Supplemental Online Content

Gračner T, Marquez-Padilla F, Hernandez-Cortes D. Changes in weight-related outcomes among adolescents following consumer price increases of taxed sugar-sweetened beverages. *JAMA Pediatr*. Published online December 13, 2021.
doi:10.1001/jamapediatrics.2021.5044

This supplemental material has been provided by the authors to give readers additional information about their work.

eSuplement 1. Data and Study Sample

1.1 Data overview

1.1.1 Price Data

We used data on nominal prices measured at the individual product level based on a unique universal product code across stores in 39 Mexican cities across all 32 states monthly between 2011-2016 by the National Institute of Statistics and Geography (INEGI). Price data, originally collected for the purpose of computing the national consumer price index (CPI) by INEGI, were representative of urban areas with more than 20,000 inhabitants and were provided to us free of charge. Data are also publicly available at https://www.inegi.org.mx/programas/inpc/2018/. All prices were expressed in either liter or kilogram and include taxes to reflect consumer prices. They include city and municipality identifiers. Though price data by INEGI were collected for food and beverage products, the main prices of interest are all tax-eligible beverage prices.

Prices were collected from several stores (convenience or corner stores, mini-markets or family-owned grocery stores, informal retail and street stalls, supermarkets, and warehouse clubs). Convenience or corner stores sell a wide range of everyday items, such as groceries, snack foods, soft drinks, as well as tobacco products, and alcoholic beverages; like 7-Eleven in the US. Mini-markets are smaller grocery stores, often family-owned that sell the same range of products as convenience stores but may also sell basic household appliances. A supermarket is a self-service store offering food and household products, organized into aisles, like Safeway or Trader's Joe in the US. A warehouse club usually sells a wide variety of merchandise offered in large, wholesale quantities, such as Sam's club and Costco in the US. This data was collected by INEGI surveyors and not by the retailers or purchasers.

Products were tracked using a unique product identifier continuously within stores over time. Importantly, store price data were collected with a barcode-equivalent product's descriptions.

These included products' name and brand, packaging type and weight, such as "Fanta, de Sabor, Botella de 1LT" or "Coca Cola, Refrescos, Light, Botella de 600ml". This detailed item's description enabled us to determine whether beverages were tax eligible (with added sugars) or tax exempt (without added sugars). We first identified tax-eligible beverages. Tax-eligible beverages were non-alcoholic beverages with added sugars (with the exception of medical beverage products), such as carbonated SSBs (i.e., soft drinks or regular sodas), non-carbonated SSBs (i.e., industrialized juices, flavored waters or *aguas frescas*), as well as energy drinks, powders and concentrates. Data recorded prices for 826 distinct tax-eligible beverage products or SSBs (or 52,117 observations in total). In our second step, we divided products into carbonated vs non-carbonated SSBs and found that the majority or 562 (or 35,217 observations in total) of SSBs were carbonated SSBs or regular sodas. While prices of these were available in all cities, prices of other SSBs were not (e.g., prices of flavored water were only available in 22 cities). For these reasons and the fact that prices of regular sodas changed the most post tax[1], we focus our analyses on carbonated SSBs or regular sodas only.

Finally, we calculated yearly city-level mean prices for regular sodas (excluding bottom and top percentile in prices across products). We used the 2011 Mexican CPI to obtain real prices. We log-transformed real prices of soda and obtained their one and two-year lagged values for the purpose of our analyses. Figure 1 shows the trajectory of soda prices over time between 2011 and 2016 (panel A) and how they changed before vs after the tax across cities (panel B). We calculated the latter by regressing logged prices of soda on a binary variable indicating periods after (vs before) 2014, interacted with city-level indicators.

1.1.2 Data on objectively measured height and weight

This data was provided to us by IMSS (Institute of Mexico's Social Security) under the academic DUA collaboration agreement. The data provided was from their electronic medical records

(EMR) for a universe of non-pregnant teenagers born between 1999-2002 who were insured by IMSS and visited one of its primary care clinics at least once between 2012 and 2017. The IMSS is Mexico's main public healthcare provider covering more than 50% of the population, the majority of which (i.e., 78%) live in urban areas. IMSS healthcare services are available to formal workers and their families who tend to have higher income and education levels on average than informal workers.[1] eTable 2 shows unadjusted means and standard deviations for continuous variables and shares for binary variables for the sample of adolescents included in our (IMSS) sample and for adolescents surveyed in the National Health and Nutritional Survey (ENSANUT) in 2012. While IMSS data correspond to information from medical appointments and it thus affected by selection into the healthcare system, ENSANUT data is a nationally representative survey of Mexican households. Data was provided to us free of charge.

The IMSS data contain information on patients' age, sex, objectively measured height, weight, and a primary reason for a patient's medical visit, recorded using the ICD-10 codes. Weight, height, and ICD-10 diagnoses were measured and recorded by a health care provider during medical visits; following the national health care standards described in the Norma Oficial Mexicana, NOM-024-SSA3-2010. Teens' height and weight measurements were used to calculate their BMI (weight in kilograms divided by height in meters squared). We obtained their average BMI in each year between 2012-2017. Z-scores for BMI were then calculated according to the WHO Growth Reference chart with Stata package "zanthro" command using sex, age (in years), and BMI. BMI z-scores were then converted into percentiles based on a standard Normal distribution. Finally, teens were then classified as overweight or obese (OWOB) or at risk for OWOB for BMI at or above 85th or 75th percentile, respectively. We used ICD-10 codes to

[1] Marquez-Padilla, Fernanda. "When less is more: Can reduced health monitoring improve medication adherence?." Journal of Health Economics 75 (2021): 102387; Llamas, Rogelio Varela, and Mayra Yesenia Nava Rubio. "Salarios e informalidad laboral en México: Una perspectiva regional y empresarial." Revista de Estudios Regionales 118 (2020): 15-46.

identify other diseases, such as endocrine (ICD-10 chapter 4, codes E00-E89) and digestive diseases (ICD-10 chapter 11, codes K00-K95), neoplasms or Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism (ICD codes C00-D89), or external diseases (ICD-10 chapters 19 and 20, codes S00-Y99).

