Short-Term Health Effects of Soda Taxes in Areas with Low Access to Safe Drinking Water*

Emilio Gutierrez† ITAM

Adrian Rubli‡ ITAM

This version: October 2019
First draft: May 2017

We ask whether taxing sugary beverages may have negative health effects in areas where clean water is unavailable. Focusing on a tax in Mexico, we find a significant but localized 7% increase in outpatient gastrointestinal disease rates throughout the first year of the tax, with evidence of a declining impact by the end of the second year. We provide evidence of avoidance behavior by affected households through differential consumption of bottled beverages two years post-tax. The costs implied by our results are small. However, our findings inform the need for accompanying soda taxes with interventions that guarantee safe drinking water.

JEL codes: I18, I15, H23
Key words: sin taxes; obesity; water quality; gastrointestinal diseases; avoidance behavior

*We thank Anna Aizer, Jay Bhattacharya, Janet Currie, Andrew Foster, Joshua Graff Zivin, Michelle Marcus, Emily Oster, Jody Sindelar, and seminar participants at the Frontiers of Health Economics Research in Latin America workshop (IHEA 2017), Latin American and Caribbean Economic Association (LACEA 2017), Northeastern Universities Development Consortium (NEUDC 2018), and ITAM-UCSD conference for their helpful comments. Dr. Isabel Martínez at the Mexican Ministry of Health (SSA-DGE) provided invaluable help in accessing, processing, and understanding the health datasets. Both authors acknowledge support from the Asociación Mexicana de Cultura. All errors are our own.

†Instituto Tecnológico Autónomo de México, Department of Economics. Camino a Santa Teresa 930. Colonia Heroes de Padierna, Magdalena Contreras CDMX 10700, Mexico. Email: emilio.gutierrez@itam.mx

‡Corresponding author. Instituto Tecnológico Autónomo de México, Department of Business Administration. Camino a Santa Teresa 930. Colonia Heroes de Padierna, Magdalena Contreras CDMX 10700, Mexico. Email: adrian.rubli@itam.mx

1 Introduction

Supported by a growing literature (see Cawley, 2015 and Allcott et al., 2019, for reviews), policy-makers and academics have advocated for taxing sugary beverages (SBs), known colloquially as soda taxes, as a way to contain the world-wide obesity epidemic. For example, the World Health Organization actively promotes these taxes, both in developed and developing countries (WHO, 2016). Out of a total of 39 countries that have implemented a national tax on SBs since 1940, 31 have done so over the last decade, and 22 of these correspond to developing countries (see Figure 1a). Proponents of these taxes often cite public health and economic benefits as reasons for their implementation (Brownell et al., 2009; Brownell and Frieden, 2009).

The empirical literature estimating the relationship between SB prices and consumption has focused on cross-state variations in tax rates (e.g., Fletcher et al., 2010b), SB taxes introduced in some US cities and European countries (e.g., Cawley et al., 2019), and simulations using estimated demand elasticities (e.g., Finkelstein et al., 2013).[1] However, little is known about how specific conditions in the developing world, such as low access to clean drinking water, may interact with the introduction of SB taxes and lead to unintended consequences.

Therefore, this paper fills this gap by asking whether the implementation of an SB tax in a developing country could lead to negative health effects for individuals in areas where access to safe drinking water is low.

The effectiveness of taxing SBs depends on consumers' substitution patterns. Allcott et al. (2019) emphasizes the role of substitution both across different types of SBs and other untaxed goods as key elements for designing optimal SB taxes. Some studies argue that substituting toward other high-caloric foods and beverages may diminish the impact of these taxes on dietary changes (Fletcher et al., 2010b, 2013; Aguilar et al., 2019). Others

[1] Estimating the effect on health outcomes, such as obesity, has proved to be more difficult due to the lack of a clean long-term identification strategy. Some papers, however, do find effects on obesity, albeit small ones (see, for example, Fletcher et al., 2010a).

have found evidence in favor of substitution toward non-sugary beverages, mainly water (Nakhimovsky et al., 2016; Colchero et al., 2016, 2017).

However, a large fraction of the world's population, particularly in developing countries, may not have regular and affordable access to safe substitutes. This in turn is one of the main forces behind the high prevalence of gastrointestinal diseases (GIDs) and infant mortality in these countries (Kremer et al., 2011; Ashraf et al., 2017; Dupas and Miguel, 2017).[2] Figure 1b presents a descriptive picture of the under 5 mortality rate (measured in 2016) for all countries that have implemented an SB tax by year of implementation, distinguishing between developed and developing countries. As expected, the average under 5 mortality rate for the developing countries that have implemented a tax is significantly higher than that of the developed countries.

We focus our attention on Mexico, where a nation-wide tax on SBs was introduced on January 1st, 2014.[3] To the best of our knowledge, the excise tax was not accompanied by any campaigns for clean water or public service announcements with simple strategies for disinfecting water (such as using iodine tablets). This particular setting - the first large-scale soda tax in a developing country - provides a unique opportunity to explore this question.

Despite being a middle-income country, many regions in Mexico still lack widespread access to piped water and have substandard surface water quality (CONAGUA, 2016; DHAyS, 2017). According to the 2010 census, 37% of households do not have piped water at home. The under 5 death rate due to diarrhea in Mexico is 0.42 per thousand live births, below other developing countries such as the Philippines' 1.95 and South Africa's 3.60, but four times higher than that for the US at 0.09 (UNICEF, 2016).

[2]For example, 57% of all GIDs worldwide are attributed to water and sanitation issues, and diarrheal disease accounts for 20% of all deaths of children under five (Prüss-Ustün et al., 2016). Furthermore, there is evidence in favor of implementing targeted policies. Interventions that improve drinking water, access to sanitation, and hygiene reduce GID morbidity by 45, 28, and 23%, respectively (Freeman et al., 2014).

[3]Since 2010, Mexico has implemented many policies to tackle the rising obesity rates. Some of these strategies include introducing healthy food options at schools, regulating the marketing of high-caloric food items to children, and requiring a clear, front-of-package labeling of nutritional facts on all foods and beverages (Barquera et al., 2013).

Existing literature has documented that the SB tax in Mexico led to a roughly 10% increase in SB prices (Colchero et al., 2015; Grogger, 2017; Aguilar et al., 2019), and a 6% reduction in consumption (Colchero et al., 2017; Aguilar et al., 2019). Furthermore, Colchero et al. (2016) shows that for low-income households, consumption of taxed beverages declined by 9% post-tax, while untaxed beverages, mostly bottled water, increased by only 2%.[4] These numbers suggest that low-income households may have substituted toward non-bottled drinking water, which might be detrimental if unclean.

We obtain data on piped water access at the electoral precinct level from the census, data on surface water quality from government monitoring stations, and health data from all public outpatient clinics.[5] We use Thiessen polygons to extrapolate water quality measures from over 2,000 monitoring stations to the whole country. To answer our question, we first identify areas with low access to safe drinking water. Due to data restrictions, we assume that all piped water is sufficiently clean for human consumption. This is in line with studies on a large piped water chlorination program introduced in Mexico in the 1990s (Bhalotra et al., 2017). We therefore focus on areas with low access to piped water *and* bad surface water quality as those where an increase in consumption of non-bottled water due to the tax may lead to negative health effects.

Using a difference-in-differences framework, we contrast GID rates over time in clinics located in areas with low access to piped water and bad surface water quality against all other clinics in Mexico. Our specifications include time period fixed effects to account for seasonality, and estimate the impact from changes within clinics over time by including clinic fixed effects. We provide supportive evidence that the parallel pre-trends assumption holds, and show that our main results are robust to including additional controls, alternative definitions of the treatment areas, and placebo checks on unrelated conditions.

[4]Higher income households decreased their SB consumption by 6% and increased their consumption of other beverages by 4%. It should be noted that the low income households in their panel correspond to those in urban areas, since the scanner data they use are unavailable for rural populations.

[5]Electoral precincts are the smallest administrative unit in Mexico (over 64 thousand in 2010).

We document a statistically significant but localized effect of the SB tax on GID rates of 6.6% for individuals in our treatment areas during the first year of the tax, equivalent to about 64 thousand additional GID cases in 2014 relative to 2013. We find mixed evidence for the second year, with much smaller point estimates, some of which are not statistically significant at conventional levels. This suggests that the increase in GID rates may have tapered off by the end of 2015. We do not find any effects on hospitalization rates, indicating that although the soda tax increased outpatient GID rates, these did not translate into complications leading to inpatient care.

To shed light on the pattern of the effects over time, we present evidence of avoidance behavior, suggesting that the estimated effect is short-lived due to households in our treatment areas differentially increasing their consumption of bottled beverages relative to control households two years after the tax. This seems to be driven by both changes in soda and bottled water consumption.

To contextualize our results, we find that at most 92 thousand GID cases in the first two years could be attributed to the tax, at a cost of around 4.75 million USD. This cost, relative to the SB tax revenues and to the potential health gains from dietary changes, is small. Therefore, these results do not warrant an argument against introducing these taxes.

However, they do indicate that in contexts where individuals lack safe drinking water, SB taxes may have unintended consequences, highlighting the importance of understanding substitution patterns as discussed by Allcott et al. (2019). This issue may be more salient in countries where more households lack clean water and mortality rates due to diarrhea are higher, such as the Philippines and South Africa, where SB taxes were introduced in

2018.[6] We recommend accompanying these taxes with aggressive, targeted policies aimed at guaranteeing clean water for vulnerable individuals.

We are not the first to suggest a link between soda consumption and diarrheal disease in contexts of low access to safe drinking water. In a review of the potential effects of soda taxes, Roache and Gostin (2017) recognizes the possibility of negative impacts in areas without clean water. Onufrak et al. (2014) presents the first piece of empirical evidence, documenting that Hispanics in the US that mistrust their local tap water are twice as likely to consume SBs than those who perceive it to be clean.

The paper most similar to ours is Ritter (2019). This study analyzes soda price changes across different regional markets in Peru, documenting that a decrease in soda prices is positively correlated with soda consumption and negatively correlated with self-reported GIDs. We believe that our paper complements these findings in at least three ways. First, our results inform about the generalizability of the relationship between GIDs and soda prices. Second, we exploit the introduction of a tax on SBs as the source of variation in prices and consumption. Apart from the fact that a national tax is more likely exogenous to local economic conditions, we believe that it is particularly important to directly explore whether these taxes – which are becoming increasingly popular and are primarily aimed at improving long-term health outcomes – could have undesired consequences in specific settings. Finally, by exploiting high-frequency administrative data at a fine geographic level, we are able to explore in detail the dynamics in medically-diagnosed GIDs.[7]

[6] The WHO has been a vocal proponent of the SB taxes introduced in both countries (https://www.reuters.com/article/us-philippines-health/philippine-taxes-on-sugary-drinks-could-avert-thousands-of-deaths-who-study-says-idUSKBN1O41NP, last accessed March 5, 2019; http://www.afro.who.int/en/south-africa/press-materials/item/9347-who-supports-proposed-sugar-sweetened-beverages-tax-in-south-africa.html, last accessed August 28, 2018). Given that the under 5 GID mortality rate is five and nine times higher in each country than in Mexico, the potential health costs in the absence of policies that guarantee safe drinking water could be far greater.

[7] Related work has also analyzed the interaction between breastfeeding and water cleanliness in developing settings. Keskin et al. (2017) finds that Bangladeshi women breastfeed longer in areas without access to safe drinking water, while Anttila-Hughes et al. (2018) documents increases in infant mortality rates following the introduction of baby formula in areas where availability of clean water is low.

Our main contribution is to add to the literature identifying the link between SBs and diarrheal disease, by causally estimating the impact of changes in SB prices on GID rates. In particular, we are the first to link taxes on SBs to increases in GIDs in areas with low access to safe drinking water. The fact that these excise taxes have – until recently – only been implemented and analyzed in developed countries poses a challenge for the generalizability of findings to developing settings, and we contribute by filling this gap. Given that around 60% of the world's obese population lives in the developing world (Ng et al., 2014), and to the extent that SB taxes may prove to be a useful tool to combat obesity, it is necessary to understand the link between these taxes and GIDs where safe drinking water is lacking.

The remainder of the paper is organized as follows. Section 2 presents some context. Section 3 describes the data sources. Section 4 introduces the identification strategy. Section 5 presents the results. Section 6 provides evidence of avoidance behavior. Section 7 discusses our findings. Section 8 concludes.

2 Background

2.1 Mexico's Soda Tax

Mexico has long struggled with obesity, with 40% of adults considered overweight, and a third obese (National Health Survey ENSANUT, 2012). In late 2013, the Mexican Congress approved a fiscal reform, effective January 1st, 2014. An important new item included in the reform was an excise tax on sugary beverages (SBs) as part of the Special Tax on Production and Services (IEPS, by its Spanish acronym). The reform established that all SBs in the country would now be subject to a 1 peso (0.06 USD) per liter tax, which on average amounted to about 10 to 12% of the average price (Colchero et al., 2016).

IEPS defines SBs as sodas, nectars and concentrates with added sugar, and powdered drink mixes. Beverages sweetened with non-caloric sugar substitutes and dairy products were exempt from the tax. Although many types of SBs are taxed, sodas garnered the most

media attention, and the tax is commonly referred to as simply the "soda tax" (*impuesto a los refrescos*). We refer to it as such throughout this paper.

