2015. We show that consumers exhibit both rational addiction and present bias in their SSB purchasing decisions, providing justification for SSB taxes.

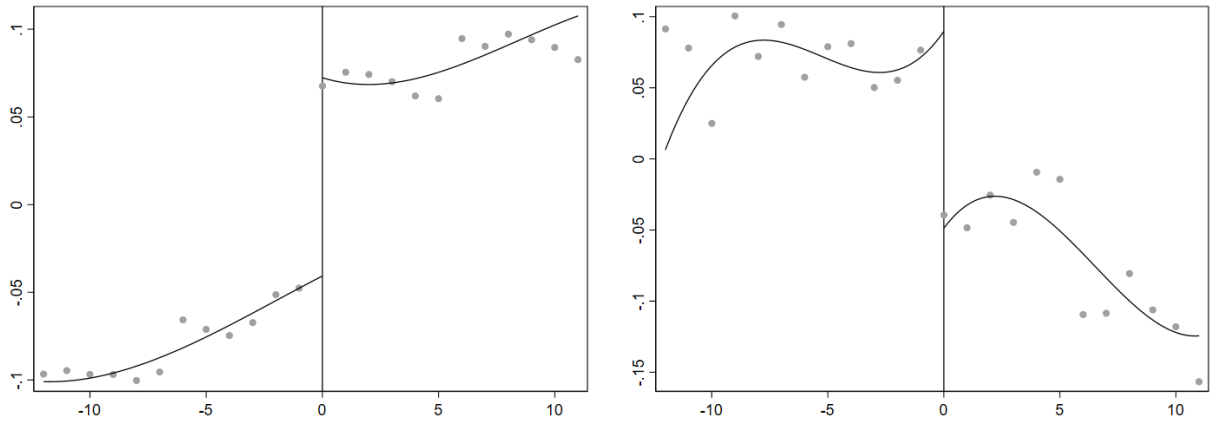
Under addiction, consumption adjusts to a tax over time, implying that long-run equilibrium price elasticities of demand for SSBs are larger than their short-run elasticities. The question is "by how much?" Using the structural model, we estimate a short-run price elasticity of -1.0, but find that the estimated long-run equilibrium price elasticity is 50% higher. To assess the accuracy of the structural estimates, we benchmark the short run elasticities against estimates published in other work and obtained from a reduced-form RDD approach and a naive demand models using the same set of instruments. We show that the short-run price elasticities obtained using these approaches are very similar to the estimates produced from our structural model. Hence, we conclude that empircal models that do not account for internalities underestimate the long-term effect of a tax on SSB consumption.

Figure 1: Price tax passthrough by city



This Figure depicts coefficients for the interactions between city indicators and the tax enactment. The dependent variable is the weighted average price per liter calculated at the household level. Data is aggregated at the two months frequency. The weights are given by the purchases of SSDs made by households during 2013 (measured in liters) and kept constant throughout 2014. Since volumes are kept constant at their 2013 level, the increase in prices is only attributable to changes in prices, not changes in consumer behavior. The regression equation includes city fixed effects. The vertical lines represent 95% confidence intervals. Standard errors are clustered at the household level.