The data include each patient's clinic location which we associated to a municipality based on the clinic geographic coordinates. As patients are assigned to clinics based on their address of residence, we use clinics' municipality to proxy for individuals' municipality of residence.

1.1.3 Data on socioeconomic characteristics

We used data from additional sources to compare different socioeconomic and environments' characteristics across cities that experience different levels of price changes. The data sources and variables used are described in eTable 1. We obtained other health and nutrition data from ENSANUT 2012 (National Health Survey), which is a nationally representative survey of Mexican households that reports characteristics for all household members, such as weight, height, age, as well as household socioeconomic characteristics, in addition to different food items consumed per day. We calculated means at the municipality level and obtained the average values for the municipalities within each city. We also obtained municipal level socioeconomic characteristics (percentage of households in extreme poverty, percentage of population working in the informal economy) for year 2010 from CONEVAL (National Evaluation Council) and INEGI and averaged them across cities to create means for low/medium/high-price change areas. Table 1 shows descriptive statistics (unadjusted mean and SD for continuous variables and shares for dummy variables) for each variable listed in eTable 1.

1.2.1.1 Study sample

We used patients' clinics' municipality identifier to merge IMSS health data to price data to assign price variables (measured at the city level) to individuals living in different municipalities. Using QGIS software, we first calculated the Euclidian distance between each city's municipality's centroid to every other municipality's centroid in the country. We restricted our sample to adolescents who live in urban municipalities close to the cities to minimize measurement error in our key variable of exposure (i.e., price change over time). Furthermore, we restricted the analysis to the patients that stay in the same municipality during our study period (although we did not exclude children who change clinics within the same municipality). Specifically, we limited the distance between individuals' municipality and the city where prices were measured to be at most 30km to minimize measurement error in price exposure. For a similar reason, we excluded rural municipalities given that the price data is representative of urban areas only. The 25th percentile distance between the clinic and city where prices were measured was 8 kilometers and the 75th percentile was 66.7 kilometers. Finally, we restricted the sample to the adolescents observed over the entire period of our study, that is, on those who visited a clinic at least once and up to ten times each calendar year to ensure a balanced panel and avoid the concern related to compositional changes to our sample over time. Our final sample consists of 12,654 individuals, each observed exactly six times.

To better understand how our selected sample of IMSS patients compare to a general population of adolescents in Mexico, we compared the characteristics of adolescents receiving healthcare from IMSS in our sample to adolescents' outcomes as recorded in the nationally representative survey, the ENSANUT 2012. Specifically, we compared characteristics of our data to the nationally representative characteristics of adolescents born between 1999-2002 in ENSANUT 2012. eTable 2 shows that the sample of teenagers observed in IMSS data is on average younger and included more girls than in the general population, reflecting the fact that teenage girls tend to use healthcare services more frequently than boys. Raw height, weight, and z-BMI measures differ due to the age and sex composition of both samples. However, at risk for OWOB and OWOB prevalence is higher in the IMSS sample, arguably reflecting that heavier individuals are more likely to attend a clinic.

We observe a negative unadjusted ecological association between city-level soda price changes (2013 vs 2014) and percentage changes in mean BMI percentile at the city level between 2013 and 2017 in eFigure 2 (correlation of -0.078).

Tables

eTable 1. Description of Data Sources for Individual and Environment Characteristics

Variables	Data sources	Description
Age	IMSS EMR 2013-2017	Obtained from IMSS medical records at each visit.
Gender	IMSS EMR 2013-2017	IMSS medical records report sex for each patient.
Weight (kgs)	IMSS EMR 2013-2017	IMSS medical records report weight at the time of the medical visit recorded by a physician.
Height (m)	IMSS EMR 2013-2017	IMSS medical records report height at the time of the medical visit recorded by a physician.
z-BMI	IMSS EMR 2013-2017	IMSS medical records report height, weight, age, and sex. We obtained z-BMI measure using zanthro STATA command (see Vidmar, Cole & Pan (2013) and details below.)
BMI (pct)	IMSS EMR 2013-2017	Obtained using the z-BMI measure, which was converted into percentiles based on a standard Normal distribution.
At risk for OWOB	IMSS EMR 2013-2017	Defined at or above 75th percentile using the BMI (pct) variable.
OWOB	IMSS EMR 2013-2017	Defined at or above 85th percentile using the BMI (pct) variable.
Digestive diseases (%)	IMSS EMR 2013-2017	IMSS medical records report diagnosis at the time of the visit. We defined digestive diseases based on ICD-10 codes K00-K95 and calculated the percent of visits due to digestive diseases as a share of total visits.
Endocrine diseases (%)	IMSS EMR 2013-2017	IMSS medical records report diagnosis at the time of the visit. We defined digestive diseases based on ICD-10 codes E00-E89 and calculated the percent of visits due to endocrine diseases as a share of total visits.
N of visits per year	IMSS EMR 2013-2017	IMSS medical records report the date of the visit, and we calculated the total visits to IMSS centers per year.
Daily soda intake (ml/day)	ENSANUT 2012	Reported by ENSANUT 2012 for ages 12 or above. We limited the sample for ages 12-18 and calculated total ml/day from the reported number of soda glasses consumed per day (soda glass is defined as 240 ml) and obtained average daily soda intake per city.
Education level (%)	ENSANUT 2012	Education level reported by ENSANUT 2012, We present the share of people with completed middle school (*secundaria*) or more.
Extreme Poverty (%)	CONEVAL 2010	Calculated by CONEVAL at the municipality level using the national extreme poverty line in 2010. We obtained average poverty percentage for municipalities within each city.
Population working in informal economy	CONEVAL 2010	Calculated by CONEVAL at the municipality level using information on social security access in 2010. We obtained average informality percentage for municipalities within each city.
Number of supermarkets per 100 convenience stores	DENUE 2012 (INEGI)	Total economic units per municipality obtained from INEGI's Statistical Directory of Economic Unites (DENUE). DENUE lists existing convenience stores, supermarkets, and soda manufacturers with geographic location. Calculated the number of stores (by store type) per municipality and obtained the ratio between supermarkets and convenience stores (x100).
2013 SSB price (2011 MXN$)	INEGI	Collected monthly by INEGI to calculate Mexican CPI. We calculated the average yearly SSB price at the city level. Data provided to us by INEGI.