Grogger (2017) analyzes the effect of the soda tax on the average price of SBs, finding a passthrough of over 100%. Additional studies corroborate this finding (Aguilar et al., 2019; Colchero et al., 2015).[8] Colchero et al. (2016) calculates the effect of the tax on the consumption of taxed beverages, concluding that they declined 6% on average (around 12 mL per capita per day, equivalent to a regular-sized can of soda per month). Aguilar et al. (2019) confirms the magnitude of this decline, but also finds that substitution patterns across all food and beverage purchases led to a null effect on total calories consumed.[9] Table A1 in the online appendix provides a review of the literature in this setting.

2.2 Beverage Consumption Patterns in Mexico

Previous literature has characterized beverage consumption patterns in Mexico over the past years. Stern et al. (2014) uses dietary recall surveys in 1999 and 2012 to describe trends in caloric beverages consumed by demographic groups, finding that SBs increased among both children and adults. Strikingly, this study calculates that in 2012, SBs accounted for 17.5% of the total daily caloric intake for children and adolescents ages 1 to 19. Barquera et al. (2010) also finds that pre-school and school children obtain 28 and 21% of their energy from caloric beverages, respectively.

We address two concerns. First, anecdotally, the market for bottled water in Mexico is large.[10] Therefore, one may worry that individuals in Mexico are only consuming bottled wa-

[8]Note that total passthrough of the tax has not always been observed in other settings. For example, Cawley and Frisvold (2017) finds a passthrough of only 43% for the Berkeley, CA tax, attributed to consumers' avoidance behavior by making SB purchases in other jurisdictions. This is a particular advantage of the nation-wide implementation of the tax in Mexico.

[9]Notwithstanding these declines, there is still a sizable amount of taxed beverages being consumed. For example, in 2014, the total tax revenues from the soda tax amounted to 18 billion pesos, which averages to 163 liters of SBs per capita or about 1.25 regular-sized cans of soda per day (http://finanzaspublicas.hacienda.gob.mx/es/Finanzas_Publicas/Estadisticas_Oportunas_de_Finanzas_Publicas, last accessed May 13, 2017).

[10]See for example https://www.forbes.com.mx/agua-embotellada-el-negocio-multimillonario-que-mexico-no-necesita/, last accessed September 11, 2018.

ter, rendering other water sources innocuous. To address this, we present evidence of trends in purchasing behavior of beverages prior to the tax, exploiting data from the nationally-representative 2012 National Household Income and Expenditures Survey (ENIGH).

Table 1 shows summary statistics for households' purchases of both taxed drinks and bottled water prior to the tax. Taxed drinks include all sodas and energy drinks; bottled water includes all presentation sizes of bottled water (including large 20-liter jugs), and club soda.[11] We stratify our descriptive statistics into terciles by total household income.

Table 1 shows that 67% of high socioeconomic status (SES) and 56% of low SES households purchased soda over the last week prior to the survey, compared to 44 and 23% of households making purchases of bottled water, respectively. Low SES households purchased about a third of the volume in liters of bottled water and more than half the liters of soda purchased by high SES households. These statistics show that not all households purchase bottled water, and less so if they fall on the left side of the SES distribution. Note that in contrast to the ENIGH data, retail panels typically do not include rural areas, where bottled water consumption is less common.

The second concern relates to whether SBs and water could be substitutes in this setting. We rely on evidence provided by Colchero et al. (2016). This study uses retail panel data to show that the consumption of taxed beverages fell by 6% on average, and up to 9% among low-income households. Untaxed beverages on the other hand increased by 4% on average, and just 2% for low-income households. Note that the data in Colchero et al. (2016) does not include non-bottled water consumption. Therefore, given their findings – low-income households experienced both a larger decrease in SBs and a smaller increase in untaxed beverages – it is likely that some households substituted toward non-bottled drinking water, particularly since bottled water consumption is not as prevalent among low SES households (see Table 1).

[11] We exclude juice since the survey only distinguishes between "natural" and bottled juice, while the tax applies only to juices with added sugar.

Additional evidence is provided in Table A2 in the online appendix, which shows the distribution of individuals reporting that their water and SB consumption went down, stayed the same or went up, for the top and bottom SES terciles, as reported in the 2016 ENSANUT.[12] Although we recognize that this is an imperfect measure, we believe that it sheds light on the possibility that water and SBs are substitutes to some degree in our context. Table A2 indicates that the majority of the increase in water consumption in the two years after the tax was implemented corresponds to individuals who decreased their SB consumption. Furthermore, the numbers suggest that substitution occurred across all SES groups.[13]

3 Data

3.1 Water Access and Quality

We gather data on households' access to piped water from the 2010 census. This information is representative at the electoral precinct level.[14] For each precinct, we observe the fraction of households in 2010 obtaining their water from a source outside the home. This includes households obtaining piped water from a neighbor or a communal tap as well as non-piped water (from vendors, surface water sources, such as rivers, lakes, and dams, or from wells). See online appendix B for more information.

We assume that all piped water in Mexico is sufficiently safe for consumption. In general, publicly available data sources measuring tap water quality are not available, especially at a sufficiently disaggregated level and with enough cross-sectional variation. Furthermore, research has shown that a large national water program in Mexico in the early 1990s saw

[12] The 2016 ENSANUT was a special round of the usual health survey administered every six years. This particular round focused on questions regarding consumption of beverages, which had not been previously recorded.

[13] Since the 2016 ENSANUT is also a nationally-representative survey, the top and bottom terciles of the SES distribution here are comparable to those of the ENIGH reported in Table 1.

[14] Administratively, Mexico is divided into 32 states, which are in turn composed of municipalities, with a total of 2,456 in the whole country. For electoral purposes, municipalities are further divided into smaller geographic areas called precincts (64,559 in our 2010 data). This is effectively the smallest administrative unit in Mexico, with an average of 383 households per precinct, a median of 320, and a maximum of 18,125.

important impacts in chlorinating piped water, in turn associated with large declines in child mortality (Bhalotra et al., 2017). We therefore shift our water quality focus toward surface water sources.

We obtain surface water quality data from government monitoring stations throughout the country belonging to the regulator CONAGUA (National Water Commission). Data are available for 2012, 2013 and 2014, with a single observation per station per year. From a total of 3,610 monitoring stations, we exclude stations that are located at salt water sources. This leaves us with 2,071 stations in our sample. All stations are geocoded.[15]

These data provide three distinct quality measures: biochemical oxygen demand (BOD), chemical oxygen demand (COD), and total suspended solids (TSS). For each, CONAGUA reports the precise measure, as well as a classification into five categories (very polluted, polluted, acceptable, good, and excellent), based on CONAGUA's established thresholds. See online appendix B for more details on water quality.

3.2 Health Outcomes

The health outcomes measuring outpatient cases and rates of gastrointestinal diseases (GIDs) come from the Ministry of Health's Reported Cases Dataset, obtained directly from SSA's offices. This information is collected by the Ministry of Health (SSA) on a weekly basis, and contains all new GID diagnoses at the outpatient clinic level. Each public outpatient clinic is legally required to report this information. Although some private clinics also report this data, their compliance rates are extremely low and there is no way to corroborate the data they may misreport. Therefore, we restrict our attention to public outpatient clinics.

The public healthcare system in Mexico is divided into separate, disjoint subsystems targeting different segments of the population. We restrict these data to the four principal

[15] We are unable to precisely characterize what determines the government's location choice for these stations, although their location was chosen in 1996 and is supposed to provide "reliable and representative information". See official government documents available at http://dgeiawf.semarnat.gob.mx:8080/ibiapps/WFServlet?IBIF_ex=D3_R_AGUA05_03&IBIC_user=dgeia_mce&IBIC_pass=dgeia_mce (last accessed October 14, 2019).

subsystems: SSA (through *Seguro Popular* insurance), IMSS, IMSS-Oportunidades, and ISSSTE.[16] This amounts to 15,634 clinics, with the excluded ones making up around 1% of public healthcare services (ENSANUT 2012). Note that individuals eligible for health services from a given subsystem are automatically assigned to the clinic closest to their home, and are expected to seek care only at that location.[17] All clinics are geocoded by merging information from SSA's publicly-available Infrastructure Dataset for 2014.

The Reported Cases data records GIDs directly from doctors' diagnoses, based on ICD-10 codes. We use SSA's classification of GIDs, including all ICD-10 codes from A00 to A09. This is consistent with the literature (in particular for Mexico, see Agüero and Beleche, 2017). We also collect data on other unrelated diagnoses for a placebo exercise.

We also obtain hospital discharge records for a subset of public hospitals, corresponding to those administered directly by SSA. These publicly-available data contain each patient's date of admission, as well as the final diagnosis based on ICD-10 codes. There are 766 SSA hospitals in this dataset. Unfortunately, the hospital discharge data for the other public subsystems only registers the year, and not the actual dates of hospitalization, and do not provide ICD-10 codes. It should be noted however, that SSA tends to provide healthcare to lower SES groups, and that according to the 2012 ENSANUT, 40% of all hospitalizations occurred at an SSA hospital.[18]

[16] IMSS provides healthcare for formal workers and their families; IMSS-Oportunidades is the rural branch of IMSS, linked to the cash transfer program Oportunidades; ISSSTE corresponds to government workers; and *Seguro Popular* provides coverage for informal workers and the unemployed, through SSA's own network of clinics and hospitals. The remaining smaller subsystems are for workers of the national oil company, the marines, and the army.

[17] Unless there is a life-threatening condition, individuals may only obtain healthcare at the subsystem and clinic they have been assigned to.

[18] Private hospitalizations account for 15% of all inpatient care, which means that almost half of all public hospitalizations are at an SSA hospital.

4 Empirical Strategy

4.1 Assigning Water Characteristics to Clinics

To answer the question of whether the soda tax may negatively impact diarrheal disease, we need to be able to observe both access to safe drinking water and GID rates. Our characterization of safe drinking water will rely on both the access to piped water measure and the surface water quality measures from the monitoring stations. We begin by geographically assigning these measures. We then map them to the outcomes observed at each clinic.

First, we observe access to tap water inside the home according to the 2010 census at the electoral precinct level. For each public outpatient clinic, we assign the fraction of households without access to piped water inside the home from the precinct in which the clinic is located. Although clinic catchment areas may extend beyond the limits of the precinct, this method allows us to consistently assign the access variable without making any assumptions about the size and shape of catchment areas.

Figure 2a shows a map of this variable at the precinct level. Figure B5 in online appendix B shows the location of clinics in their corresponding electoral precinct. Note that it is possible to have more than one clinic per precinct (or no clinics), especially given the healthcare subsystems targeting different segments of the population. There are on average 1.4 clinics per precinct, with a median of one, conditional on having a clinic.

We now turn to the quality measures. We observe measures from 2,071 monitoring stations located at fresh water sources. These stations are unevenly distributed across Mexico. To extrapolate the data, we create Thiessen polygons. The idea is to partition the plane into disjoint areas – one for each monitoring station – such that all coordinates within an area are closer to the corresponding monitoring station than any other station. We then assign quality measures to each clinic from the Thiessen polygon in which the clinic is located.

Effectively, each clinic is assigned the quality measures from the monitoring station that is closest to it. Three maps in Figure 2 show these Thiessen polygons as well as the spatial

distribution of the BOD, COD, and TSS measures, taking the average over 2012 and 2013. Note that we exclude 2014 since the tax was implemented in that year. Figure B2 in online appendix B further shows the location of these monitoring stations. Figure B6 in the online appendix shows the distribution of the distance between the clinics and their assigned station, with a range from 0.04 to 199.67 km. As a robustness check, we exclude clinics for which the assigned monitoring station is at a relatively large distance.

We collapse our weekly clinic data into quarterly observations by summing over the 13 weeks that make up a quarter. This reduces some of the natural noise in GIDs due to small fluctuations from one week to another. Altogether, we are left with a balanced panel of 15,634 clinics $\times$ 28 quarters.

4.2 Defining Treatment and Control Groups

Having assigned the water variables to each clinic, we now turn our attention to identifying suitable treatment and control groups. The treatment should be the areas where an increase in water consumption could lead to more diarrheal disease. Due to the lack of reliable data on tap water quality, we assume that all piped water is sufficiently safe for human consumption. This is a reasonable assumption, as a large water program in the early 1990s greatly improved chlorination levels of piped water in Mexico. Bhalotra et al. (2017) finds important declines in infant mortality due to this program.

Therefore, our hotspots of unsafe drinking water are those that lack access to piped water *and* have bad quality surface water.

Since we are incorporating two dimensions into our treatment status definition that are measured in different units, we need to summarize our multiple continuous measures into a single classification. For the sake of transparency, we stick to a simple binary classification for the main specification, and show results on sensitivity checks that vary the definition of our treatment group, including a continuous score of availability of safe drinking water.

Piped water access f_e measures the fraction of households in an electoral precinct e that do not have access to piped water inside the home. Since each clinic c is mapped onto a single precinct, we can substitute the subscript e for c. Figure B1 in online appendix B shows the distribution of this variable for all precincts in our sample. We classify clinics into those with low and high access to tap water based on the median of this distribution f^{p50}. Formally, allowing $\mathbb{1}_{[\cdot]}$ to represent the indicator function,

$$Low_c = \mathbb{1}_{[f_c > f^{p50}]}$$

Note that we use the distribution across precincts and not clinics, since we want our classification to correspond to household-level risk. For our sensitivity analysis, we also calculate our results using the mean of this distribution instead.

The surface water quality measures are a bit more complex, since there are three different measures (BOD, COD, and TSS) and two different pre-treatment years (2012 and 2013) to consider. We include all three measures assuming that together they minimize any noise and strengthen the signal of the true underlying water quality. Likewise, we include both years of available data to reduce noise and increase power. Note that including 2014 data would only be a problem if the tax coincided with differential changes in water quality in treatment areas relative to the control. To the extent that this may be unlikely and since more data may improve the signal strength, we also present results including the 2014 measures as part of our sensitivity analysis.