Figure 2: RD effects of the tax on SSB prices and liters consumed

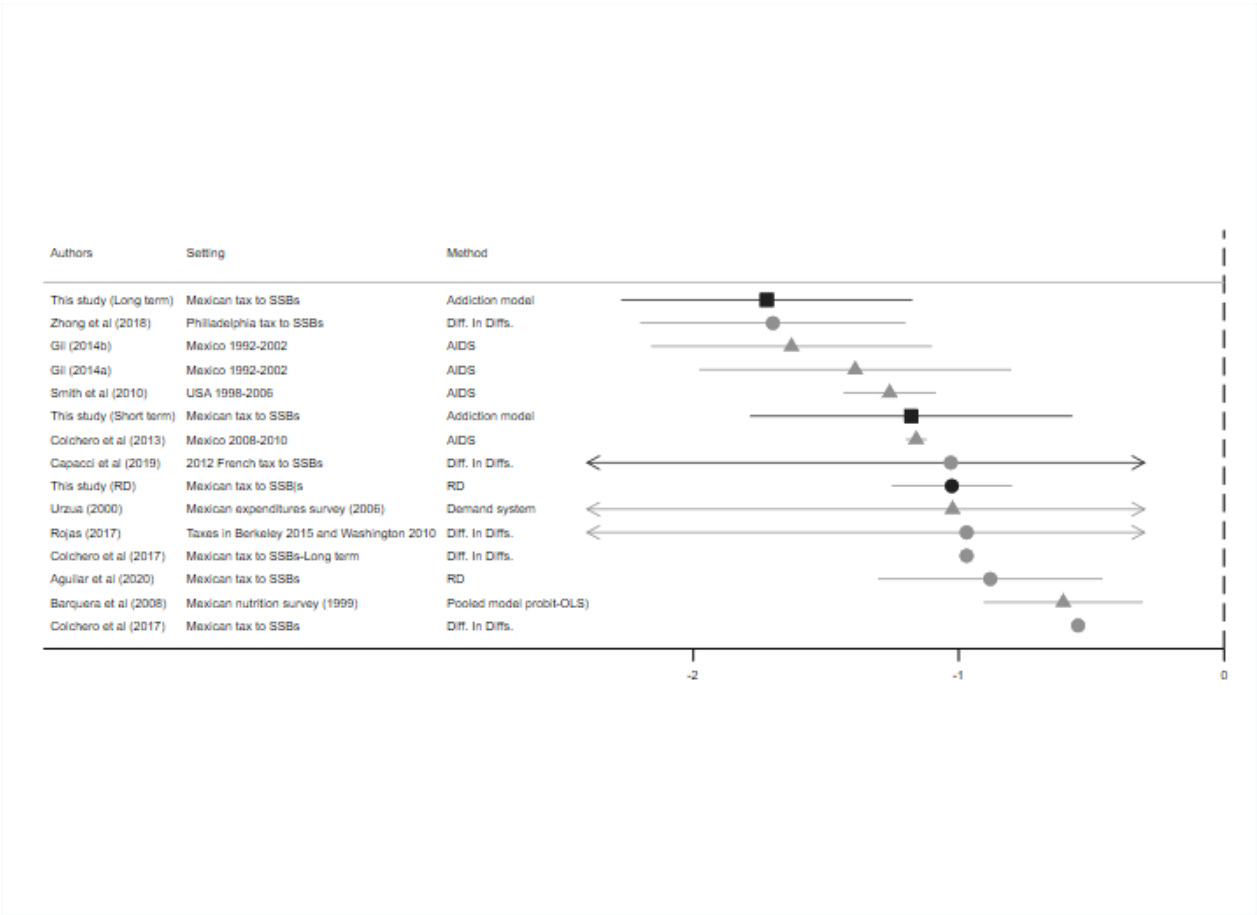


(a) Price per liter of sugar sweetened beverages

(b) Liters of sugar sweetened beverages

These figures show the evolution of the price per liter defined in equation 18 for sugar sweetened beverages (SSb) and purchased liters of SSb. Each graph's horizontal axis covers from 2012 to 2015. The data is aggregated at the two-months frequency. Period 1 corresponds to January and February 2014, when the tax started being collected. The price variable is an average weighted price calculated using households' purchases os SSb during 2013 as weights. Since quantities are kept constant at their 2013 level, the jump observed in January 2014 is only attributable to changes in prices, not changes in consumer behavior. Graphs show RD plots on de-meaned and seasonally adjusted residuals. These residuals were obtained from an OLS regression of the outcome variable on household fixed effects and within year two months period dummies. We then use these residuals at the household week level to plot a month bin scatterplot, and fit a separate quadratic polynomial prediction to weeks corresponding to 2013 and those corresponding to 2014. The level change at the discontinuity corresponds to the inferred effect of the tax.

Figure 3: Comparison with previous work



Notes: This Figure reports estimated or implied elasticities for 11 studies that have either estimated the elasticity of demand for SSBs or measured the reduced form effect of price changes on consumption. By order of appeareancem the studies included in this analysis are (Zhong et al., 2018), (Gil, 2006), (Smith, 2010), (Colchero et al., 2015b), (Capacci et al., 2019), (Urzúa, 2008), (Rojas and Wang, 2017), (Colchero et al., 2017), (Aguilar et al., 2016) and (Barquera et al., 2008)

Table 1: Summary statistics and sample representativeness

Data source	ENSANUT		Kantar
Location	Full sample	Urban	Urban
SSBs (lts per bimester)	59.775	62.26	63.756
	(62.253)	(62.368)	(48.581)
Household size	3.887	3.846	4.369
	(1.852)	(1.814)	(1.765)
Females	2.021	2.007	2.311
	(1.206)	(1.191)	(1.192)
Adults	2.531	2.561	2.748
	(1.178)	(1.196)	(1.146)
HH has Higher Education	0.182	0.22	0.285
	(0.385)	(0.414)	(0.452)
HH Age	49.45	49.4	48.212
	(31.099)	(31.32)	(13.701)
Households	50528	34461	4,022

Notes: This table reports sociodemographic characteristics and average SSBs consumption at the household level for our final sample in the Kantar dataset and ENSANUT 2012. ENSANUT is a survey collected by the Mexican National Institute of Public Health. This survey records sociodemographic, health and nutritional data at the household and individual level for a nationally representative sample of Mexican households. This information includes consumption of sugar sweetened beverages (labeled as Juices and Soft drinks with added sugar) and multiple household characteristics such as the gender, education and age of its members. Consumption of SSBs is reported at he bimester level to facilitate comparisons with our main sample.