eTable 2. Descriptive statistics as observed in IMSS and ENSANUT in 2012[1]

	Data source		
	IMSS 2012	**ENSANUT 2012**	**Difference**
	(mean (s.d.))	(mean (s.d.))	
Age	10.88	12.39	-1.51***
	(0.96)	(0.59)	(0.02)
Male (%)	45.87	50.72	-4.85***
	(49.83)	(50.01)	(1.12)
Weight (kgs)	41.00	51.24	-10.24***
	(11.74)	(12.96)	(0.26)
Height (m)	1.43	1.54	-0.11***
	(0.10)	(0.08)	(0.002)
z-BMI[2]	0.84	0.76	0.07*
	(1.30)	(1.28)	(0.02)
BMI (pct)	69.75	68.16	1.59*
	(30.60)	(30.55)	(0.68)
At risk for OWOB (%)[3]	56.20	53.15	3.04**
	(49.62)	(49.91)	(1.11)
OWOB (%)[3]	46.45	42.75	3.69***
	(49.88)	(49.48)	(1.12)
IMSS affiliation		34.02	
		(47.39)	
Seguro popular affiliation		41.73	
		(49.32)	
No healthcare affiliation		17.39	
		(37.91)	
Observations	12,654[4]	2,346[5]	

[1] Descriptive statistics are calculated for teenagers in the IMSS sample in 2012, and for teenagers in ENSANUT in 2012. Mean value and standard deviation reported for continuous variables, and shares are reported for binary variables.
[2] z-BMI was calculated for the IMSS sample using the WHO growth charts[2]. z-BMI as reported by the dataset in ENSANUT 2012. *** implies p<0.001, ** p<0.01, and * p<0.05.
[3] At risk for OWOB and OWOB are defined at the 75th or 85th BMI percentile or above, respectively.
[4] 2012 sample restricted to cohort born between 1999 and 2002 and patients in urban areas.
[5] Sample restricted to patients in urban areas for teenagers ages 10-13 as reported in the ENSANUT 2012 survey. Age restriction applied given that there are no individuals in 2001 and 2002 in the teenagers ENSANUT sample.

eAppendix 2. Statistical Analyses

2.1 Statistical Models

We exploit non-random variation in prices at the city level following the tax to analyze associations between price changes and weight-related outcomes. Prior work finds multiple reasons why prices may respond differently to taxes including demand-side factors (taste for SSB, health awareness or income) and supply-side factors (market power, strategic behavior by retailers). While it is beyond the scope of this paper to explain the mechanisms leading to differential price responses to taxes, Table 1 shows individual, population, and environmental characteristics for regions with small (<5%), medium (5-10%), and high (≥10%) price changes before vs after the tax, and shows that in cities with greater price changes adolescents were heavier on average and that the share of larger vs smaller retail (e.g., supermarkets vs convenience stores) was increasing with price change, consistent with prior work.[2]

Exploiting city-level variation in prices, we estimated the sensitivity of weight-related outcomes to changes in prices of soda using the following regression model for each outcome:

$$(1)\ Y_{ict} = \alpha_0 + \beta_1 log(P_{soda})_{c,t-1} + \beta_2 log(P_{soda})_{c,t-2} + X'_{ict}\delta + \gamma_i + \sigma_t + \epsilon_{ict}$$

Y_{ict} describes an outcome variable (either continuous for BMI percentile or binary for prevalence in of overweight and obesity, OWOB, or at risk for OWOB). In this model, we regressed outcomes of interest, Y_{ict}, on one and/or two-year lags of prices of soda (log-transformed), and key covariate(s) of interest such as age fixed effects, indicators on whether the patient had at least one digestive or endocrine diagnosis during the year, and categories for the number of individual visits to a clinic in a year (i.e., number of visits<6 or visits between 6-10). This model includes individual fixed effects, γ_i, calendar-year (σ_t) fixed effects to control for common

[2] Campos-Vázquez, R.M. and Medina-Cortina, E.M., 2019. Pass-through and competition: the impact of soft drink taxes as seen through Mexican supermarkets. Latin American Economic Review, 28(1), pp.1-23; Salgado JC, Ng SW. Understanding heterogeneity in price changes and firm responses to a national unhealthy food tax in Mexico. Food policy. 2019;89:101783; Cawley, J., Thow, A.M., Wen, K. and Frisvold, D., 2019. The economics of taxes on sugar-sweetened beverages: a review of the effects on prices, sales, cross-border shopping, and consumption. Annual review of nutrition, 39, pp.317-338; Weyl, E.G. and Fabinger, M., 2013. Pass-through as an economic tool: Principles of incidence under imperfect competition. Journal of Political Economy, 121(3), pp.528-583; Bonnet, C. and Réquillart, V., 2013. Impact of Cost Shocks on Consumer Prices in Vertically-Related Markets: The Case of The French Soft Drink Market. *American Journal of Agricultural Economics*, *95*(5), pp.1088-1108; Colchero, M.A., Salgado, J.C., Unar-Munguía, M., Molina, M., Ng, S. and Rivera-Dommarco, J.A., 2015. Changes in prices after an excise tax to sweetened sugar beverages was implemented in Mexico: evidence from urban areas. *PloS one*, *10*(12), p.e0144408.

yearly national shocks. In addition to other covariates described in X'_{ict} in the model (1) above, vector X'_{ict} may also include one- and/or two-year lags prices of other foods (i.e., pastries, cake, cereal, candy, cookies, and savory snacks). Standard errors are clustered at the city level. These models were estimated for all, and by sex.