Let m_{sjy} be the measure of type j in year y at monitoring station s, where type is BOD, COD, or TSS. We begin by taking the average over the pre-tax years:

$$\bar{m}_{sj} = \frac{1}{N_{sj}} \sum_{y} m_{sjy}$$

where N_{sj} is the number of years for which there is a non-missing value of m_{sjy} for measure j at station s. Whatever the underlying true quality is, taking this two-year average will tend to minimize any measurement error in a particular year.

Next, we use the median of the distribution of each two-year average measure for BOD, COD, and TSS across monitoring stations to classify into good and bad surface water quality. Note that these measures have a heavily skewed distribution, yielding a mean that is significantly higher than the median (see Table B2 in online appendix B). In our sensitivity analysis, we use the official thresholds provided by the government regulator CONAGUA for this binary classification. Let m_j^{p50} represent the median of the two-year average of measure type j across all stations s. Then formally we have:

$$M_{sj} = \mathbb{1}_{[\bar{m}_{sj} > m_j^{p50}]}$$

We now have three binary measures per monitoring station.

Finally, to summarize into a single binary measure of water quality, we conservatively take the minimum over these three measures:

$$Bad_s = \min M_{sj}, \quad \forall j \in \{\text{BOD, COD, TSS}\}$$

This means that a clinic only has bad quality assigned if all three binary summaries of BOD, COD, and TSS fall above their respective median.

Having summarized the access variable and the quality variables into two binary measures, we define treatment clinics as those for which access to piped water is low and surface water quality is bad. Since each clinic is assigned one monitoring station, we can substitute the s subscript for c, such that:

$$Treat_c = Low_c \times Bad_c$$

This gives us a total of 1,596 treated clinics. We use all other clinics in the sample as our control group (14,038 units).

Figure 3a shows a map of our treatment and control areas using the baseline definition of treatment. Additional maps show alternative definitions for the treatment areas that make up our sensitivity analysis. Figure 3b considers the official thresholds provided by the government regulator CONAGUA for the binary classification of water quality at each monitoring station, Figure 3c includes the 2014 water quality data, and Figure 3d uses the mean of the distribution of piped water access as the cutoff instead of the median. These maps show that although there are similarities in the treatment and control classifications across definitions, there is also significant variation. We will show that the main results are not sensitive to the definition of treatment that we use.

Table 2 shows summary statistics at the clinic level for the treatment and control clinics under our main definition. Panel A shows clinic outcomes and characteristics. Panel B presents clinic-level descriptives of household characteristics. Lastly, Panel C shows clinic-level descriptives of water quality from monitoring stations, as well as the location of the stations by water source.

As expected by construction, clinics in treatment and control areas differ importantly by access to piped water inside the home and by water quality measures from the assigned monitoring station. We also find important differences in terms of access to electricity, sewerage and a bathroom inside the home, with treatment clinics significantly poorer in this regard. This is also related to the fact that treatment clinics are more likely to be SSA clinics. We find small differences in the water source for the monitoring stations, and importantly we detect no difference between our treatment and control clinics in terms of the distance to its assigned monitoring station. Overall, Table 2 suggests that treatment clinics are in low SES and more rural areas relative to control clinics. Crucially for us, our identification strategy will rely on assuming (and partially testing for) similar trends over time across groups of clinics, not similar levels.

4.3 Identification

We are interested in estimating the effect of the soda tax on outpatient GID rates, using areas with both low access to piped water *and* bad quality surface water as our treatment, and places with either high access to piped water and/or good quality surface water as our control. Our strategy relies on the tax having had an impact on the consumption of SBs.

Given our data and the lack of an obvious control group, we cannot present direct evidence of the first stage effect of the tax on prices and consumption. However, we address this in two ways. First, we rely on previous literature. Table A1 in online appendix A summarizes those findings. These studies consistently find an increase of around 10% in SB prices (Colchero et al., 2015; Grogger, 2017; Aguilar et al., 2019), and a 6% decrease in consumption during the first year of the tax (Colchero et al., 2016, 2017; Aguilar et al., 2019).

Our second piece of evidence comes from the 2008, 2010, 2012, 2014 and 2016 ENIGH survey rounds. We regress weekly soda purchases per household on indicators for each year, the log of total household income, indicators for whether the house has access to electricity, sewerage and a bathroom inside, indicators for household head's years of education, indicators for household size, fraction of household members that are adults, indicators for size of the locality, and municipality fixed effects. The results show an 8-9% decline in soda purchases from 2012 to 2014, and an additional 2.5% from 2014 to 2016 (see Figure A1 in online appendix A). While these estimates are purely descriptive, they echo the main findings in the literature.

We now turn to the question of estimating the impact on GIDs in areas with low access to safe drinking water. We follow a difference-in-differences (DD) specification for this purpose. We begin with a simple DD equation of the following form:

$$rate_{ct} = \beta(Treat_c \times post_t) + \lambda_c + \theta_t + \varepsilon^0_{ct} \tag{1}$$

where $rate_{ct}$ is the GID rate per 100,000 at public outpatient clinic c in year-quarter t, $Treat_c$ is an indicator for whether the clinic is in a low piped water access and bad surface water quality area, $post_t$ is an indicator for post-tax years 2014 and 2015, λ_c are clinic fixed effects, θ_t are year-quarter dummies, and ε^0_{ct} is the error term. We winsorize the outcome variable at the 5% in our main specifications to address potentially spurious outliers. Robustness checks will consider more conservative winsorization levels. Standard errors are clustered at the clinic level to allow for serial correlation.

The coefficient of interest β denotes the differential change in GID rates in treatment vs control areas, before vs after the tax was implemented. The main identifying assumption is that control and treatment clinics follow similar trends over time. When the treatment is a binary variable, then β identifies a weighted average of the treatment effect in each clinic and at each period (de Chaisemartin and D'Haultfoeuille, 2019).

To provide a clear dynamic picture of the effects and to estimate the treatment effects at different points in time (which also allows for a direct inspection of the common trends assumption), we estimate the following equation as our main specification:

$$rate_{ct} = \sum_{\tau=2009}^{2015} \beta_\tau (Treat_c \times \mathbb{1}_{[y=\tau]}) + \lambda_c + \theta_t + \varepsilon^1_{ct} \tag{2}$$

where $\mathbb{1}_{[\cdot]}$ is the indicator function, y indexes years, and everything else is as defined above.

Our coefficients of interest are now given by β_τ, particularly for 2014 and 2015, as they represent differential changes in GID rates for treatment clinics relative to our control for each year. The clinic and time period fixed effects imply that we are estimating changes within clinics over time, net of overall seasonal effects.

Additional specifications include other controls. First, we include household characteristic controls by interacting an indicator for each year-quarter with indicators for being in quartile q of the precinct-level distribution of access to electricity, sewerage, and a bathroom inside the home. This will account for any differential trends in areas that are more rural or low

SES. Second, we include geographic controls where we interact indicators for each year-quarter with an indicator for a grid cell defined by latitude and longitude degrees. These controls will absorb any weather shocks or regional epidemiological shocks.

Once again, the key identifying assumption for a causal estimate is that trends in GID rates would be the same in both treatment and control clinics in the absence of the soda tax, regardless of any differences in levels, captured by the clinic fixed effect. The fact that we observe multiple periods prior to the implementation of the tax allows us to partially test for the validity of the common trends assumption, by identifying whether the coefficients β_τ during the pre-tax period are small and statistically indistinguishable from zero.

To further analyze the time trends, we also present estimates from a more dynamic DD approach. We estimate the following equation analogous to equation 2:

$$rate_{ct} = \sum_{\tau=1}^{T} \beta_\tau (Treat_c \times \mathbb{1}_{[t=\tau]}) + \lambda_c + \theta_t + \varepsilon_{ct}^2 \tag{3}$$

where T represents the total number of quarters in our panel (28 quarters from 2009 to 2015), and everything else is as defined above. Once again, we are interested in the β_τ coefficients, which now represent differential changes in GID rates for treatment clinics relative to our control for each quarterly time period. This specification allows us to distinctly analyze the parallel pre-trends assumption. It also allows us to see the post-tax dynamics more clearly.

The DD estimators would be biased if clinics "self-select" into our groups, as defined by access to tap water and surface water quality. By using the average quality over the 2012 to 2013 period, we are minimizing this issue. The effect will also be confounded if other policies that affected GID rates or healthcare in general were introduced in 2014.[19] To the best of our knowledge, this was not the case. Moreover, this would only be an issue if these additional policies affected our treatment and control clinics differentially. We provide

[19] A related concern would be the presence of unrelated GID outbreaks in the treatment areas. We were unable to find any evidence of this in the media, and the inclusion of our additional geographic controls should account for this.

evidence against this by performing a placebo check on unrelated diagnoses. Additionally, we use staffing and infrastructure data at the clinic level to check for differential changes in healthcare supply between our treatment and control groups in online appendix C. These results provide convincing evidence that we are not confounding the effect of the tax with supply-side changes.

An additional concern is measurement error in our variables. We limit extreme values by winsorizing the outcome variable, and present alternative winsorization levels as robustness checks. However, our outcome effectively measures GID cases at outpatient clinics, not the overall epidemiological prevalence of GIDs. This may be a concern only if the likelihood of seeking medical care at the clinic conditional on being sick changes differentially over time between treatment and control clinics. Using available survey data, we show that this is not the case in online appendix D.

In terms of our independent variable, we may be misclassifying clinics along the piped water access and surface water quality dimensions due to the different moments in time when our data were collected. First, we argue that over such a short time period, any gains in these dimensions must be relatively small. As such, any measurement error is likely to be classical, with an equal probability of classifying the clinic correctly or incorrectly, thus leading to attenuation bias. Second, if improvements in access and quality are indeed sufficiently large, then we would tend to classify high access/good quality clinics as low access/bad quality. This, too, would attenuate our results.

5 Results

5.1 Main Effect on Outpatient GID Rates

We begin by estimating the simple DD described in equation 1 on our balanced panel of 15,634 clinics over 28 quarters. We show this result in column 1 of Table 3. Our coefficient of interest – the interaction between our indicator for treatment and an indicator for the

post-tax period – is positive and statistically significant. Given the mean GID rate per 100,000 in treatment clinics prior to the tax, we estimate a differential 3.28% increase in the rate of GIDs in areas with low access to piped water and bad quality surface water, relative to the rest of the country, which amounts to a total of 40 additional GID cases per treated clinic over this two-year period.[20]

In order to provide evidence on the common trends assumption and to unmask any heterogeneity in the average treatment effects over time, we estimate the dynamic DD in equation 2, treating the interaction with 2013 as the excluded category. The remaining columns in Table 3 show the point estimates for the β_τ coefficients. Columns 2-4 consider the main treatment definition provided above, with each subsequent column including additional controls. Columns 5-8 present estimates from alternative definitions of treatment.

We present the β_τ coefficients in reverse chronological order, distinguishing between post and pre-tax years. The baseline specification in column 2 shows a large, positive and significant effect for the interaction of the treatment and 2014 indicators. Given the mean GID rate per 100,000 in treatment clinics during the pre-tax years, this effect amounts to a differential 6.6% increase in GIDs from 2013 to 2014 in treatment clinics relative to control clinics, or an average of 40 additional GID cases per treated clinic, consistent with the bulk of the effect concentrated during the first year of the tax. The estimate for 2015 is smaller and statistically indistinguishable from zero. Taking the point estimate, this amounts to 17.5 additional cases per treated clinic in the second year of the tax. A test allows us to reject that the post-tax estimates are equal. Focusing on the pre-tax estimates, they are all statistically zero, show no particular trend, and are about four times smaller than the 2014 estimate. We cannot reject that these pre-tax estimates are all simultaneously zero, allaying any concerns about existing pre-trends and providing us with reassurance that our estimates may be interpreted as causal.

[20] According to our data, the average population for each clinic is 139,495 individuals. Therefore, to calculate the total additional cases take the estimate 3.597 $\times$ 4 quarters $\times$ 2 years $\times$ 1.39 individuals.

To further make sense of the magnitudes, we calculate the local average treatment effect (LATE). This is the effect for compliers, individuals who would only drink unsafe water if they live in our treatment areas after the tax. The LATE can be obtained by dividing our estimates by the share of compliers in the treatment. Since surface water quality is the same for all households at a given location, the compliers are determined by access to piped water. From Table 2, the share of households in treatment areas without piped water inside the home is 88%. Given our estimated 6.6% increase in 2014, the LATE is a 7.5% increase in the GID rate. If instead we focus on the share of households obtaining water from surface water and wells (29% in treatment areas), the LATE would be a 23% increase, while considering only those who report using surface water (7% in treatment areas) would result in a LATE of 89%.[21]

These magnitudes are in line with the literature. In the study most similar to ours, Ritter (2019) calculates that a 10% decline in the price of sodas in Peru is associated with a 0.5 percentage point decrease in self-reported diarrheal disease. Given the baseline prevalence, this amounts to a 71% decline. In a different setting but with clinic-level data, Ashraf et al. (2017) finds that for a 24-day period of piped water outages in Lusaka, there are 23.7 additional GID cases at the local clinic, equivalent to a 16% increase given the baseline mean. Focusing on infant mortality, Anttila-Hughes et al. (2018) finds that the introduction of baby formula increased mortality by 10%, while Bhalotra et al. (2017) shows that the program that chlorinated piped water in Mexico in the early 1990s led to an average 50% decline in mortality, and up to 80% in areas with good quality pipes.