Table 2: Structural model of demand estimates

Instruments	City indicators			Local and regional retail market concentration		
	(1)	(2)	(3)	(4)	(5)	(6)
Variables	a_t	a_t	a_t	a_t	a_t	a_t
Panel A: Addiction and forward looking behavior						
Price effects						
P_t	-5.108***	-5.945***	-6.489***	-5.005***	-5.890***	-6.440***
	(0.891)	(0.948)	(1.060)	(0.896)	(0.933)	(1.077)
P_{t+1}	-1.560*	0.0622	-0.351	-1.577*	0.151	-0.345
	(0.828)	(1.044)	(1.026)	(0.844)	(1.015)	(1.031)
P_{t+2}		-2.335***	-0.577		-2.352**	-0.492
		(0.905)	(1.167)		(0.925)	(1.157)
P_{t+3}			-2.258*			-2.395*
			(1.293)			(1.308)
Addictive stock effect						
a_{t-1}	0.320***	0.320***	0.320***	0.321***	0.320***	0.320***
	(0.022)	(0.022)	(0.022)	(0.0214)	(0.0214)	(0.0214)
a_{t-2}	0.034***	0.035***	0.036***	0.0338***	0.0351***	0.0356***
	(0.012)	(0.012)	(0.012)	(0.0116)	(0.0118)	(0.0117)
Panel B: Identification						
IV Identification statistics						
Post-LASSO Weak-ID statistic	22.84	17.40	12.78	20.84	15.59	11.52
Contemporary price instruments F-stat	9.910e+07	31718	15392	238.6	160.2	220
1st lead of price instruments F-stat	218789	12900	2273	206	242.4	137.6
2nd lead of price instruments F-stat		7653	1957		830.3	682.1
3rd lead of price instruments F-stat			453.4			617.7
Empirical specification						
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
City and region time trends	Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic status trends	Yes	Yes	Yes	Yes	Yes	Yes
Sample						
Observations	48,264	48,264	48,264	48,264	48,264	48,264
Number of Households	4,022	4,022	4,022	4,022	4,022	4,022

Notes: This table reports estimates of consumers' response to current and future price changes and the effect of their addictive stock of sugar sweetened beverages (SSb). The estimates are obtained by instrumenting prices with the tax under two specifications, the first exploits the interactions between tax enactment indicators and its leads with city indicators. The second uses the tax and its leads interacted with local and regional market structure statistics such as Herfindahl indexes and the market shares of the most relevant retailers in our data. The dependent variable is current purchases of sugar sweetened beverages (measured in liters). The reported coefficients come from a regression that implements the post-regularization method proposed by Chernozhukov et al. (2015) using the Stata package developed by Ahrens et al. (2018). The data is aggregated at the two months frequency. The price variable is obtained by calculating the average price per liter using weights proportional to the liters purchased in the 2013 consumption basket of every household. Since weights are kept constant, all changes are attributable to price changes and not to consumers' behavior. Panel A reports the direct effect of contemporaneous and future prices on current consumption and the effect of the first and second lag of consumption. Panel B describes the regression specification and provides identification statistics. The Post-LASSO Weak ID-statistic corresponds with the weak identification statistic calculated under standard errors clustered at the city level in the post-lasso regression using the selected controls and instruments. The F-statistics for contemporary price and its leads are calculated from the first stage

regressions as selected by the LASSO regressions.

Table 3: Structural parameters and effects of permanent price changes

Instruments	City indicators		Local and regional retail market concentration	
	(2)	(3)	(5)	(6)
Change in consumption				
Contemporaneous price change	-8.217	-9.674	-8.091	-9.672
p-value	(0.01)	(0)	(0.01)	(0.002)
Expected one period ahead	-2.273	-3.185	-2.201	-3.232
p-value	(0)	(0.002)	(0)	(0)
Expected two periods ahead	-2.335	-2.834	-2.352	-2.887
p-value	(0.01)	(0.007)	(0.011)	(0.006)
Expected three periods ahead		-2.258		-2.395
p-value		(0.081)		(0.067)
Long term effect of price changes				
Short term elasticity	-1.047	-1.232	-1.031	-1.232
Long term elasticity	-1.539	-1.813	-1.609	-1.927
Short term = Long term p-value	(0)	(0)	(0)	(0)
Structural parameters				
λ	2.920	2.885	2.919	2.883
d	0.890	0.889	0.890	0.889
δ	0.790	0.923	0.778	0.935
Time consistency test				
Distance from zero	0.688	0.492	0.734	0.490
p-value	(0.101)	(0.086)	0.0896	0.0862
Panel C: Identification				
Empirical specification				
Leads of price	2	3	2	3
Lags of consumption	2	2	2	2
Sample				
Observations	48,264	48,264	48,264	48,264
Number of Households	4,022	4,022	4,022	4,022