β_1 and β_2 are the key coefficients of interest. All coefficients presented in tables were multiplied by 10, as we interpret all our results at the average price increase, which was 10%. We interpret β_2, for instance, as a change in the OWOB prevalence in time *t* by (β_2/100)*[100 percentage points (pp)] for a 10% change in prices of soda in *t*-2.

The main results from this model are presented in Table 2, but alternative specifications are estimated and presented in eTables 3 and 4 (see Sensitivity analyses section below). We find that our conclusions do not change and that at large, soda prices at t-1, and prices of other foods are not associated with changes in weight-related outcomes.

To examine whether changes in soda prices are associated with changes in BMI percentile for heavier teens, we run our main specification separately for individuals that were at risk for OWOB in 2012 and those that were not. We present these results in Table 3.

We then performed heterogeneity analyses to examine whether changes in weight-related outcomes before vs after the tax varied by areas that experienced a low (below 5%), medium (5-10%), or high (at least 10%) increase in soda prices after the tax. We regressed whether a patient was at risk for OWOB on an interaction term between an after-tax indicator variable that equals one in calendar years 2014-2017 (0 otherwise), $AfterTax_t$, and indicator variables for cities in which percentage change in SSB prices between 2013 and 2014 was below 5% (D_c^L), 5-10% (D_c^M), or at least 10% (D_c^H). The same covariates as listed above are included, except that we include linear time trends instead of year fixed effects to avoid collinearity with the “AfterTax” indicator. The coefficients associated with $\mathrm{AfterTax_t} \times \mathrm{D_c^s}$ are presented in eFigure 1 for the outcome variable at risk for OWOB.

$$\mathrm{Y_{ict}} = \alpha_0 + \sum_{s\in\{L,M,H\}} \beta^s(\mathrm{D_c^s} \times \mathrm{AfterTax_t}) + \mathrm{X'_{ict}}\delta + \gamma_i + \mathrm{t} + \epsilon_{\mathrm{ict}},$$

where i indexes an individual, c a city, and t a year. $\mathrm{Y_{ict}}$ is an indicator for being at risk for OWOB. $\mathrm{AfterTax_t}$ is an indicator variable that equals 1 for years 2014-2017, and D_c^S are indicator variables that equal 1if city c had a price change between 2013 and 2014 of below 5% (s=L), 5-10% (s=M), and at least 10% (s=H). $\mathrm{X'_{ict}}$ is a vector of covariates, which include age indicators, categories for the number of medical visits per year, and indicator variables if an adolescent was diagnosed with endocrine or digestive diseases each year. All regression models included a vector of individual indicators, γ_i, (i.e., individual fixed effects) to adjust for potential confounding from time-invariant individual (e.g., genetics, preferences, baseline differences in socioeconomic characteristics), clinic or city-level factors, and a linear trend, t. β is the key coefficient of interest and measures a difference in outcome variables before vs after the tax for cities with a low/medium/high price change. Standard errors are clustered at the city level. These models are estimated for all teenagers and by sex and estimated coefficients (β^s), multiplied by 100, are plotted in eFigure 1.

Sensitivity analyses

We describe our sensitivity analyses in eTables 3-4. To control for differential trends in outcomes across regions, we include state time trends by including a dummy variable for the region of the country where city *c* is located φ_s, interacted with a linear time trend, *t*, thus effectively adding the $\varphi_s t$ term to equation (2). We consider four regions (South, North and North Pacific, Central Pacific and Highlands, Gulf Coast, and Yucatan Peninsula and South).

Results remain significant at the 5% level and of a similar magnitude (eTable 4, Concern *i*)). For other sensitivity analyses we modified the specific sample on which we run the regressions. Specifically, we verified whether results are robust to including prices of other high calorie foods (namely pastries, cookies, cake, candies, cereals, fries, savory snacks). Our findings remained largely unchanged (Concern *ii*, eTable 4). We also reproduced the analysis excluding children experiencing conditions likely influencing outcomes irrespective of prices. We excluded patients with a primary diagnosis related to neoplasms or diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism; identified by Chapters II or III (C00-D89) using the ICD-10 codes and our findings remain mostly unchanged (Concern iii, eTable 4), and we controlled for the number of (yearly) appointments with a primary diagnosis of digestive issues to which our results are also robust (Concern i*v*, eTable 4).

Finally, to address the concern of spurious correlation, we applied a falsification test by using an alternative outcome that should arguably be unaffected by the SSBs tax or soda prices—in particular, whether patients visited a clinic due to an external cause (i.e. injuries and/or accidents) as our outcome variable, Y_{ict} (Concern *v,* eTable 4). A positive association between soda prices and external cause diagnoses would suggest our results are driven by other unobservable variables or secular trends in health outcomes. The insignificant association and a null coefficient lend support to our finding that higher soda prices are associated with lower prevalence of overweight or obesity among teens. eTable 5 presents results for alternative weight-related outcomes for heavier teens, as concerns with using BMI in adolescents exist, particularly among those with severe obesity.[3]

---

[3] Lokling et al. Monitoring children and adolescents with severe obesity: body mass index (BMI), BMI z-score or percentage above the International Obesity Task Force overweight cut-off? Acta Paediatr, 2019; Woo. Using body mass index Z-score among severely obese adolescents: a cautionary note. Int J Pediatr Obes 2009; Zaniqueli et al. Ponderal index classifies obesity in children and adolescents more accurately than body mass index z-scores Pediatr Res 2019; Gomez-Campos et al. Accuracy of Body Mass Index Cutoffs for Classifying Obesity in Chilean Children and Adolescents. Int J Environ Res Public Health 2016.