We present additional specifications in columns 3 and 4 of Table 3. Since treatment clinics are in areas where households are more likely to lack access to electricity, sewerage, and a bathroom inside the home, one may worry that the estimated effects are confounding the tax with contemporaneous changes over time related to these characteristics. To address this,

[21]Note that the census only asks households their main source of water. Our assumption is that any household with piped water inside will report it as their main source. However, households without piped access inside may obtain water from a variety of sources, leading to potential misreporting. As such, we always consider the share without piped water inside the home as a more reliable measure.

we include time-varying household controls in column 3 of Table 3, by interacting indicators for each year-quarter with indicators for being in quartile q of the precinct-level distribution of these household characteristics.

An additional concern is that the estimated effects may be driven by weather shocks in treatment areas, such as changes in rainfall or temperatures. Alternatively, there may other regional shocks that may be resulting in differential epidemics in these areas that coincided with the tax. To address this, we add time-varying geographic controls in column 4, by interacting indicators for each year-quarter with indicators for each latitude-longitude grid cell, defined by degrees only.[22]

When adding the household controls, the estimate in column 3 of Table 3 for the interaction of the treatment and 2014 indicators is positive, significant, and slightly larger than the baseline estimate in column 2. Adding geographic controls leads to a very similar estimate in magnitude and significance relative to the baseline effect. The estimates for the interaction with 2015 are slightly larger, and are now significant at the 90% level. Across both specifications, we can reject that the 2014 and 2015 effects are the same, and we cannot reject that all pre-tax coefficients are jointly equal to zero (and that each individual coefficient is equal to zero).

We present a sensitivity analysis to our definition of treatment areas in columns 5-8 of Table 3. Column 5 considers government-established thresholds for the binary classification of water quality.[23] Column 6 includes water quality data from 2014. Column 7 uses the mean of the precinct-level distribution of access to piped water for the binary classification of access. Lastly, column 8 creates a continuous index using principal components analysis, and standardizing so that the index has a mean of zero and a standard deviation of one.[24]

[22]Given the location of Mexico, these grid cells measure approximately 103 by 110 kilometers, with a total of 5,908 cells spanning the entire country. Note also that including these controls results in 20 control clinics dropped from the estimation procedure due to being singletons within each grid cell - quarterly date.

[23]The regulator CONAGUA follows a five-tier classification system ranging from "excellent" to "very polluted". We use the "good" category as our cutoff for good quality (see Table B3 in online appendix B).

[24]Figure B7 in online appendix B shows the distribution of this continuous index separately for treatment and control clinics as defined at baseline. The plot shows that there is considerable overlap between the two, although there is a clear mean-shift between them.

The results of our sensitivity analysis show that the effects are stable and similar across specifications. For the binary definitions in columns 5-7, the estimated effect in 2014 ranges from a 4.7 to a 6% differential increase in GID rates, relative to the 6.6% at baseline. With the exception of column 5, the 2015 estimates are insignificant and statistically different from the ones for 2014. Using the continuous index, the estimates in column 8 show that for a one standard deviation change in the treatment variable, GID rates increase by 0.6 in 2014. The smaller magnitudes obtained under this continuous metric suggest that this measure is less precise in classifying exposure. However, we obtain positive and significant coefficients for both 2014 and 2015. Across all specifications, the pre-tax coefficients are individually and jointly indistinguishable from zero.

To visualize the effects over time more clearly, Figure 4 shows graphical representations of the estimates of equation 3, with the last quarter of 2013 as the excluded category. We present the estimates for eight quarters before and after the introduction of the tax for clarity. For each plot, we include error bars that indicate 95% confidence intervals. The darker coefficient series corresponds to the baseline specification, using the main treatment definition and without including any additional controls.

Figure 4a contrasts the estimates between the baseline specification and specifications that include additional controls (see above). Figure 4b compares the baseline specification with specifications that use alternative definitions of treatment. For both plots, the point estimates for the two years prior to the tax are small and statistically insignificant. The estimates then become positive and (mostly) significant throughout the first five quarters after the tax. By the end of 2015, the estimates are once again small and indistinguishable from zero. These results echo the findings in Table 3.

Overall, Table 3 and Figure 4 show that clinics in treatment areas experienced a differential increase in GID rates in 2014, ranging from 4.7 to 7.5%. The evidence for 2015 is much weaker, with coefficients that are smaller, less significant, and statistically different from the 2014 estimates. This suggests that the effect was short-lived, and we explore a potential

mechanism in Section 6. We also show that results are stable and robust to the inclusion of controls and to alternative treatment definitions. Lastly, our pre-tax coefficients provide supporting evidence that the common trends assumption holds.

5.2 Robustness Checks on the Main Effect

We now present a series of robustness checks on the main results presented above. We present three main exercises: exclude clinics that are far away from their assigned monitoring station, use alternative winsorization levels, and perform placebo tests on unrelated diagnoses that should be unaffected by the tax in our treatment areas.

First, we consider excluding clinics based on distance to the assigned water monitoring station. We exclude the top 99th, 95th and 90th percentiles of the clinic-level distribution of distance to the station, and estimate equation 2. Table 4 presents these results in columns 2-4, while column 1 replicates the baseline result for reference. The findings are consistent with the main results: large and significant coefficients for the first year of the tax, smaller and insignificant estimates for 2015 that we can reject are equal to the 2014 estimates, and smaller and insignificant coefficients for the pre-tax years, for which we cannot reject that they are all jointly zero. Figure E1 in online appendix E presents similar graphical results from estimating equation 3.

Second, we winsorize our outcome variable at more conservative levels, and estimate equation 2. Column 5 of Table 4 winsorizes the data at the 2.5%, while column 6 winsorizes at the 1%. The main results hold under these alternatives. We find somewhat larger magnitudes for the 2014 effect: a 7.5% differential increase in column 5, and 8.1% in column 6. Once again, there are no significant impacts in 2015. For the pre-tax years, we cannot reject that the coefficients are jointly zero at the 95% level (although we reject the null at the 90%). Figure E1 in online appendix E also presents graphical results for this exercise from estimating equation 3.

Lastly, we perform a series of placebo checks on conditions unaffected by the soda tax. Since other infectious diseases may be affected by changes in GIDs (Agüero and Beleche, 2017), we abstain from using them in this analysis. Instead, we first consider accidents and external conditions, including all ICD-10 codes from S00 to T98.[25] We then consider sexually transmitted diseases (STDs), including ICD-10 codes from A53 to A60, as well as B18, B20 and B97.[26] Finally, we include chronic conditions, defined as ICD-10 codes E11-E14, I10-I15, and M15-M19.[27] We also present results for the sum of these three placebo conditions.

Table 5 shows the results from estimating equation 2 on these outcomes, which have also been winsorized at the 5%. Each of the first three columns corresponds to a different condition, while column 4 considers the sum of all three. We find no significant effects for the first year of the soda tax. We also find no evidence of any significant increases in 2015. As with the previous checks, we present graphical results for the placebo tests from estimating equation 3 in Figure E2 in online appendix E. Overall, these placebo checks suggest that our results are not confounded by other healthcare policies that might have increased public supply or general demand for healthcare. We present further evidence in online appendix C that there was no differential change in clinic staffing and infrastructure from 2013 to 2014 in treatment clinics relative to our controls.

5.3 Effect on Hospitalizations

Our main results correspond to the effect of the tax on GID rates at public outpatient clinics. We now turn our attention to a worse health outcome: hospitalizations. Our goal is to assess whether the increase in GID rates at the outpatient level were severe enough to necessitate inpatient care. As outlined in Section 3, we focus on SSA hospitals due to data availability, keeping in mind that they represent 40% of all hospitalizations and about half of all public sector hospitalizations.

[25] This classification covers conditions such as fractures, sprains, dislocations, open wounds, burns, and drug overdoses, among others.

[26] This covers STDs such as syphilis, gonorrhea, HPV, and HIV, among others.

[27] These codes correspond basically to diabetes, hypertension, and osteoarthritis.

We assign piped water access and surface water quality measures to each hospital as before, resulting in a balanced panel of 766 hospitals $\times$ 28 quarters. We consider two outcomes. First, we use hospitalization rates due to GIDs, winsorized at the 5%. We distinguish between all hospitalizations, and hospitalizations of children under 6 years old, since they are the most at-risk group for dehydration. Second, we calculate the average length of hospital stays for GIDs only and for all hospitalizations. We winsorize this outcome at the 1%. The first measure captures the extensive margin, while the second focuses on the intensive one. Our rationale for including all hospitalizations in one of the length of stay measures is that more diarrheal disease may also impact the severity of illness for those who were hospitalized for reasons unrelated to GIDs.

Table 6 presents the results from estimating equation 2 on this dataset. Across all four columns (corresponding to the four outcomes outlined above), we find insignificant and small coefficients for the post-tax years. We also find insignificant effects during the pre-tax years. Figure E3 in online appendix E shows the analogous estimates from equation 3 graphically.

Taken together, the findings in Table 6 indicate that the SB tax had no discernible effect on hospitalization rates and length of stay in areas with low access to tap water and bad surface water quality relative to the control. Although the increase in soda prices did lead to more GIDs, as evidenced by our main findings, these outbreaks were successfully controlled and contained at the outpatient level.

6 Evidence of Avoidance Behavior

Although there is strong evidence of an effect in 2014, the evidence for a continued effect in 2015 is much weaker and the point estimates are smaller. Some specifications indicate that the effect died out two years after the tax, while others simply do not allow us to reject that the magnitude of the effect is the same in both years. A plausible explanation is that individuals learn about their local water quality and adjust through some avoidance behavior.

Potentially, these individuals could be switching back to bottled beverages after they realize that water made them sick.[28] Although constrained by data availability, we present some suggestive evidence that individuals without access to safe drinking water switched back to consuming bottled beverages two years after the tax was introduced.

We use repeated cross-sections of household data from the 2008, 2010, 2012, 2014 and 2016 ENIGH rounds. A disadvantage of these data relative to the data used above is that we only observe each household's location at the municipality level. An advantage over retail panels is that we can observe rural areas, and that these surveys are representative at the national level. Another advantage is that we now observe access to piped water at the household level.

We assign water quality as before, aggregating up to municipalities. Effectively, we take the weighted average of the measures defined by the Thiessen polygons based on how they overlap with each municipality. Although important heterogeneity within municipalities may be lost, this is the only alternative given the granularity of the data. For piped water access inside the home, this survey records access to tap water for each household. We make use of this valuable information by defining piped water access directly at the household level. Interacting these two variables provides us with our treatment indicator as before.

We estimate equation 2 on the repeated cross-sections. We substitute the clinic and year-quarter fixed effects with municipality and survey round fixed effects, respectively. We consider three outcomes of interest. First, we calculate the liters of bottled water consumed by a household in a given municipality and survey year. Bottled water includes all presentation sizes of regular water (including large 20-liter jugs), as well as all types of sparkling water and club soda. Second, we calculate the liters of all taxed drinks, including sodas and energy drinks. The survey does not distinguish between regular and diet sodas, which were not taxed. However, sodas with artificial sweeteners represent a small fraction of total soda

[28] Avoidance behavior by increasing consumption of bottled water has been identified in other contexts. See, for example, Graff Zivin et al. (2011).

consumption.[29] Furthermore, the survey only distinguishes between "natural" and bottled juice, while the tax applies only to juices with added sugar. Therefore, we exclude all juices. Lastly, we consider the sum of bottled water and taxed beverages. All outcomes are transformed using the inverse hyperbolic sine function, and standard errors are clustered at the municipality level.

Table 7 presents the results. Even-numbered columns include additional household-level controls (log income, whether the house has access to electricity, sewerage, and a bathroom inside, indicators for household head's years of education, indicators for household size, fraction of members that are adults, and indicators for the size of the household's locality). The first two columns correspond to all bottled beverages (water and SBs). We find small and insignificant coefficients for the interaction of the treatment indicator and 2014. This indicates that households in treatment areas did not differentially change their consumption of bottled beverages in the first year of the tax.

However, we find large, positive, and significant coefficients for the interaction with 2016. This indicates that households in treatment areas differentially increased their consumption of bottled beverages by 18-19% relative to households in control areas. A test allows us to reject that the 2014 and 2016 effects are equal. We also find, reassuringly, small and insignificant coefficients for the pre-tax survey rounds, and a test of coefficients does not allow us to reject that they are all jointly zero.

Columns 3 and 4 of Table 7 explore the effects for bottled water only. We find a similar pattern, with large and significant effects only in 2016 that indicate a differential increase of 17-20%. However, our test does not allow us to reject that the 2014 and 2016 coefficients are the same magnitude, even though the former are less than half the size of the latter. Columns 5 and 6 consider taxed beverages only. We find no significant effects. The point

[29] For example, diet sodas produced by Coca Cola in Mexico represent only 4% of their total soda sales (https://www.elfinanciero.com.mx/empresas/refrescos-sanos-le-dan-punch-a-coca-cola-mexico, last accessed October 16, 2019).

estimates for 2016 are positive, while those for 2014 are negative. This results in a significant difference between these coefficients.

Overall, the results in Table 7 suggest that households in areas without access to safe drinking water differentially increased their consumption of bottled beverages two years after the tax, consistent with the decline in the effect on GIDs that we estimated above. We cannot conclusively determine whether the effect is driven by bottled water or SBs. An alternative explanation for the short-lived effect of the soda tax on GIDs is that doctors at public clinics inform GID patients of simple measures they can take to decrease their likelihood of infection. We conducted a few informal interviews with public clinic doctors, confirming that it is common practice for doctors to share simple strategies, such as boiling water or using disinfectants, with GID patients.[30] Official government guidelines from SSA also suggest that this is the case.[31] Unfortunately, we cannot confirm individuals' knowledge on water quality pre-tax, although it seems unlikely that they are well-informed.