Notes: This table reports estimates of consumers' response to current and future permanent price changes as well as the obtained structural parameters. The responses and parameters are derived from our estimates for the structural demand model. Panel A reports responses for permanent price changes up to three periods into the future as well as immediate and long term elasticities. The response to a permanent price change m periods into the future $\Delta\mu_m$ is obtained by adding the effect of the accumulated contemporary responses. The immediate effect is obtained from panel A. The change in steady state consumption is obtained by solving the steady state equations $a^* = \lambda S^* + \mu$ and $S^* = (1-d)(S^* + a^*)$ for consumption and using $\Delta\mu_m$ as an estimate for the change in μ. Then, elasticities are computed with respect to baseline pre-tax average consumption and price. The structural paramenters are reported in panel B. The marginal propensity to addiction λ and the depreciation rate of the addictive stock d are obtained from the law of motion of the addictive stock. The parameters obtained are used to perform the test of time consistency. The null hypothesis of time consistency yields with p-values that range between 0.087 to 0.108.

Table 4: RD estimates of the tax effects and naive demand estimates

Panel A: RD effect of the tax enactment

	(1)	(2)
Variables	SSDs price per lt	Purchased lts
Tax enactment	0.834***	-6.086***
	(0.117)	(0.688)
Bias corrected effect	0.993	-6.086
Polynomial order	3	3
Bandwidth	3.370	3.595
Effective obs.	28154	28154
Implied elasticity	-1.03	

Panel B: Naive demand elasticity (2SLS)

Instruments	Tax	Tax#Retail competition
	(1)	(2)
Variables	Purchased lts	Purchased lts
SSDs price per lt	-8.061***	-8.188***
	(0.955)	(0.970)
First stage F-statistic	77	52.36
Implied elasticity	-1.133	-1.151

Notes: This table shows reduced form and naive estimates for the response of consumers to the tax enactment. Panel A reports the effect of the tax enactment on consumers' sugar sweetened beverages (SSb) prices and liters purchased estimated using a regression discontinuity strategy. The running variable is time and the tax enactment is the treatment. In column (1) the dependent variable is the price per liter of sugar sweetened beverages (SSb) and in column (2) purchased liters of SSb. The price variable is obtained by calculating the average price per liter using weights proportional to the liters purchased in the 2013 consumption basket of every household. Since weights are kept constant, all changes are attributable to price changes and not to consumers' behavior. Columns 3 and 4 report the demand response to price changes instrumented by the tax enactment and its interactions. We use to sets of instruments, the first is comprised by the tax and its interactions with city indicators. The second comprises the tax and its interactions with retail industry competition and concentration indices, calculated using information from before the enactment. The implied elasticity is calculated as the quotient between the estimated percent change in consumption and the estimated percent change in prices. All regressions include household fixed effects. Standard errors are clustered at the city level.

References

AGUILAR, A., E. GUTIERREZ, AND E. SEIRA (2014): "Taxing to Reduce Obesity," *Working paper*. [8]

——— (2016): "Taxing to reduce obesity," *Wroking Paper*, 1–39. [8, 9, 21]

——— (2021): "The effectiveness of sin food taxes: Evidence from Mexico," *Journal of Health Economics*, 77, 102455. [2, 5, 9]

AGUILAR, A., E. GUTIERREZ, E. SEIRA, ET AL. (2019): "The effectiveness of sin food taxes: evidence from Mexico," *Instituto Tecnológico Autónomo de México (ITAM) Working Paper*. [17]

AHRENS, A., C. B. HANSEN, AND M. E. SCHAFFER (2018): "PDSLASSO: Stata module for post-selection and post-regularization OLS or IV estimation and inference," Statistical Software Components, Boston College Department of Economics. [11, 23]

ALLCOTT, H., B. B. LOCKWOOD, AND D. TAUBINSKY (2019a): "Regressive Sin Taxes, with an Application to the Optimal Soda Tax," *Quarterly Journal of Economics (Forthcoming)*. [1, 2]

——— (2019b): "Should We Tax Soda? An Overview of Theory and Evidence," *Journal of Economic Perspectives*, 33. [1]

ANDERSON, S. P., A. DE PALMA, AND B. KREIDER (2001): "Tax incidence in differentiated product oligopoly," *Journal of Public Economics*, 81, 173–192. [10]

ANDREYEVA, T., M. W. LONG, AND K. D. BROWNELL (2010): "The impact of food prices on consumption: a systematic review of research on the price elasticity of demand for food," *American journal of public health*, 100, 216–222. [2, 18]

AVENA, N. M., J. A. GOLD, C. KROLL, AND M. S. GOLD (2012): "Further developments in the neurobiology of food and addiction: update on the state of the science," *Nutrition*, 28, 341–343. [1]

AVENA, N. M., P. RADA, AND B. G. HOEBEL (2008): "Evidence for sugar addiction: behavioral and neurochemical effects of intermittent, excessive sugar intake," *Neuroscience & Biobehavioral Reviews*, 32, 20–39. [1]