Tables

eTable 3. Association between soda prices and changes in weight-related outcomes for 1999-2002 birth cohort between 2013-2017 using alternative model specifications.[1]

	All			Girls			Boys		
Outcomes:	**BMI (percentiles)**	**At risk for OWOB (≥75th pct)**	**OWOB (≥85th pct)**	**BMI (percentiles)**	**At risk for OWOB (≥75th pct)**	**OWOB (≥85th pct)**	**BMI (percentiles)**	**At risk for OWOB (≥75th pct)**	**OWOB (≥85th pct)**
Prices included as regressors:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: All (Unique patients= 12654)	Reported below are coefficients for Log(Price Soda)t-2, multiplied by 10								
Model 1: Log(Price of soda) at t-2	-0.148	**-0.830***	**-0.786***	-0.481	**-1.344***	**-1.065***	0.291	-0.167	-0.410
	(-0.458 - 0.162)	(-1.512 - -0.147)	(-1.471 - -0.102)	(-0.997 - 0.036)	(-2.364 - -0.324)	(-1.977 - -0.153)	(-0.371 - 0.954)	(-1.072 - 0.739)	(-1.320 - 0.501)
Model 2: Log(Price of soda) at t-1 and t-2	-0.137	-0.662	**-0.752***	-0.452	**-1.273****	**-1.260****	0.244	0.081	-0.130
	(-0.459 - 0.185)	(-1.336 - 0.013)	(-1.442 - -0.063)	(-0.944 - 0.040)	(-2.192 - -0.355)	(-2.137 - -0.383)	(-0.375 - 0.863)	(-0.839 - 1.001)	(-1.111 - 0.850)
Model 3: Log(Price of soda) and other prices at t-2	-0.120	**-0.813***	-0.677	-0.422	**-1.319****	-0.865	0.295	-0.124	-0.388
	(-0.466 - 0.226)	(-1.563 - -0.062)	(-1.398 - 0.044)	(-0.845 - 0.002)	(-2.184 - -0.455)	(-1.773 - 0.043)	(-0.458 - 1.048)	(-1.303 - 1.056)	(-1.422 - 0.646)
Model 4: Log(Price of soda) and other prices at t-1 and t-2	-0.238	**-0.848***	**-0.911***	**-0.489***	**-1.381****	**-1.246****	0.096	-0.150	-0.447
	(-0.616 - 0.140)	(-1.538 - -0.158)	(-1.600 - -0.222)	(-0.896 - -0.081)	(-2.089 - -0.673)	(-2.017 - -0.475)	(-0.669 - 0.861)	(-1.386 - 1.085)	(-1.559 - 0.665)
Panel B: 2013 vs 2016/2017 (Unique patients=12654)	Reported below are coefficients for Log(Price Soda)t-2, multiplied by 10								
Model 5: Log(Price of soda) at t-2	-0.217	**-1.248***	**-0.966***	-0.597	**-1.970****	**-1.458****	0.286	-0.294	-0.304
	(-0.698 - 0.264)	(-2.223 - -0.273)	(-1.859 - -0.072)	(-1.279 - 0.085)	(-3.272 - -0.668)	(-2.534 - -0.383)	(-0.611 - 1.183)	(-1.698 - 1.109)	(-1.646 - 1.038)
Model 6: Log(Price of soda) at t-1 and t-2	-0.213	**-1.059***	**-0.906***	-0.478	**-1.775****	**-1.604****	0.098	-0.173	-0.059
	(-0.719 - 0.292)	(-2.053 - -0.066)	(-1.783 - -0.029)	(-1.161 - 0.205)	(-3.090 - -0.461)	(-2.568 - -0.641)	(-0.664 - 0.859)	(-1.491 - 1.145)	(-1.519 - 1.401)
Model 7: Log(Price of soda) and other prices at t-2	-0.102	-1.068	-0.537	-0.389	**-1.621****	-0.674	0.321	-0.252	-0.255
	(-0.703 - 0.500)	(-2.171 - 0.035)	(-1.632 - 0.557)	(-1.129 - 0.351)	(-2.820 - -0.422)	(-1.824 - 0.477)	(-0.796 - 1.437)	(-2.106 - 1.601)	(-1.943 - 1.432)
Model 8: Log(Price of soda) and other prices at t-1 and t-2	-0.295	-0.969	-0.772	-0.408	**-1.352***	**-1.094***	-0.115	-0.413	-0.316
	(-0.911 - 0.320)	(-2.020 - 0.083)	(-1.745 - 0.200)	(-1.041 - 0.224)	(-2.384 - -0.321)	(-1.961 - -0.227)	(-1.151 - 0.922)	(-2.265 - 1.439)	(-1.939 - 1.308)

[1] Coefficients are reported for a two-year log-transformed soda prices. Regressions include individual fixed effects, age and year fixed effects, categories for the number of individual visits to a clinic in a year, and indicator variables if the teenager has endocrine or digestive diseases as reported. 95% confidence intervals calculated using clustered robust standard errors at the city level presented in parentheses. *** p<0.001,** p<0.01,*p<0.05. Health outcomes in these regressions are measured in 2013 and beyond, and price data are measured with a two-year lag, starting in 2011 and beyond. Sample includes a balanced panel of patients during 2012-2017, observed at least once per year and up to ten times per year in urban areas for the teenage cohort born between 1999-2002. Coefficients marked in bold are deemed significant at p<0.05. All coefficients are multiplied by 10, as we interpret all our results at the average price increase of 10%. Total observations: 63,270 (all), 34,250 (girls), and 29,020 (boys).