7 Discussion

We have identified an effect of the soda tax on GIDs for individuals in areas with low access to piped water and bad surface water quality. Taking the estimates presented in Table 3, we calculate a differential 4.7-7.5% increase in GID rates in treatment areas during the first year of the tax, and a differential 1.7-5.6% increase in the second year. Based on the point estimates from the baseline specification (Table 3 column 2), this amounts to 40 additional GID cases per clinic in 2014, and 17.5 in 2015. Given the number of treated clinics, this is less than 64,000 additional cases in the whole country in 2014, and 28,000 in 2015.

We do not attempt to quantify the welfare gains or losses from the tax. To fully address welfare implications, we would first need to take a stance on the relevant welfare function from

[30] We conducted 12 telephone interviews with IMSS doctors asking about common recommendations for GID patients. More details are available upon request.

[31] See for example, https://www.gob.mx/salud/prensa/la-secretaria-de-salud-emite-recomendaciones-para-evitar-enfermedades-diarreicas-colera-y-golpe-de-calor, last accessed August 30, 2017.

the social planner's perspective. We would also have to quantify how consumer surplus is affected by the price increase, taking a stance on the individuals' utility function, and how it compares to the individual welfare effects of increased diarrheal disease, taking heterogeneity into account. We would also have to consider all possible externalities of SB consumption, as well as "internalities" due to consumers' lack of information or self-control. We would also need to consider the correlation between the regressivity of the tax and our estimated effects. All this would require a full estimation of the demand for SBs and for clean drinking water, which is outside the scope of this paper.

Instead, we simply contrast the cost of the increase in GIDs relative to the observed tax revenues, and relative to the potential health gains (and associated healthcare savings) from the tax. The objective of this discussion is to put our estimated effects into perspective.

We perform a simple back-of-the-envelope calculation of the average cost of a GID episode. According to the 2012 ENSANUT, GID patients at public clinics paid on average 41 pesos in transportation, fees, and medicines, and spent an average of 129 minutes getting to the clinic, waiting, and with the doctor. Therefore, with an average hourly wage of 33.40 pesos, this amounts to 112.81 pesos.[32] Assuming an additional two full days of unpaid sick leave, which is a likely upper bound, the total cost of an episode of diarrheal disease is at most 647 pesos (52 USD).

Given our estimates, the total cost of the SB tax due to GIDs during the first two years is roughly 59.4 million pesos or about 4.75 million USD. Considering the total revenue from the SB tax in 2014 alone of 18 billion pesos, this cost is at most 0.33% of the government's respective revenue.[33] This simple comparison indicates that these non-life-threatening spells could easily be covered with less than one percent of tax revenues. Furthermore, a small

[32]We use the average hourly wage in the first trimester of 2014 as reported in the National Occupation and Employment Survey (ENOE).

[33]Total IEPS revenue in 2014 was 124 billion pesos, with 18 billion pesos directly from the SB tax (http://finanzaspublicas.hacienda.gob.mx/es/Finanzas_Publicas/Estadisticas_Oportunas_de_Finanzas_Publicas, last accessed May 13, 2017).

fraction of tax revenues could be put toward policies that prevent these bouts of disease, such as awareness campaigns and subsidized iodine tablets.

We can also compare the GID costs due to the tax with the relative health gains attributed to the tax and their related healthcare savings. Estimating the direct effect of the tax on weight, and ultimately obesity-related conditions such as diabetes and heart disease, is difficult given the lack of an obvious control group and the chronic nature of obesity-related diseases. However, the literature has shown large effects on the price of taxed goods and the consumption of taxed beverages (Grogger, 2017; Colchero et al., 2016; Aguilar et al., 2019). Considering that obesity costs Mexico around 120 billion pesos per year (Molina et al., 2015), these additional GID cases represent only 0.05% of this yearly cost during the first two years of the tax.

Overall, our findings do not warrant against implementing taxes on SBs in this context or others. Nevertheless, they do inform the need for accompanying these taxes with targeted policies that guarantee affordable access to safe drinking water for populations that may be negatively impacted by the tax in order to avoid these unintended consequences. This insight may matter even more in countries where access to clean water is lower than in Mexico (for example, the Philippines and South Africa, where SB taxes were introduced in 2018), since the potential cost in those settings may be far greater. The lesson we espouse is that policies attempting to incentivize water consumption over SBs can have negative health impacts on individuals who lack access to safe drinking water.

8 Conclusion

This paper asks whether a soda tax could potentially lead to unexpected negative health impacts in areas without access to safe drinking water. Focusing on Mexico, a middle-income country where clean water is still an issue in some areas, we find a significant but temporary increase in GID rates at public outpatient clinics in places with low access to

tap water and bad surface water quality. We provide evidence suggesting that affected households differentially increased their consumption of bottled beverages two years after the tax, consistent with avoidance behavior that has been observed in other settings.

We perform a simple back-of-the-envelope calculation to provide some context for our results. We attribute an upper bound of 92 thousand additional GID cases in 2014 and 2015 to the introduction of the tax, of which 64 thousand correspond to 2014 alone. The relative cost implied by these estimates is less than one percent of the 2014 SB tax revenues. It is also incredibly small relative to the potential healthcare savings from an improved diet. Therefore, our results are definitely not an argument against the introduction of such a tax in Mexico or other contexts.

Our results do indicate however that in settings where some individuals do not have access to clean water, a soda tax may have pernicious effects on this population. As such, this paper emphasizes the issues associated with implementing policies from developed countries in a developing context, and the associated unintended consequences. Our findings inform the need for aggressive, targeted interventions that guarantee safe water access to these populations when introducing taxes aimed at incentivizing water consumption. The magnitude of this negative effect may be exacerbated in lower income countries, where a larger fraction of the population lacks regular access to clean water.

References

Agüero, J. M. and T. Beleche (2017). Health shocks and their long-lasting impact on health behaviors: Evidence from the 2009 H1N1 pandemic in Mexico. *Journal of Health Economics 54*, 40–55.

Aguilar, A., E. Gutierrez, and E. Seira (2019). Taxing to reduce obesity. Technical report, Working Paper.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019). Should we tax soda? An overview of theory and evidence. *Journal of Economic Perspectives 33*(2).

Anttila-Hughes, J. K., L. C. Fernald, P. J. Gertler, P. Krause, and B. Wydick (2018). Mortality from Nestle's marketing of infant formula in low and middle-income countries. Technical report, National Bureau of Economic Research.

Ashraf, N., E. Glaeser, A. Holland, and B. M. Steinberg (2017). Water, health and wealth. Technical report, National Bureau of Economic Research.

Barquera, S., F. Campirano, A. Bonvecchio, L. Hernández-Barrera, J. A. Rivera, and B. M. Popkin (2010). Caloric beverage consumption patterns in Mexican children. *Nutrition journal 9*(1), 47.

Barquera, S., I. Campos, and J. Rivera (2013). Mexico attempts to tackle obesity: the process, results, push backs and future challenges. *Obesity Reviews 14*(S2), 69–78.

Bhalotra, S. R., A. Diaz-Cayeros, G. Miller, A. Miranda, and A. S. Venkataramani (2017). Urban water disinfection and mortality decline in developing countries. Technical report, National Bureau of Economic Research.

Brownell, K., T. Farley, W. Willett, B. Popkin, F. Chaloupka, J. Thompson, and D. Ludwig (2009). The public health and economic benefits of taxing sugar-sweetened beverages. *The New England Journal of Medicine 361*(16), 1599–1605.

Brownell, K. D. and T. R. Frieden (2009). Ounces of prevention—the public policy case for taxes on sugared beverages. *New England Journal of Medicine 360*(18), 1805–1808.

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.

Cawley, J., D. Frisvold, A. Hill, and D. Jones (2019). The impact of the Philadelphia beverage tax on purchases and consumption by adults and children. *Journal of health economics 67*, 102225.

Cawley, J. and D. E. Frisvold (2017). The pass-through of taxes on sugar-sweetened beverages to retail prices: The case of Berkeley, California. *Journal of Policy Analysis and Management 36*(2), 303–326.

Colchero, M. A., M. Molina, and C. M. Guerrero-López (2017). After Mexico implemented a tax, purchases of sugar-sweetened beverages decreased and of water increased: Difference by place of residence, household composition, and income level. *The Journal of Nutrition*, jn251892.

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.

Colchero, M. A., J. C. Salgado, M. Unar-Munguía, M. Molina, S. Ng, and J. A. Rivera-Dommarco (2015). Changes in prices after an excise tax to sweetened sugar beverages was implemented in Mexico: evidence from urban areas. *PLoS One 10*(12), e0144408.

CONAGUA (2016). Atlas del agua en Mexico. Technical report, Comision Nacional del Agua.

de Chaisemartin, C. and X. D'Haultfoeuille (2019). Two-way fixed effects estimators with heterogeneous treatment effects. Technical report, National Bureau of Economic Research.

DHAyS (2017). Informe sobre violaciones a los derechos humanos al agua potable y al saneamiento en Mexico. Technical report.

Dupas, P. and E. Miguel (2017). Impacts and determinants of health levels in low-income countries. *Handbook of Economic Field Experiments 2*, 3–93.

Finkelstein, E. A., C. Zhen, M. Bilger, J. Nonnemaker, A. M. Farooqui, and J. E. Todd (2013). Implications of a sugar-sweetened beverage (SSB) tax when substitutions to non-beverage items are considered. *Journal of Health Economics 32*(1), 219–239.

Fletcher, J., D. Frisvold, and N. Tefft (2013). Substitution patterns can limit the effects of sugar-sweetened beverage taxes on obesity. *Preventing chronic disease 10*.

Fletcher, J. M., D. Frisvold, and N. Tefft (2010a). Can soft drink taxes reduce population weight? *Contemporary economic policy 28*(1), 23–35.

Fletcher, J. M., D. E. Frisvold, and N. Tefft (2010b). The effects of soft drink taxes on child and adolescent consumption and weight outcomes. *Journal of Public Economics 94*(11), 967–974.

Freeman, M. C., M. E. Stocks, O. Cumming, A. Jeandron, J. Higgins, J. Wolf, A. Prüss-Ustün, S. Bonjour, P. R. Hunter, L. Fewtrell, et al. (2014). Hygiene and health: systematic review of hand-washing practices worldwide and update of health effects. *Tropical Medicine & International Health 19*(8), 906–916.

Graff Zivin, J., M. Neidell, and W. Schlenker (2011). Water quality violations and avoidance behavior: Evidence from bottled water consumption. *American Economic Review 101*(3), 448–53.

Grogger, J. (2017). Soda taxes and the prices of sodas and other drinks: evidence from Mexico. *American Journal of Agricultural Economics 99*(2), 481–498.

Keskin, P., G. K. Shastry, and H. Willis (2017). Water quality awareness and breastfeeding: evidence of health behavior change in Bangladesh. *Review of Economics and Statistics 99*(2), 265–280.

Kremer, M., J. Leino, E. Miguel, and A. P. Zwane (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions. *The Quarterly Journal of Economics 126*(1), 145–205.

Molina, H. S., I. A. Pérez, A. A. Alonso, J. P. D. Martínez, M. P. Castellanos, C. F. del Valle Laisequilla, J. G. R. García, B. B. Estrada, and A. A. Alonso (2015). Carga económica de la obesidad y sus comorbilidades en pacientes adultos en México. *PharmacoEconomics Spanish Research Articles 12*(4), 115–122.

Nakhimovsky, S. S., A. B. Feigl, C. Avila, G. O'Sullivan, E. Macgregor-Skinner, and M. Spranca (2016). Taxes on sugar-sweetened beverages to reduce overweight and obesity in middle-income countries: A systematic review. *PloS one 11*(9), e0163358.

Ng, M., T. Fleming, M. Robinson, B. Thomson, N. Graetz, C. Margono, E. C. Mullany, S. Biryukov, C. Abbafati, S. F. Abera, et al. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet 384*(9945), 766–781.

Onufrak, S. J., S. Park, J. R. Sharkey, and B. Sherry (2014). The relationship of perceptions of tap water safety with intake of sugar-sweetened beverages and plain water among US adults. *Public health nutrition 17*(01), 179–185.

Prüss-Ustün, A., J. Wolf, C. Corvalán, and M. Neira (2016). Preventing disease through healthy environments. a global assessment of the burden of disease from environmental risks. *Geneva: World Health Organization*.

Ritter, P. I. (2019). Soda consumption in the tropics: The trade-off between obesity and diarrhea in developing countries. Technical report, Working Paper.

Roache, S. A. and L. O. Gostin (2017). The untapped power of soda taxes: incentivizing consumers, generating revenue, and altering corporate behavior. *International journal of health policy and management 6*(9), 489.

Stern, D., C. Piernas, S. Barquera, J. A. Rivera, and B. M. Popkin (2014). Caloric beverages were major sources of energy among children and adults in Mexico, 1999–2012. *The Journal of nutrition 144*(6), 949–956.

UNICEF (2016). One is too many: Ending child deaths from pneumonia and diarrhoea. Technical report, United Nations' Children Fund.

WHO (2016). Fiscal policies for diet and the prevention of noncommunicable diseases. *Ginebra: World Health Organization*.