——— (2009): "Sugar and fat bingeing have notable differences in addictive-like behavior," *The Journal of nutrition*, 139, 623–628. [1]

BARQUERA, S., I. CAMPOS-NONATO, L. HERNANDEZ-BARRERA, A. PEDROZA, AND J. A. RIVERA-DOMMARCO (2013): "Prevalence of obesity in Mexican adults 2000-2012," *Salud publica de Mexico*, 55, S151–60. [1]

BARQUERA, S., L. HERNANDEZ-BARRERA, M. L. TOLENTINO, J. ESPINOSA, S. W. NG, J. A. RIVERA, AND B. M. POPKIN (2008): "Energy intake from beverages is increasing among Mexican adolescents and adults," *The Journal of nutrition*, 138, 2454–2461. [2, 17, 18, 21]

BECKER, G. S. AND K. M. MURPHY (1988): "A theory of rational addiction," *Journal of political Economy*, 96, 675–700. [1, 3]

BERNHEIM, B. D. AND A. RANGEL (2004): "Addiction and cue-triggered decision processes," *American economic review*, 94, 1558–1590. [1]

BIRCH, L. L., L. MCPHEE, L. STEINBERG, AND S. SULLIVAN (1990): "Conditioned flavor preferences in young children," *Physiology & behavior*, 47, 501–505. [1]

BUSEY, E. (2020): "Sugary drink taxes around the world," . [1]

CALONICO, S., M. D. CATTANEO, M. H. FARRELL, AND R. TITIUNIK (2017): "rdrobust: Software for regression-discontinuity designs," *The Stata Journal*, 17, 372–404. [17]

CALONICO, S., M. D. CATTANEO, AND R. TITIUNIK (2014): "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 82, 2295–2326. [16]

CAMPOS-VÁZQUEZ, R. M. AND E. M. MEDINA-CORTINA (2019): "Pass-through and competition: the impact of soft drink taxes as seen through Mexican supermarkets," *Latin American Economic Review*, 28, 3. [10, 30]

CAPACCI, S., O. ALLAIS, C. BONNET, AND M. MAZZOCCHI (2019): "The impact of the French soda tax on prices and purchases. An ex post evaluation," *PloS one*, 14, e0223196. [2, 9, 17, 21]

CAWLEY, J. (2015): "An economy of scales: A selective review of obesity's economic causes, consequences, and solutions," *Journal of health economics*, 43, 244–268. [9]

CAWLEY, J., D. FRISVOLD, A. HILL, AND D. JONES (2019a): "The impact of the Philadelphia beverage tax on purchases and consumption by adults and children," *Journal of health economics*, 67, 102225. [2, 17]

CAWLEY, J., A. M. THOW, K. WEN, AND D. FRISVOLD (2019b): "The economics of taxes on sugar-sweetened beverages: a review of the effects on prices, sales, cross-border shopping, and consumption," *Annual review of nutrition*, 39, 317–338. [1]

CHERNOZHUKOV, V., C. HANSEN, AND M. SPINDLER (2015): "Post-selection and post-regularization inference in linear models with many controls and instruments," *American Economic Review*, 105, 486–90. [11, 23]

COLCHERO, M. A., B. M. POPKIN, J. A. RIVERA, AND S. W. NG (2016): "Beverage purchases from stores in Mexico under the excise tax on sugar sweetened beverages: observational study," *bmj*, 352, h6704. [2]

COLCHERO, M. A., J. RIVERA-DOMMARCO, B. M. POPKIN, AND S. W. NG (2017): "In Mexico, evidence of sustained consumer response two years after implementing a sugar-sweetened beverage tax," *Health Affairs*, 36, 564–571. [1, 9, 21]

COLCHERO, M. A., J. SALGADO, M. UNAR-MUNGUÍA, M. HERNANDEZ-AVILA, AND J. RIVERA-DOMMARCO (2015a): "Price elasticity of the demand for sugar sweetened beverages and soft drinks in Mexico," *Economics & Human Biology*, 19, 129–137. [2, 17, 18]

COLCHERO, M. A., J. C. SALGADO, M. UNAR-MUNGUÍA, M. MOLINA, S. NG, AND J. A. RIVERA-DOMMARCO (2015b): "Changes in prices after an excise tax to sweetened sugar beverages was implemented in Mexico: evidence from urban areas," *PloS one*, 10, e0144408. [21]

DELIPALLA, S. AND M. KEEN (1992): "The comparison between ad valorem and specific taxation under imperfect competition," *Journal of Public Economics*, 49, 351–367. [10]

DINICOLANTONIO, J. J., J. H. O'KEEFE, AND W. L. WILSON (2018): "Sugar addiction: is it real? A narrative review," *Br J Sports Med*, 52, 910–913. [1]