[2] At risk for OWOB and OWOB are defined at the 75^{th} or 85^{th} BMI percentile or above, respectively.

eTable 4. Association between soda prices and changes in weight-related outcomes using alternative model specifications to address endogeneity concerns[1]

	All			Girls			Boys		
	BMI (percentiles)	**At risk for OWOB (≥75th pct)**	**OWOB (≥85th pct)**	**BMI (percentiles)**	**At risk for OWOB (≥75th pct)**	**OWOB (≥85th pct)**	**BMI (percentiles)**	**At risk for OWOB (>75th pct)**	**OWOB (>85th pct)**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CONCERN	Reported below are coefficients for Log(Price Soda)t-2, multiplied by 10								
i) Regional trends in outcomes[2]	-0.054	-0.472	-0.637	-0.356	**-1.048***	**-1.173***	0.309	0.230	0.024
	(-0.308 - 0.200)	(-1.145 - 0.202)	(-1.334 - 0.060)	(-0.819 - 0.108)	(-1.935 - -0.160)	(-2.070 - -0.276)	(-0.279 - 0.898)	(-0.726 - 1.186)	(-0.978 - 1.025)
ii) Other prices driving associations[3]	-0.238	**-0.848***	**-0.911***	**-0.489***	**-1.381****	**-1.246****	0.096	-0.150	-0.447
	(-0.616 - 0.140)	(-1.538 - -0.158)	(-1.600 - -0.222)	(-0.896 - -0.081)	(-2.089 - -0.673)	(-2.017 - -0.475)	(-0.669 - 0.861)	(-1.386 - 1.085)	(-1.559 - 0.665)
iii) Restricting to no neoplasms[4]	-0.042	-0.547	-0.706	-0.396	**-1.353***	**-1.358****	0.343	0.336	0.004
	(-0.372 - 0.288)	(-1.246 - 0.152)	(-1.421 - 0.008)	(-0.933 - 0.141)	(-2.440 - -0.266)	(-2.330 - -0.385)	(-0.262 - 0.948)	(-0.600 - 1.272)	(-0.918 - 0.927)
iv) Controls for total visits for digestive conditions[5]	-0.138	-0.665	**-0.754***	-0.453	**-1.278****	**-1.258****	0.243	0.080	-0.134
	(-0.460 - 0.184)	(-1.340 - 0.010)	(-1.442 - -0.066)	(-0.945 - 0.040)	(-2.196 - -0.359)	(-2.137 - -0.379)	(-0.375 - 0.861)	(-0.841 - 1.001)	(-1.112 - 0.844)
	Outcome: external diseases:			**Outcome: external diseases:**			**Outcome: external diseases:**		
v) Spurious correlations/secular trends in health outcomes[6]	0.030			0.071			-0.010		
	(-1.016 - 1.075)			(-1.164 - 1.306)			(-1.229 - 1.209)		

[1] Coefficients are reported for a two-year log-transformed soda prices. Regressions include individual fixed effects, age and year fixed effects, categories for the number of individual visits to a clinic in a year, and indicator variables if the teenager has endocrine or digestive diseases as reported . 95% confidence intervals calculated using clustered robust standard errors at the city level presented in parentheses. *** p<0.001,** p<0.01,*p<0.05. Health outcomes in these regressions are measured in 2013 and beyond, and price data are measured with a two-year lag, starting in 2011 and beyond. Coefficients marked in bold are deemed significant at p<0.05. All coefficients are multiplied by 10, as we interpret all our results at the average price increase of 10%. Total observations: 63,270 (all), 34,250 (girls), and 29,020 (boys).

[2] i) Model addressing this concern is equivalent to model 2 in eTable 3, but also includes region specific trends to adjust for regional trends in outcomes.

[3] ii) Model addressing this concern is equivalent to model 2 in eTable 3, but also includes a two-year lag in log-transformed prices of other foods (pastry, cookies, cake, candies, cereals, fries, savory snacks).

[4] iii) Model addressing this concern is equivalent to model 2 in eTable 3, but excludes patients that ever had a diagnosis related to neoplasms or other disorders of the blood (ICD codes C00-D89). The number of distinct patients is N=11,096

[5] iv) Model addressing this concern is equivalent to model 2 in eTable 3, but controls for *total number visits* (per year) due to digestive diseases (ICD codes K00-K95).

[6] v) Model addressing this concern is equivalent to model 2 in eTable 3 with an outcome defined as a binary variable for diseases unlikely to be affected by changes in diet: external diseases, deformations, and others.