Table 1:
Pre-tax Purchases of Beverages

	High SES			Low SES		
	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.
Fraction with taxed drink purchases	0.67	0.47	2,613	0.56	0.50	3,468
Taxed drinks purchased (L)	4.21	5.59	2,613	2.33	3.44	3,468
Taxed drinks, if purchased (L)	6.27	5.80	1,820	4.19	3.66	1,950
Fraction with bottled water purchases	0.44	0.50	2,613	0.23	0.42	3,468
Bottled water purchased (L)	15.37	35.43	2,613	5.62	14.26	3,468
Bottled water, if purchased (L)	35.28	46.69	1,137	24.63	20.58	748

Notes: This table shows weekly purchases of beverages in liters (L) by households prior to the SB tax, using household survey data from the 2012 ENIGH. We show the fraction of households with positive purchases, the average amount purchased, and the average amount conditional on purchasing a positive quantity. Taxed drinks include all sodas and energy drinks. Bottled water includes all bottled water and club soda. We present statistics for the top and bottom terciles of total household income reported in the survey. Survey weights are included in the calculations.

Table 2:
Summary Statistics

	Control		Treatment		Difference	
	Mean	SD	Mean	SD	Diff.	SE
Panel A: Clinic outcomes and characteristics						
GID rate per 100,000	117.52	(186.67)	105.25	(180.88)	-12.28***	(0.93)
Accidents rate per 100,000	6.11	(2.31)	6.09	(2.33)	-0.02	(0.01)
STD rate per 100,000	16.38	(30.57)	16.42	(31.42)	0.04	(0.15)
Chronic diseases rate per 100,000	58.38	(56.04)	53.47	(53.54)	-4.91***	(0.28)
Accidents, STDs, and chronic rate per 100,000	80.87	(74.52)	75.97	(73.95)	-4.89***	(0.37)
IMSS clinic	0.07	(0.26)	0.01	(0.11)	-0.06***	(0.00)
IMSS-Oportunidades clinic	0.23	(0.42)	0.20	(0.40)	-0.04***	(0.00)
ISSSTE clinic	0.03	(0.18)	0.00	(0.05)	-0.03***	(0.00)
SSA clinic	0.66	(0.47)	0.79	(0.41)	0.13***	(0.00)
Distance to monitoring station, kms	18.89	(20.25)	18.91	(18.11)	0.02	(0.10)
Panel B: Household characteristics at the precinct level						
Fraction households without water inside	0.53	(0.36)	0.88	(0.10)	0.34***	(0.00)
Fraction households without electricity	0.06	(0.11)	0.08	(0.15)	0.02***	(0.00)
Fraction households without sewerage	0.28	(0.30)	0.43	(0.30)	0.16***	(0.00)
Fraction households without bathroom inside	0.13	(0.17)	0.27	(0.24)	0.15***	(0.00)
Panel C: Water quality from monitoring stations						
Biochemical oxygen demand, 2012	17.69	(44.43)	36.47	(58.72)	18.79***	(0.28)
Biochemical oxygen demand, 2013	21.39	(178.21)	39.62	(153.06)	18.23***	(0.98)
Chemical oxygen demand, 2012	55.86	(116.13)	107.77	(166.97)	51.91***	(0.74)
Chemical oxygen demand, 2013	78.02	(393.97)	155.81	(405.23)	77.79***	(2.19)
Total suspended solids, 2012	30.78	(75.44)	82.25	(140.85)	51.47***	(0.47)
Total suspended solids, 2013	77.62	(207.33)	278.40	(413.73)	200.78***	(1.21)
River	0.67	(0.47)	0.68	(0.47)	0.00	(0.00)
Dam, lake, or lagoon	0.26	(0.44)	0.31	(0.46)	0.04***	(0.00)
Stream, canal, spring, or reservoir	0.06	(0.24)	0.02	(0.14)	-0.04***	(0.00)
Observations	393,064		44,688		437,752	
Total clinics	14,038		1,596		15,634	

Notes: This table presents summary statistics at the clinic level for our sample by treatment and control groups. The mean and standard deviation for each group is shown, as well as the difference and the standard error for a test of difference in means. Panel A shows clinic-level outcomes and characteristics. All rates are winsorized at the 5% level. Panel B shows household characteristics at the precinct level from the 2010 census. Panel C shows descriptives for water quality obtained from government monitoring stations, including the location of the stations by source.
*** p<0.01, ** p<0.05, * p<0.1

Table 3:
DD Effect of the Soda Tax on GID Rates

	Main treatment definition				Alternative definitions of treatment			
					CONAGUA threshold	Include 2014	Mean access	Continuous index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment × post	3.597**							
	(1.723)							
Post-tax years								
Treatment × 2015		3.154	3.704*	4.147*	5.285*	3.303	1.833	0.847***
		(2.052)	(2.163)	(2.287)	(2.870)	(2.100)	(1.936)	(0.322)
Treatment × 2014		7.188***	8.231***	7.495***	4.432**	6.911***	5.507***	0.603***
		(1.662)	(1.754)	(1.842)	(2.134)	(1.713)	(1.571)	(0.218)
Pre-tax years								
Treatment × 2012		2.145	2.753	2.675	0.565	1.181	1.171	-0.116
		(1.637)	(1.737)	(1.794)	(2.127)	(1.740)	(1.554)	(0.230)
Treatment × 2011		2.825	2.067	3.295	0.850	1.619	1.793	-0.288
		(1.998)	(2.095)	(2.160)	(2.671)	(2.046)	(1.882)	(0.281)
Treatment × 2010		3.226	1.965	1.357	3.214	2.770	2.305	0.087
		(2.170)	(2.296)	(2.433)	(2.730)	(2.281)	(2.043)	(0.337)
Treatment × 2009		-0.324	-0.579	-1.956	-3.058	0.023	-0.861	-0.158
		(2.351)	(2.476)	(2.636)	(3.204)	(2.458)	(2.224)	(0.383)
Observations	437,752	437,752	437,752	437,192	437,752	437,752	437,752	437,752
Treated clinics	1,596	1,596	1,596	1,596	720	1,558	1,818	
R-squared	0.800	0.800	0.801	0.814	0.800	0.800	0.800	0.800
Household controls			X	X				
Geographic controls				X				
Mean dependent variable	109.6	109.6	109.6	109.6	94.45	114.4	110.3	121.5
Coefficient tests:								
$H_0 : T \times 2014 = T \times 2015$		0.024	0.015	0.085	0.742	0.050	0.028	0.430
$H_0 : T \times k = 0, \forall k = 2009, ..., 2012$		0.185	0.331	0.142	0.179	0.557	0.351	0.566

Notes: This table shows the main results. Column 1 corresponds to the estimates of equation 1, while the other columns follow the dynamic DD from estimating equation 2 on a balanced panel of outpatient clinic-quarters (15,634 clinics × 28 quarters). The outcome is the clinic GID rate per 100,000, winsorized at the 5%. Coefficients for the interaction of the treatment indicator and each year in the sample are shown, with 2013 as the excluded year (column 1 interacts with an indicator for post-tax years). Household controls are indicators for each year-quarter interacted with indicators for being in quartile q of the precinct-level distribution of access to electricity, sewerage, and a bathroom inside the home. Geographic controls are indicators for each year-quarter interacted with indicators for latitude-longitude grid cells. Columns 1-4 correspond to the main definition of the treatment clinics. Columns 5-8 consider alternative definitions: column 5 uses the CONAGUA thresholds to determine surface water quality, column 6 includes water quality measures from 2014, column 7 uses the mean across precincts in access to piped water as the threshold for the access indicator, and column 8 constructs a continuous index for treatment. Robust standard errors clustered at the clinic level. We test whether the post-tax coefficients are equal to each other, and whether the pre-tax coefficients are jointly equal to zero. The mean of the dependent variable for the treatment clinics prior to the tax is shown.
*** p<0.01, ** p<0.05, * p<0.1

Table 4:
DD Effect of the Soda Tax on GID Rates: Robustness Checks on Distance to the Monitoring Stations and Winsorization Levels

		Excluding large distances			Alt. winsorization	
	Baseline	99th p.	95th p.	90th p.	2.5%	1%
	(1)	(2)	(3)	(4)	(5)	(6)
Post-tax years						
Treatment × 2015	3.154	2.741	2.414	1.892	2.862	1.926
	(2.052)	(2.033)	(2.067)	(2.120)	(2.718)	(3.649)
Treatment × 2014	7.188***	7.156***	7.068***	6.337***	9.593***	11.888***
	(1.662)	(1.672)	(1.695)	(1.734)	(2.281)	(3.047)
Pre-tax years						
Treatment × 2012	2.145	1.665	1.793	1.768	4.266*	5.481*
	(1.637)	(1.639)	(1.655)	(1.702)	(2.334)	(3.328)
Treatment × 2011	2.825	2.601	2.733	2.688	4.713*	7.697*
	(1.998)	(2.009)	(2.037)	(2.070)	(2.828)	(4.095)
Treatment × 2010	3.226	2.780	2.599	3.015	5.328*	9.464**
	(2.170)	(2.173)	(2.210)	(2.250)	(2.951)	(4.180)
Treatment × 2009	-0.324	-0.800	-1.110	-1.570	0.065	3.116
	(2.351)	(2.364)	(2.396)	(2.468)	(3.192)	(4.592)
Observations	437,752	433,384	415,856	393,960	437,752	437,752
Treated clinics	1,596	1,583	1,543	1,459	1,596	1,596
R-squared	0.800	0.802	0.802	0.803	0.792	0.774
Mean dependent variable	109.6	109.3	109.8	109.1	128.5	146.6
Coefficient tests:						
$H_0 : T \times 2014 = T \times 2015$	0.024	0.012	0.009	0.015	0.006	0.002
$H_0 : T \times k = 0, \forall k = 2009, ..., 2012$	0.185	0.230	0.194	0.099	0.065	0.096

Notes: This table shows robustness checks on the main results from estimating equation 2 on a balanced panel of outpatient clinic-quarters (15,634 clinics × 28 quarters). The outcome is the clinic GID rate per 100,000, winsorized at the 5%. Columns 2-4 exclude clinics that were assigned a far away monitoring station. Column 2 excludes the top 1% of the distance distribution (distance greater than 97 km), column 3 the top 5% (57 km), and column 4 the top 10% (44 km). Columns 5-6 consider alternative winsorization levels. Column 5 winsorizes at the 2.5%, and column 6 at the 1%. Coefficients for the interaction of the treatment indicator and each year in the sample are shown, with 2013 as the excluded year. Robust standard errors clustered at the clinic level. We test whether the post-tax coefficients are equal to each other, and whether the pre-tax coefficients are jointly equal to zero. The mean of the dependent variable for the treatment clinics prior to the tax is shown.
*** p<0.01, ** p<0.05, * p<0.1

Table 5:
DD Effect of the Soda Tax on GID Rates: Placebo Check on Unrelated Diseases and Conditions

	Accidents (1)	STDs (2)	Chronic (3)	All (4)
Post-tax years				
Treatment × 2015	-0.120**	-1.015	-0.784	-1.919*
	(0.051)	(0.640)	(0.737)	(1.105)
Treatment × 2014	-0.028	-0.769	0.754	-0.043
	(0.040)	(0.520)	(0.602)	(0.870)
Pre-tax years				
Treatment × 2012	0.017	0.230	0.399	0.646
	(0.037)	(0.517)	(0.613)	(0.838)
Treatment × 2011	-0.025	0.398	1.035	1.407
	(0.047)	(0.605)	(0.742)	(1.026)
Treatment × 2010	0.081*	-0.718	1.823**	1.187
	(0.047)	(0.610)	(0.802)	(1.090)
Treatment × 2009	0.114**	-1.754***	1.440*	-0.200
	(0.048)	(0.640)	(0.850)	(1.147)
Observations	437,752	437,752	437,752	437,752
R-squared	0.454	0.574	0.762	0.763
Mean dependent variable	6.220	15.26	55.96	77.44
Coefficient tests:				
$H_0 : T \times 2014 = T \times 2015$	0.040	0.608	0.012	0.030
$H_0 : T \times k = 0, \forall k = 2009, ..., 2012$	0.011	0.002	0.243	0.364

Notes: This table shows placebo tests of the main results from estimating equation 2 on a balanced panel of outpatient clinic-quarters (15,634 clinics × 28 quarters) for unrelated conditions. The outcomes considered are accidents and external injuries, sexually transmitted diseases, chronic diseases, and the sum of all three. All outcomes are measured as rates per 100,000, winsorized at the 5%. Coefficients for the interaction of the treatment indicator and each year in the sample are shown, with 2013 as the excluded year. Robust standard errors clustered at the clinic level. We test whether the post-tax coefficients are equal to each other, and whether the pre-tax coefficients are jointly equal to zero. The mean of the dependent variable for the treatment clinics prior to the tax is shown.
*** p<0.01, ** p<0.05, * p<0.1

Table 6:
DD Effect of the Soda Tax on GID Hospitalization Rates at SSA Hospitals

	Hospitalization Rate		Length of Stay	
	All	< 6 y.o.	GIDs	All
	(1)	(2)	(3)	(4)
Post-tax years				
Treatment × 2015	-0.329	0.562	-0.096	0.074
	(11.999)	(5.734)	(0.273)	(0.208)
Treatment × 2014	-0.660	-1.850	-0.230	-0.055
	(5.043)	(1.950)	(0.262)	(0.068)
Pre-tax years				
Treatment × 2012	1.536	0.187	-0.326	-0.094
	(4.414)	(2.337)	(0.221)	(0.126)
Treatment × 2011	4.691	-0.473	-0.215	0.603
	(6.750)	(3.190)	(0.274)	(0.775)
Treatment × 2010	5.815	0.555	-0.518*	1.302
	(8.473)	(3.818)	(0.272)	(1.146)
Treatment × 2009	5.866	1.807	-0.295	0.622
	(8.223)	(4.620)	(0.315)	(0.810)
Observations	21,448	21,448	21,448	21,448
R-squared	0.685	0.657	0.438	0.885
Mean dependent variable	71.57	35.42	1.411	2.578
Coefficient tests:				
$H_0 : T \times 2014 = T \times 2015$	0.978	0.667	0.668	0.535
$H_0 : T \times k = 0, \forall k = 2009, ..., 2012$	0.955	0.975	0.001	0.688