DREWNOWSKI, A., J. A. MENNELLA, S. L. JOHNSON, AND F. BELLISLE (2012): "Sweetness and food preference," *The Journal of nutrition*, 142, 1142S–1148S. [1]

FALBE, J., H. R. THOMPSON, C. M. BECKER, N. ROJAS, C. E. MCCULLOCH, AND K. A. MADSEN (2016): "Impact of the Berkeley excise tax on sugar-sweetened beverage consumption," *American journal of public health*, 106, 1865–1871. [2, 17]

FULLERTON, D. AND G. E. METCALF (2002): "Tax incidence," *Handbook of public economics*, 4, 1787–1872. [10]

GERTLER, P. AND T. GRACNER (2021): "The Sweet Life: The Long-Term Effects of a Sugar-Rich Childhood," *Working paper*. [1]

GIL, J. N. V. (2006): "Estimación de elasticidades e impuestos óptimos a los bienes más consumidos en México," *Estudios Económicos*, 127–176. [2, 17, 18, 21]

GRACNER, T., K. KAPNIOS, AND P. GERTLER (2021): "Associations of a National Tax on Non-Essential High Calorie Foods with Changes in Consumer Prices," *Food policy*, forthcoming. [10]

GRUBER, J. AND B. KÖSZEGI (2001): "Is addiction "rational"? Theory and evidence," *The Quarterly Journal of Economics*, 116, 1261–1303. [1, 12]

HAMILTON, S. F. (1999): "Tax incidence under oligopoly: a comparison of policy approaches," *Journal of Public Economics*, 71, 233–245. [10]

HOEBEL, B. G., N. M. AVENA, M. E. BOCARSLY, AND P. RADA (2009): "A behavioral and circuit model based on sugar addiction in rats," *Journal of addiction medicine*, 3, 33. [1]

JAYASINGHE, S. N., R. KRUGER, D. C. WALSH, G. CAO, S. RIVERS, M. RICHTER, AND B. H. BREIER (2017): "Is sweet taste perception associated with sweet food liking and intake?" *Nutrients*, 9, 750. [1]

LAIBSON, D. (1997): "Golden eggs and hyperbolic discounting," *The Quarterly Journal of Economics*, 112, 443–478. [3]

LEE, M. M., J. FALBE, D. SCHILLINGER, S. BASU, C. E. MCCULLOCH, AND K. A. MADSEN (2019): "Sugar-sweetened beverage consumption 3 years after the Berkeley, California, sugar-sweetened beverage tax," *American journal of public health*, 109, 637–639. [2, 17]

LUSTIG, R. H. (2010): "Fructose: metabolic, hedonic, and societal parallels with ethanol," *Journal of the American Dietetic Association*, 110, 1307–1321. [1]

O'DONOGHUE, T. AND M. RABIN (2006): "Optimal sin taxes," *Journal of Public Economics*, 90, 1825–1849. [1]

POWELL, L. M. AND F. J. CHALOUPKA (2009): "Food prices and obesity: evidence and policy implications for taxes and subsidies," *The Milbank Quarterly*, 87, 229–257. [2, 18]

POWELL, L. M., J. F. CHRIQUI, T. KHAN, R. WADA, AND F. J. CHALOUPKA (2013): "Assessing the potential effectiveness of food and beverage taxes and subsidies for improving public health: a systematic review of prices, demand and body weight outcomes," *Obesity reviews*, 14, 110–128. [2, 17, 18]

RIVERA, J. Á., T. G. DE COSSÍO, L. S. PEDRAZA, T. C. ABURTO, T. G. SÁNCHEZ, AND R. MARTORELL (2014): "Childhood and adolescent overweight and obesity in Latin America: a systematic review," *The lancet Diabetes & endocrinology*, 2, 321–332. [1]

ROJAS, C. AND E. Y. WANG (2017): "Do taxes for soda and sugary drinks work? Scanner data evidence from Berkeley and Washington," *Scanner Data Evidence from Berkeley and Washington (September 23, 2017)*. [21]

RTVELADZE, K., T. MARSH, S. BARQUERA, L. M. S. ROMERO, D. LEVY, G. MELENDEZ, L. WEBBER, F. KILPI, K. MCPHERSON, AND M. BROWN (2014): "Obesity prevalence in Mexico: impact on health and economic burden," *Public health nutrition*, 17, 233–239. [1]

SALGADO, J. C. AND S. W. NG (2019): "Understanding heterogeneity in price changes and firm responses to a national unhealthy food tax in Mexico," *Food policy*, 89, 101783. [10]

SILVER, L. D., S. W. NG, S. RYAN-IBARRA, L. S. TAILLIE, M. INDUNI, D. R. MILES, J. M. POTI, AND B. M. POPKIN (2017): "Changes in prices, sales, consumer spending, and beverage consumption one year after a tax on sugar-sweetened beverages in Berkeley, California, US: A before-and-after study," *PLoS medicine*, 14, e1002283. [2, 9, 17]