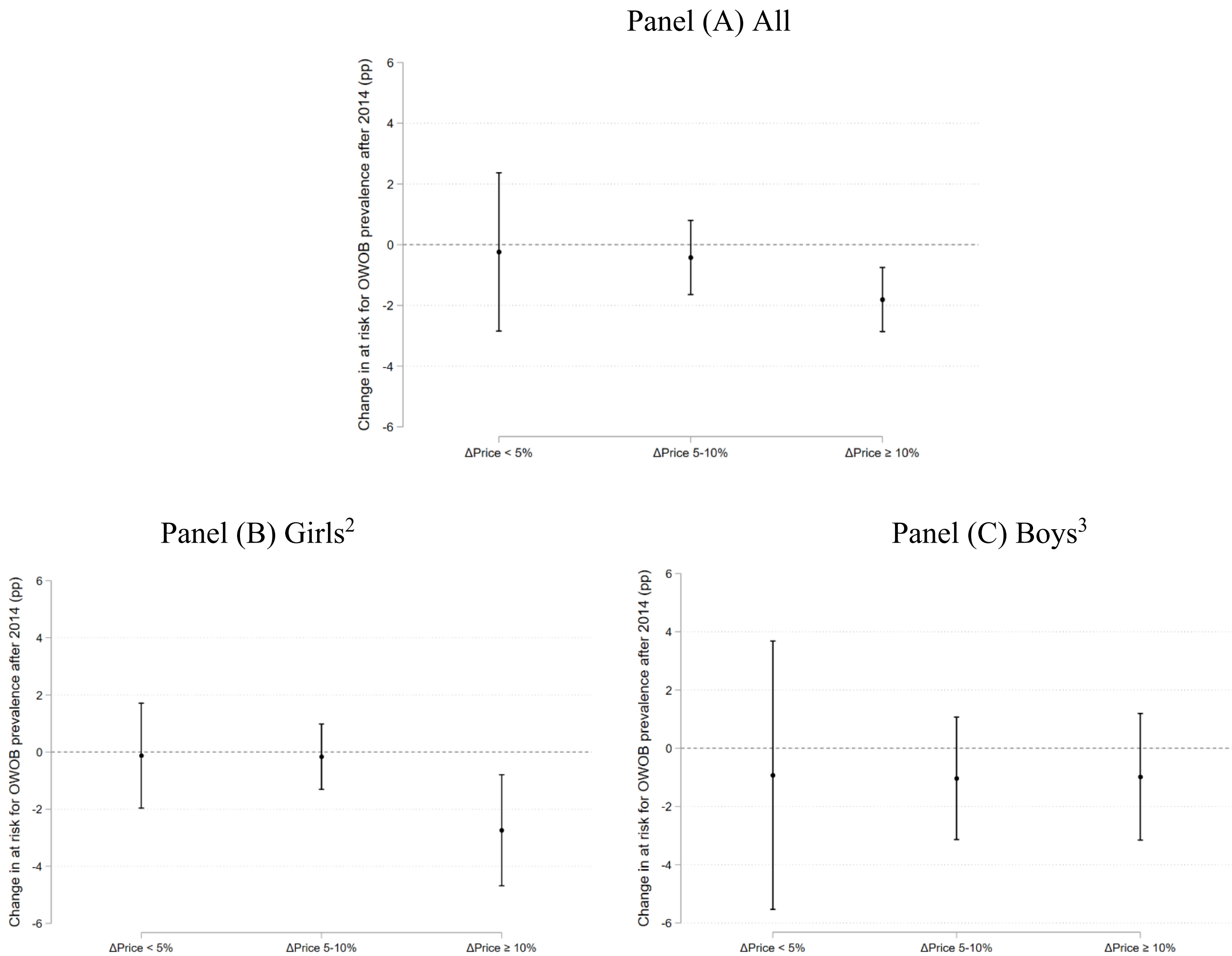
eTable 5. Association between soda prices and changes in weight-related outcomes using alternative measures of body mass by weight status prior to tax [1]

	Girls				Boys			
	z-BMI	**Log(BMI)**	**Ponderal Index**	**Log(weight)**	**z-BMI**	**Log(BMI)**	**Ponderal Index**	**Log(weight)**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not at risk for OWOB pre-tax:								
Log(Price of soda) $_{t-1}$	0.136	0.120	1.471	0.095	1.833	0.143	-0.035	0.242
95% CI	(-3.012 - 3.284)	(-0.350 - 0.590)	(-6.377 - 9.319)	(-0.372 - 0.563)	(-2.334 - 6.001)	(-0.393 - 0.680)	(-8.444 - 8.373)	(-0.252 - 0.736)
Log(Price of soda) $_{t-2}$	-1.022	-0.219	-2.224	-0.180	1.305	0.143	2.165	0.173
95% CI	(-3.542 - 1.498)	(-0.559 - 0.121)	(-8.885 - 4.438)	(-0.510 - 0.150)	(-2.496 - 5.106)	(-0.384 - 0.670)	(-6.758 - 11.088)	(-0.319 - 0.665)
Mean	-0.24	17.62	12.10	37.78	-0.23	17.23	12.03	36.07
Unique patients	3,180				2,363			
At risk for OWOB pre-tax:								
Log(Price of soda) $_{t-1}$	-0.275	-0.148	-2.836	-0.200	-0.788	-0.109	-1.280	-0.024
95% CI	(-2.928 - 2.379)	(-0.564 - 0.269)	(-11.331 - 5.660)	(-0.634 - 0.233)	(-3.870 - 2.294)	(-0.649 - 0.431)	(-12.255 - 9.696)	(-0.573 - 0.524)
Log(Price of soda) $_{t-2}$	**-2.643***	-0.378	-6.681	-0.410	0.629	-0.000	0.448	-0.128
95% CI	(-4.821 - -0.465)	(-0.804 - 0.047)	(-14.102 - 0.741)	(-0.839 - 0.019)	(-1.598 - 2.856)	(-0.428 - 0.427)	(-8.348 - 9.244)	(-0.573 - 0.316)
Mean	1.54	23.23	15.55	52.62	1.73	22.86	15.32	51.62
Unique patients	3,670				3,441			