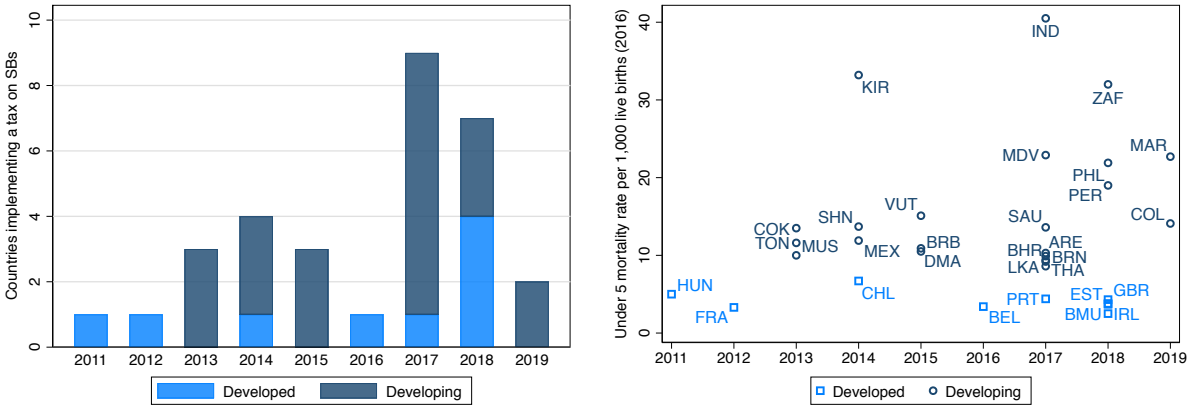
Notes: This table shows the results on SSA hospitalizations from estimating equation 2 on a balanced panel of SSA hospital-quarters (766 hospitals × 28 quarters). The outcome in columns 1 and 2 is hospitalizations due to GIDs per 100,000, for the full population in column 1 and for children under 6 in column 2, winsorized at the 5%. The outcome in columns 3 and 4 correspond to the length of hospital stays due to GIDs and all hospitalizations, respectively. All lengths of stay are winsorized at the 1%. Coefficients for the interaction of the treatment indicator and each year in the sample are shown, with 2013 as the excluded year. Robust standard errors clustered at the hospital level. We test whether the post-tax coefficients are equal to each other, and whether the pre-tax coefficients are jointly equal to zero. The mean of the dependent variable for the treatment hospitals prior to the tax is shown.
*** p<0.01, ** p<0.05, * p<0.1

Table 7:
Differential Changes in Consumption of Bottled Beverages

	SBs and water		All water		All SBs	
	(1)	(2)	(3)	(4)	(5)	(6)
Post-tax years						
Treatment × 2016	0.182**	0.192**	0.173*	0.200**	0.067	0.056
	(0.088)	(0.085)	(0.099)	(0.095)	(0.052)	(0.050)
Treatment × 2014	0.052	0.047	0.087	0.092	-0.007	-0.015
	(0.095)	(0.092)	(0.108)	(0.103)	(0.059)	(0.058)
Pre-tax years						
Treatment × 2010	0.056	0.060	0.080	0.099	0.004	-0.004
	(0.086)	(0.085)	(0.091)	(0.089)	(0.059)	(0.058)
Treatment × 2008	0.016	0.024	0.021	0.042	-0.010	-0.016
	(0.088)	(0.086)	(0.090)	(0.088)	(0.059)	(0.056)
Observations	153,064	153,064	153,064	153,064	153,064	153,064
R-squared	0.128	0.169	0.135	0.150	0.125	0.173
Controls		X		X		X
Mean dependent variable	9.66	9.66	6.89	6.89	2.77	2.77
Coefficient tests:						
$H_0 : T \times 2014 = T \times 2016$	0.041	0.020	0.262	0.138	0.039	0.046
$H_0 : T \times k = 0, \forall k = 2008, 2010, 2012$	0.938	0.892	0.710	0.536	0.953	0.907

Notes: This table shows DD estimates of changes in consumption over time for households in treatment versus control areas using data from the 2008, 2010, 2012, 2014 and 2016 rounds of the ENIGH. The outcome variable is the inverse hyperbolic sine of liters purchased. The unit of observation is a household-year. Columns 1 and 2 correspond to all bottled beverages, columns 3 and 4 are all bottled water, and columns 5 and 6 are all SBs. Even-numbered columns include household-level controls (log income, whether the house has access to electricity, sewerage and a bathroom inside, indicators for household head's years of education, indicators for household size, fraction of household members that are adults, and indicators for size of the locality). Robust standard errors clustered at the municipality level. We test whether there is a difference in the differential response of treatment households in 2014 relative to 2016, and whether the common trends assumption holds prior to the tax.
*** p<0.01, ** p<0.05, * p<0.1

Figure 1:
Implementation of SB Taxes Around the World Over Time

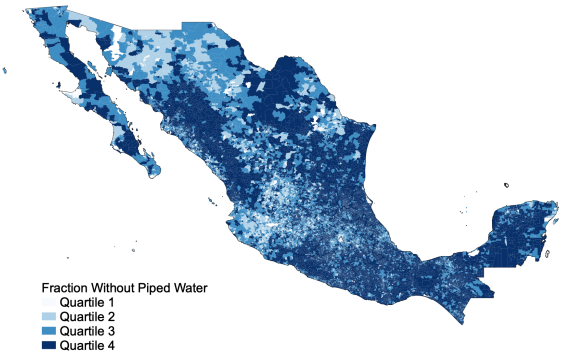


(a) Countries with SB taxes by year of implementation

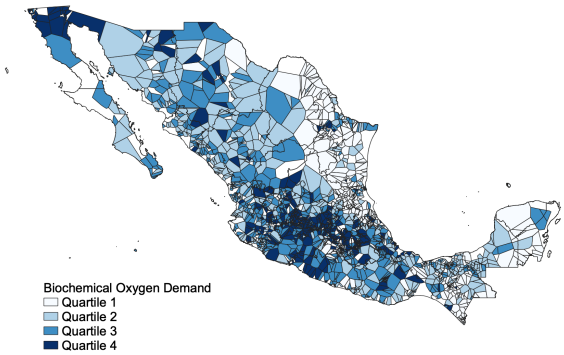
(b) Under 5 mortality rate of countries with SB taxes by year of implementation

Notes: The plot on the left shows the number of countries implementing a tax on sugary beverages (SBs) by year since 2011, using data reported in (Allcott et al., 2019). There are a total of 31 countries implementing a tax in this time frame (versus eight from 1940 to 2007). We classify these countries by development status using the UN World Economic Situation and Prospects classification. The plot on the right shows the under 5 mortality rate per 1,000 live births in countries implementing a tax on SBs by year of implementation, using measures from 2016 recovered from the CIA World Factbook.

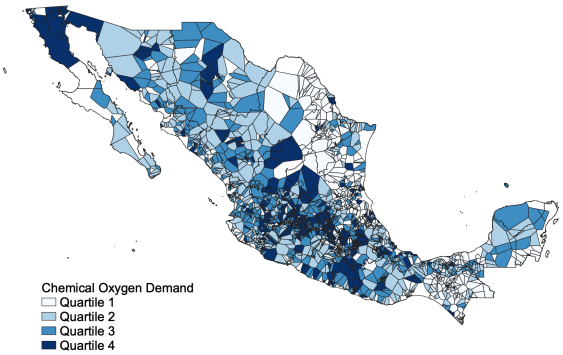
Figure 2:
Geographic Distribution of Access to Piped Water and Water Quality Measures



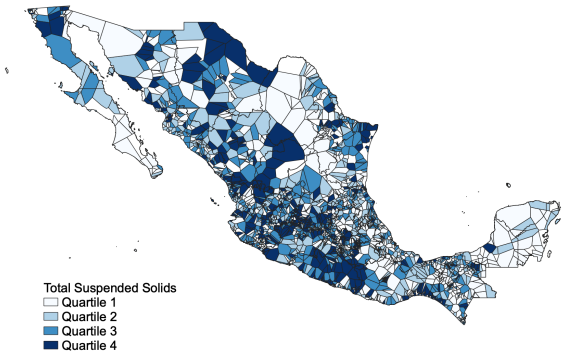
(a) Fraction without access to piped water



(b) Biochemical oxygen demand



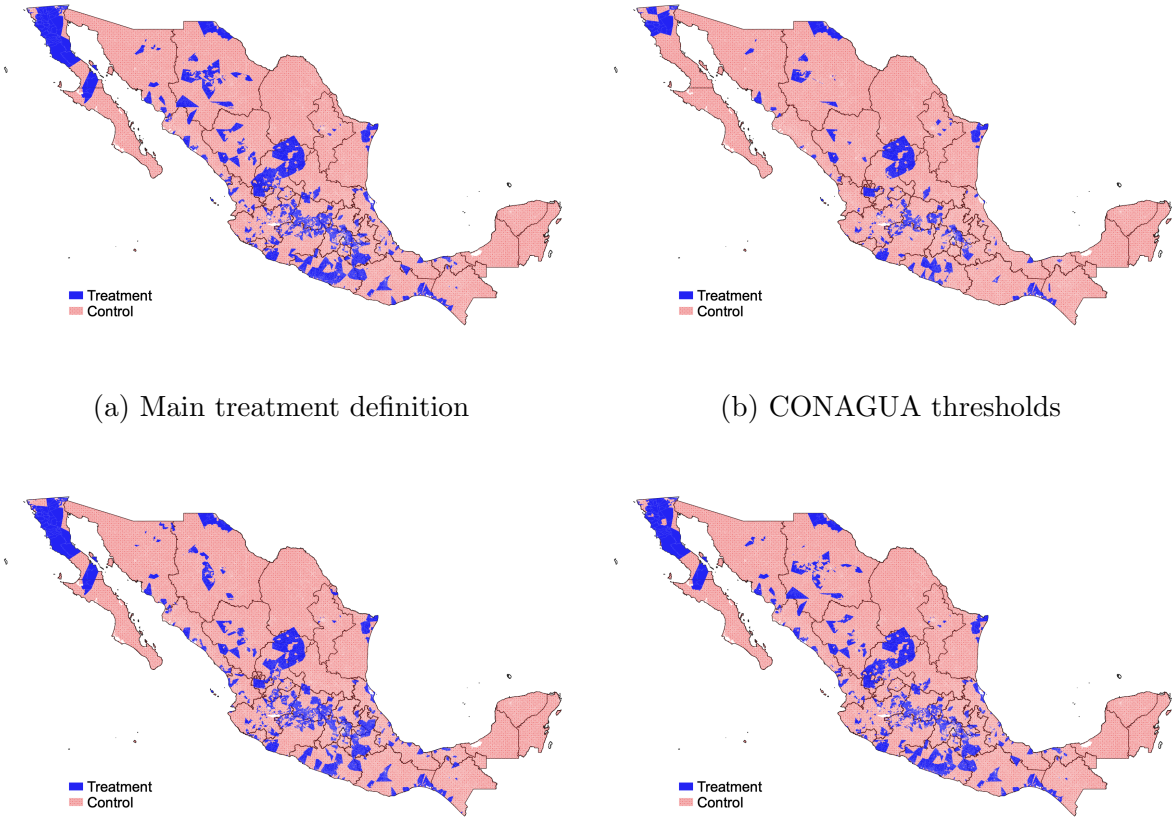
(c) Chemical oxygen demand



(d) Total suspended solids

Notes: These maps show the geographic distribution of our piped water access and surface water quality variables. The top left graph corresponds to the fraction of households in an electoral precinct without access to piped water inside the home, using data from the 2010 census. The other three graphs show each of the three water quality measures, averaged over 2012 and 2013, at the Thiessen polygon level. All graphs split the data into quartiles.

Figure 3:
Geographic Distribution of Treatment and Control Areas



(a) Main treatment definition

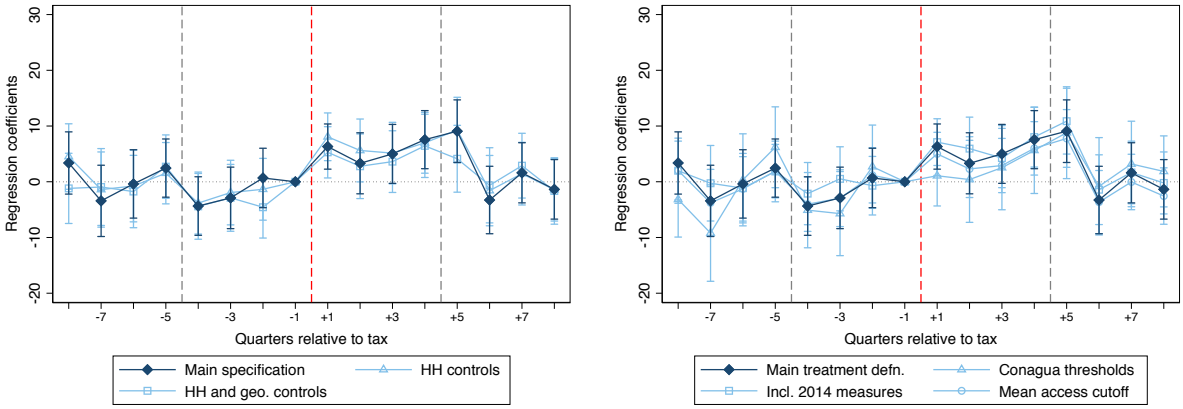
(b) CONAGUA thresholds

(c) Including 2014 measures

(d) Mean access as cutoff

Notes: These maps show the geographic distribution of our treatment and control areas. The top left map considers our main definition of treatment, while the other maps show alternative definitions. The top right map uses the CONAGUA thresholds instead of the median for quality. The bottom left map includes 2014 measures of quality. The bottom right map uses the mean of the precinct-level distribution of access to piped water as the cutoff instead of the median.