SMITH, T. A. (2010): *Taxing caloric sweetened beverages: potential effects on beverage consumption, calorie intake, and obesity*, DIANE Publishing. [2, 18, 21]

STERN, N. (1987): "The effects of taxation, price control and government contracts in oligopoly and monopolistic competition," *Journal of Public Economics*, 32, 133–158. [10]

SULLIVAN, L. AND C. BRUMFIELD (2016): "The first 1,000 days: Nourishing America's future," . [1]

THUROW, R. (2016): "The first 1,000 days: A crucial time for mothers and children—and the world," *Breastfeeding Medicine*, 11, 416–418. [1]

UNAR, M., M. COLCHERO, G. TERUEL, AND J. RIVERA DOMMARCO (2013): "The effect of food and beverage prices on weight among women participating in the Mexican Family Life Survey," *Cuernavaca: Instituto Nacional de Salud Pública*. [2, 17, 18]

URZÚA, C. M. (2008): "Evaluación de los efectos distributivos y espaciales de las empresas con poder de mercado en México," *Mexican Federal Competition Commission*. [2, 17, 18, 21]

WEYL, E. G. AND M. FABINGER (2013): "Pass-through as an economic tool: Principles of incidence under imperfect competition," *Journal of Political Economy*, 121, 528–583. [10]

WISS, D., N. AVENA, AND P. RADA (2018): "Sugar Addiction: From Evolution to Revolution," *Frontiers in psychiatry*, 9, 545. [1]

ZHONG, Y., A. H. AUCHINCLOSS, B. K. LEE, AND G. P. KANTER (2018): "The short-term impacts of the Philadelphia beverage tax on beverage consumption," *American journal of preventive medicine*, 55, 26–34. [2, 9, 17, 21]

A First stage and price variation

In this appendix we detail the supply-side variation we use to estimate the dynamic response of consumers to price changes. We start by showing our first stage results and discussing the variables selected by the post-regularization procedure. Then we show how the selected instruments correlate with observable city-level characteristics. Finally, we show the variation current and future prices induced by the instruments.

Table 5 shows the results from estimating the first stage models defined by equation 27 (the model that uses local and regional measures of industry concentration). Panel A corresponds to the model that includes two leads of the price variable whereas panel B corresponds to the model with three price leads. In both panels, each column corresponds to the first stage for the corresponding price lead (starting from contemporaneous prices P_t and then adding its leads P_{t+i}. The rows correspond to the interactions between the tax indicator (or its leads) with the selected supply-side city-level instruments.

The sources of variation chosen by the post-regularization method are the city-level Herfindahl-Hirschman index, the shares of some of the largest supermarket chains in our dataset along with the market shares that correspond to smaller but widespread traditional retailers such as corner-stores and spring water retailers.

As noted by Campos-Vázquez and Medina-Cortina (2019) local market concentration is correlated with the fraction of the tax that was passsed to consumers. While this variable is chosen by the post-regularization procedure, after including other supply-side variables we do not find a significant association between prices and the Herfindahl-Hirschman index.

On the other hand, we find a strong association between the market shares of the largest supermarket chains and higher higher portions of the tax being passed to consumers; especially at the region level. The chains with the largest effects are Bodega Aurrera and Soriana (the former being the largest supermarket chain in Mexico and the later its main competitor). Furthermore, we also find that the market shares of corner stores (which account for more more than half of the observed purchases in our data) and spring water retailers are also associated to higher price increases after the tax enactment.

The instrumental variables method requires that we include all the selected instruments in all the first stage models. This means that predictions for prices at time $t+i$ could be driven by expected taxes at time $t' > t+i$. While this is not a technical problem; identification of our empirical model not only requires cross-sectional variation in the effect of the tax on prices but also variation in how prices increase over time. To identify the response of consumers to changes in both current and future prices we need consumers to face different relative prices between current and future consumption. Figure 4 depicts the first stage predicted effects of the tax on contemporaneous and future prices. Subfigure a shows how, from the standpoint of period t, predictions for prices at time $t+i$ would change if a tax at time $t+i$ was announced. However, we are interested in the response of consumers to permanent price increases. Since our first stage models include multiple leads of the tax indicator, expectations further into the future will mechanically affect the prediction for P_{t+i}. Subfigure b reports the sum of all coefficients for interactions with L_j for $j \geq i$ on P_{t+i} which corresponds to the actual predicted change in prices if a permanent tax starting from period $t+i$ was announced at time t.

Figure 4 serves two purposes, first, it depicts the heterogeneity in current and future that results from the retail concentration instruments. Second, it shows how our predictions for realized and expected prices differ. Our data is aggregated at a long frequency (two months) which allows for differences in the price variable from one period to another. This causes differences in the models selected for predicting different leads of the price variable. Hence, we observe some differences even within the same city between contemporaneous and expected price changes. However, taking the mean of these predictions across cities, we find that on average, there are no substantial differences between realized and expected price changes which is consistent with a model of perfect information.