[1] Coefficients are reported for a two-year log-transformed soda prices. All coefficients are multiplied by 10, as we interpret all our results at the average price increase of 10%. Regressions include individual fixed effects, age and year fixed effects, categories for the number of individual visits to a clinic in a year, and indicator variables if the teenager has endocrine or digestive diseases as reported. Log(height) is included as a control in the Log(weight) regressions. 95% confidence intervals calculated using cluster robust standard errors at the city level are presented in parentheses. *** $p<0.001$,** $p<0.01$,*$p<0.05$. Health outcomes in these regressions are measured in 2013 and beyond, and price data are measured with a two-year lag, starting in 2011 and beyond. Sample includes a balanced panel of patients during 2012-2017, observed at least once per year and up to ten times per year in urban areas for the teenage cohort born between 1999-2002. Coefficients marked in bold are deemed significant at $p<0.05$. z-BMI represents BMI z-scores calculated according to Vidmar, Cole & Pan (2004). Log(BMI) is log transformed BMI. Ponderal Index is weight/height^3. Log weight is log-transformed weight in kilograms. Total observations: 18,350 (girls at risk for OWOB pre-tax), 15,900 (girls not at risk for OWOB pr-tax), 17,205 (boys at risk for OWOB pre-tax), 11,815 (boys not at risk for OWOB pre-tax). and 29,020 (boys). Girls at risk for OWOB pre-tax's outcomes for Log(Price of soda)$_{t-2}$ are statistically significant at the 10% level: z-BMI (p-value: 0.019), Ponderal Index (p-value: 0.076), Log(BMI) (p-value: 0.080), and Log(weight) (p-value: 0.061).

Figures

eFigure 1. Changes in weight-related outcomes before vs after the tax across cities with different price changes [1]

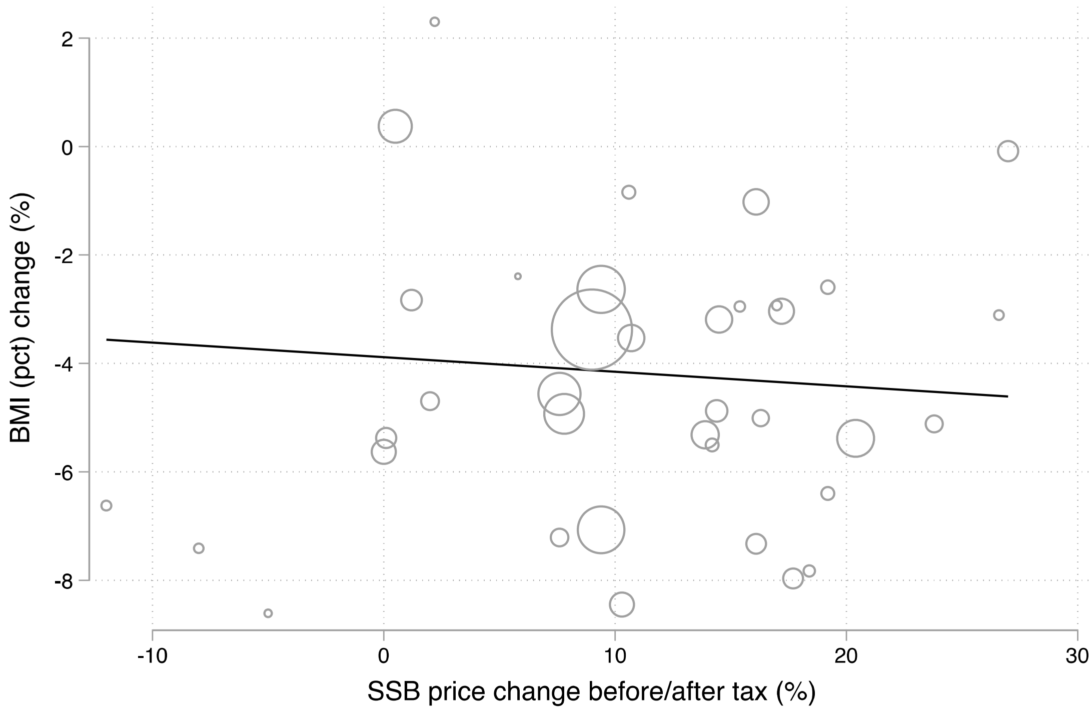


[1] Panel A shows the change in prevalence for teenagers at risk for OWOB in years 2014 and after (vs years prior) for cities with different price change between 2013 and 2014. Coefficients show the interaction term between an after-tax indicator variable that equals one in calendar years 2014-2017 and indicator variables for cities in which percentage change in SSB prices between 2013 and 2014 was below 5% (ΔPrice <5%), between 5-10% (ΔPrice 5-10%), or at least 10% (ΔPrice≥10%). They are multiplied by 100 to interpret them as percentage point (pp) change in the outcome after the tax. Sample is restricted to patients observed six times during 2012-2017, at least once per year in urban areas for the teenage cohort born between 1999-2002. Regressions include individual fixed effects, age fixed effects, a linear time trend, categories for the number of individual visits to a clinic in a year, and indicator variables if the teenager has endocrine or digestive diseases during the year.

[2] Panel B shows the change in prevalence of girls who are at risk for OWOB. The only significant coefficient corresponds to cities in which SSB price changes before vs after the tax were at least 10%, implying a 2.7 pp (95% CI, -4.68 to –0.08, P<0.01) absolute, or about 4.8% relative decrease in prevalence of at risk for OWOB post-tax for girls exposed to these price changes. ΔPrice≥10% is significantly different than ΔPrice 5-10% (p-value: 0.025) and ΔPrice <5% (p-value: 0.018).

[3] Panel C shows the change in prevalence of boys who are at risk for OWOB.

eFigure 2. Ecological association between SSB price changes and changes in BMI percentile[1]



[1] Each circle represents a city. The x-axis shows the percent change in real prices of regular soda before vs after the tax, calculated by multiplying the estimate on post-tax-city indicator by 100 and the y-axis shows the percentage change in mean BMI (percentile) between 2013 and 2017. Size of circles is weighted by the number of patients observed at city-level. Cities with fewer than ten patients are excluded from the figure (two cities). Corr=-0.778