Finally, we explore how the variation we use for identification correlates with observable household and city characteristics. We do this by regressing the observed tax effects, our predicted tax effects, and all the supply-side variables that were selected as instruments, on a set of city-level characteristics comprised by: population according to the 2010 census, average household pre-tax SSDs consumption in liters, average household size and median socio-economic status as reported by Kantar.

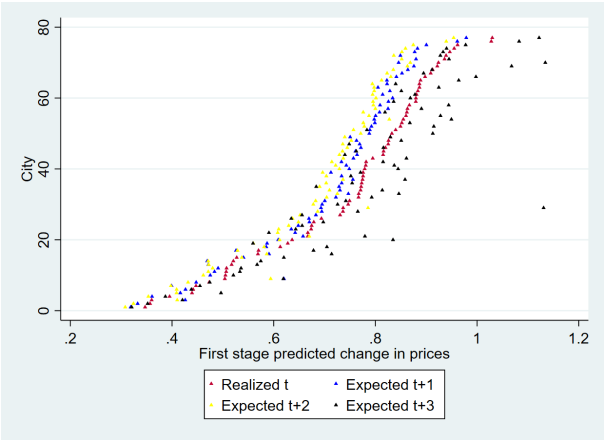
Table 5: First stage results - Local and regional retail industry concentration

Variables	P_{t+0}	P_{t+1}	P_{t+2}
HHI-Region#L_0	2.054***	0.216	-0.332
	(0.428)	(0.279)	(0.307)
HHI-Region#L_1	-0.00547	1.821***	0.360*
	(0.183)	(0.357)	(0.217)
HHI-Region#L_2	-0.00414	-0.0748	1.610***
	(0.325)	(0.339)	(0.448)
HHI-City#L_0	0.567*	-0.0916	-0.00697
	(0.319)	(0.227)	(0.233)
HHI-City#L_1	0.0662	0.623**	-0.0784
	(0.159)	(0.256)	(0.188)
HHI-City#L_2	0.560*	0.542*	1.110***
	(0.291)	(0.306)	(0.318)
Aurrera#L_0	0.562*	-0.0236	-0.249
	(0.337)	(0.228)	(0.270)
Aurrera#L_1	-0.237	0.532	0.186
	(0.208)	(0.350)	(0.156)
Aurrera#L_2	0.0842	-0.161	0.421
	(0.246)	(0.306)	(0.347)

Variables	P_{t+0}	P_{t+1}	P_{t+2}	P_{t+3}
HHI-Region#L_0	1.944***	0.202	-0.338	-0.0398
	(0.451)	(0.312)	(0.324)	(0.254)
HHI-Region#L_1	-0.0339	1.812***	0.350	-0.313
	(0.184)	(0.357)	(0.220)	(0.335)
HHI-Region#L_2	0.158	-0.221	1.670***	0.448*
	(0.248)	(0.207)	(0.357)	(0.229)
HHI-Region#L_3	-0.742***	-0.280	-0.439	1.315***
	(0.274)	(0.334)	(0.343)	(0.399)
HHI-City#L_0	0.638*	-0.0467	-0.00880	-0.309
	(0.353)	(0.282)	(0.259)	(0.200)
HHI-City#L_1	0.0830	0.634**	-0.0759	0.158
	(0.158)	(0.257)	(0.194)	(0.217)
HHI-City#L_2	0.241	0.132	0.664***	-0.118
	(0.199)	(0.162)	(0.249)	(0.194)
HHI-City#L_3	0.541**	0.621**	0.615**	1.118***
	(0.248)	(0.302)	(0.306)	(0.295)
Aurrera#L_0	0.737*	0.217	0.00423	-0.0907
	(0.405)	(0.215)	(0.242)	(0.225)
Aurrera#L_1	-0.208	0.574	0.229	0.169
	(0.221)	(0.361)	(0.156)	(0.209)
Aurrera#L_2	0.273	-0.0822	0.640*	0.182
	(0.213)	(0.241)	(0.356)	(0.151)
Aurrera#L_3	0.156	0.338	0.140	0.928**
	(0.231)	(0.265)	(0.342)	(0.379)
Chedraui#L_0	-0.434	-0.445	-0.583*	-0.793***
	(0.351)	(0.337)	(0.322)	(0.294)
S-Mart -Region#L_2	8.492***	9.571***	5.115	-5.504
	(2.907)	(3.522)	(4.528)	(4.975)
S-Mart -Region#L_3	8.847***	7.196***	9.743**	19.35**
	(2.520)	(2.535)	(4.082)	(7.789)
Soriana-Region#L_3	3.910*	3.618	1.425	-4.632
	(2.078)	(2.318)	(2.707)	(3.340)

Figure 4: Heterogeneity in predicted observed and expected prices changes

(a) Only contemporaneous effects



(b) Full first stage models