Figure 4:
DD Effect of the Soda Tax on GID Rates



(a) Main treatment definition

(b) Alternative treatment definitions

Notes: These graphs show the main results from estimating equation 3 on a balanced panel of outpatient clinic-quarters (15,634 clinics × 28 quarters). The outcome is the clinic GID rate per 100,000, winsorized at the 5%. Coefficients for the interaction of the treatment indicator and each quarter for two years before and after the tax was introduced are shown, with quarter 4 of 2013 as the excluded period. The plot on the left corresponds to the main treatment definition, with the lighter-colored coefficient series corresponding to specifications with additional controls (see main text and Table 3). The plot on the right contrasts the main treatment definition with alternative definitions (see main text and Table 3). Robust standard errors clustered at the clinic level. Error bars show 95% confidence intervals. The mean of the dependent variable for the treatment clinics prior to the tax under the main treatment definition is 110.

Appendices for Online Publication

Contents

Appendix A shows evidence of the first stage effect of the soda tax. We first summarize the existing literature that has evaluated the effect of the tax on prices and consumption. We then show descriptive evidence from multiple household survey rounds.

Appendix B looks more carefully at our data sources. First, we show descriptives on access to piped water at the electoral precinct level. Second, we present more information on surface water quality from the monitoring stations. Lastly, we show clinic-level descriptives relating to the assignment of water variables. This helps paint a complete picture of our data sources, the variation that we exploit in the empirical analysis, and how our final dataset at the clinic level relates to the multiple water measures.

Appendix C uses clinic-level data on staffing and infrastructure to explore whether there was any differential change in treatment clinics from 2013 to 2014 relative to the controls. We estimate regressions similar to the main specification in the text for a variety of outcomes measuring clinic resources, obtaining small and statistically insignificant estimates. This

reassures us that our main results are not driven by supply-side changes in treated clinics that coincided with the introduction of the tax.

Appendix D explores whether the local prevalence of GIDs has an effect on the likelihood of seeking outpatient care using survey data from before the tax. We estimate regressions of an indicator for seeking care at a public (or private) clinic on an indicator for being sick with a GID, the local GID rate, and the interaction between the two. Our main test is whether the chances of seeking care when sick with a GID rate change when the local prevalence of GIDs is changing. We find that this is not the case. This exercise allows us to claim that our measure of GIDs, which relies on clinic reports and not the total epidemiological prevalence, is not introducing a bias in our estimates.

Appendix E presents estimates of both the robustness checks and additional results from estimating equation 3. We show these estimates in graphical form as a complement to the point estimates reported in table format in the main text.

A First Stage Evidence of the Soda Tax

Due to data limitations, we are unable to show a direct first stage indicating the impact of the tax on SB prices and/or consumption. However, we can present two pieces of information that speak to this issue. First, we rely on previously estimated impacts. Second, we present some descriptives from the National Household Income and Expenditures Survey (ENIGH).

For our literature review, we were able to identify five papers that analyzed the effects of the tax on prices and consumption. A summary of the main findings are shown in Table A1. Overall, these studies consistently find a 100% pass-through of the tax on SB prices throughout the first year of the tax. There is little evidence on the long-run effects on prices. Grogger (2017) finds similar estimates in the first six month of 2015 as the full pass-through of 2014.

In terms of consumption, findings are consistent across studies with an estimated 6% decline in SB purchases during 2014. However, there are mixed results in terms of the dynamics of this effect throughout the first year. Using an extrapolation method, Colchero et al. (2016) finds that there is an initial drop in consumption that gets larger over time. On the other hand, Aguilar et al. (2019) implement a synthetic control to find the opposite: after an initial sharp drop, the decrease in SB purchases gets smaller in absolute value over time. A large drawback is that neither study analyzes data after 2014, limiting our sense of how SB purchases were affected by the tax in the long run. Identifying the long-run effects of a nation-wide policy is not a simple task. This is likely the reason for why studies have not attempted to calculate them.

Overall, the existing literature allows us to confidently infer that the tax had a strong effect of around 10% on prices, as well as a 6% decrease in purchases of SBs during 2014, the first year of the tax. However, we are less sure about effects in 2015, where the literature is silent with respect to consumption effects, although at least it seems that pass-through was maintained.

To address these shortcomings, we turn to data from the 2008, 20010, 2012, 2014 and 2016 ENIGH survey rounds. This is a repeated cross-section that allows us to observe weekly household purchases of different beverages. We present results for taxed beverages and bottled water. Taxed beverages includes sodas and energy drinks. Note that the survey does not distinguish between regular and diet sodas. However, the latter seem to be a small fraction of total soda consumption. Note also that the survey does not allow us to distinguish between taxed and untaxed juice. Bottled water includes all bottled water (in all presentation sizes) and club soda. For this exercise, we are not attempting to assign a causal interpretation, as we have no control group. However, we believe these patterns are suggestive of the potential impact of the tax on consumption.

Figure A1 shows the point estimates from a regression of purchases on survey round indicators, log income, indicators for whether the house has access to electricity, sewerage and a bathroom inside, indicators for household head's years of education, indicators for household size, fraction of household members that are adults, indicators for size of the locality, and municipality fixed effects. These controls are consistent with what has been used in the literature. Survey weights are included in the estimation. The plot on the left considers taxed beverages, while the one on the right is for bottled water.

Overall, these descriptive trends show that consumption of taxed beverages decreased in the post-tax years, while the trends for bottled water are not as clear. Although the point estimates for 2008 and 2010 are smaller than the others, the post-tax estimates are not significantly different from the 2012 mean.

In order to provide one last piece of evidence that may indicate substitution patterns between soda and water, we present summary statistics from the 2016 National Health Survey (ENSANUT). This survey asks about perceived changes in consumption after the implementation of the tax. Although we recognize that this is an imperfect measure, we believe that it sheds light on the possibility that water and SBs are substitutes to some degree in our context.

Table A2 shows the distribution of individuals reporting that their water and SB consumption went down, stayed the same or went up, for the top and bottom SES terciles. This table indicates that the majority of the increase in water consumption in the two years after the tax was implemented corresponds to individuals who decreased their SB consumption. Furthermore, the numbers suggest that substitution occurred across all SES groups, although perhaps to a lesser degree among low SES individuals.

Table A1:
Previously Estimated Impacts of the SB Tax in Mexico on SB Prices and Consumption

	Short-Run Effects	Long-Run Effects
Aguilar et al. (2019)	100% pass-though of the tax on prices, and 6% decrease in purchases throughout 2014; an initial sharp drop in consumption is followed by an increasing trend	N.A.
Colchero et al. (2017)	6.3% reduction in purchases and 2% reduction in probability of purchasing SBs during 2014	N.A.
Colchero et al. (2016)	average 6% reduction in purchases throughout 2014, decreasing at an increasing rate with a 12% decline in December 2014	N.A.
Colchero et al. (2015)	0.95-1.12 pesos per liter increase in SB prices, 11% increase in carbonated SB prices throughout 2014; evidence of full pass-through of the tax	N.A.
Grogger (2017)	12.3-14.1% price increase during 2014; over 100% pass-through	estimates hold up to June 2015, no increasing or decreasing trend

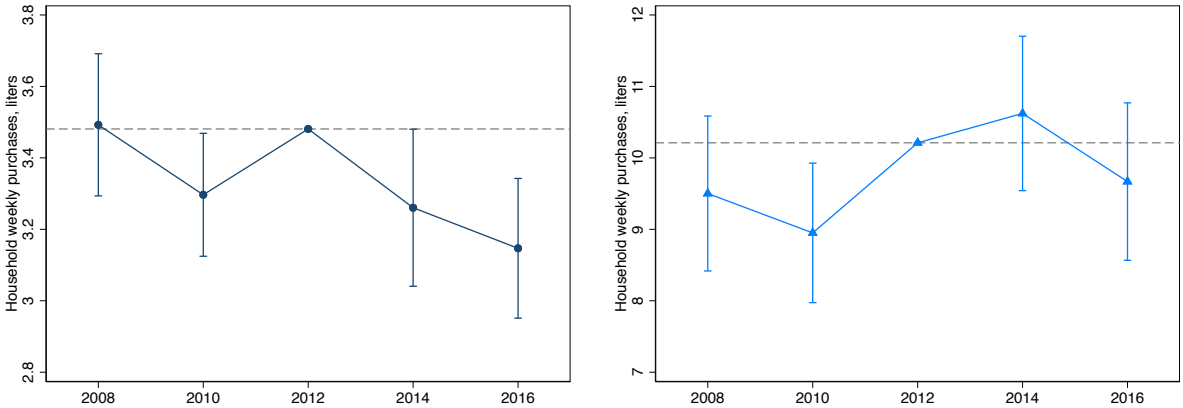
Notes: This table summarizes the main findings of the papers that have studied the effect of the SB tax in Mexico on both prices and consumption of SBs. We distinguish between short and long-run effects (roughly one year after the implementation of the tax). Cells with "N.A." indicate that the paper did not estimate those effects.

Table A2:
Perceptions of Consumption Changes After the Tax

		Water consumption			
	Panel A: High SES				
		Went down	The same	Went up	Total
Sugary beverage consumption	Went down	2%	9	34	45
	The same	3	19	18	40
	Went up	5	5	5	15
	Total	10	33	57	
	Panel B: Low SES				
		Went down	The same	Went up	Total
	Went down	3%	12	23	38
	The same	3	31	13	47
	Went up	4	6	5	15
	Total	10	49	41	

Notes: This table shows the distribution of individuals answering questions about how they have perceived changes in their consumption of both water and sugary beverages during the two years after the tax, according to the 2016 ENSANUT. Panel A shows high socioeconomic status individuals, corresponding to those in the top tercile (definition provided in the survey). Panel B shows the bottom tercile. Survey weights are included in the calculations.

Figure A1:
Evolution of Household Purchases of Taxed Beverages and Bottled Water over Time



(a) Taxed beverages

(b) Bottled water

Notes: These graphs show trends in purchases of taxed beverages and bottled water using the 2008, 2010, 2012, 2014 and 2016 ENIGH rounds. Taxed beverages are defined as all sodas (including diet sodas which were untaxed, but the survey does not allow for this distinction), and energy and sports drinks. Bottled water includes regular and sparkling water. The graphs show the point estimates of a regression of household-level purchases on survey round indicators, log income, indicators for whether the house has access to electricity, sewerage and a bathroom inside, indicators for household head's years of education, indicators for household size, fraction of household members that are adults, indicators for size of the locality, and municipality fixed effects. Survey weights are included in the estimation. Error bars show 95% confidence intervals using robust standard errors clustered at the municipality level.

B Access to Piped Water, Surface Water Quality, and Public Outpatient Clinics

B.1 Access to Piped Water

Table B1 shows descriptive statistics at the electoral precinct level for the variables measuring access to piped water in the 2010 census for the total number of precincts in Mexico. This table shows that on average 63% of households within a precinct have access to piped water inside their home. The remaining 37% is then broken down by other sources: around 24% obtain piped water from neighbors or a communal tap, 1% buy water from vendors, and the remaining 12% use ground and surface water from wells (8%) and from rivers, lakes and dams (3%).

Figure B1 shows histograms of lack of access to piped water inside the home, using precinct-level data from the 2010 census. Not all precincts have a public outpatient clinic. For our main sample, we have 13,732 precincts. These plots show the distribution for our sample. Figure B1a corresponds to the density of the fraction of households in an electoral precinct without piped water at home. This distribution has two humps at each extreme. This is consistent with a large number of precincts having full or almost full access to piped water, with another relatively big mass of precincts that have almost no access to piped water at home. The latter are mostly rural. Figure B1b shows the same distribution over the number of households instead of the fraction.

B.2 Surface Water Quality

As described in the main text, our data for surface water quality comes from monitoring stations belonging to the Mexican government. While there are over 3,000 stations, we discard those that are located at salt water sources. In total, we are left with 2,071 monitoring stations in our main sample, from which we construct our Thiessen polygons. Figure B2

shows a map with the location of these monitoring stations, as well as the corresponding Thiessen polygons.

These stations are located across a variety of water sources. Figure B3 shows the distribution of these water sources for our monitoring stations. The graph shows that the vast majority (62%) of monitoring stations are located at rivers. This is followed by dams (15%), lagoons (9%), and lakes (9%).

Table B2 shows descriptive statistics of the surface water quality measures at these stations. As described in the main text, there are three measures per monitoring station: biochemical oxygen demand (BOD), chemical oxygen demand (COD), and total suspended solids (TSS). The mean, standard deviation, and median for each year (2012, 2013, and 2014) are shown for each measure. Note that some stations have missing values for some measures in some years. We also report summary statistics for the 2012-2013 average and the three-year average of each measure.

Figure B4 shows how each measure has changed over time. Each plot corresponds to one of the three measures, and shows point estimates from a regression of the continuous measure on year indicators and monitoring station fixed effects. The plot for BOD shows that there was not a lot of change in this measure over the course of these three years. The plot for COD shows an increase from 2012 to 2013. Lastly, the plot for TSS shows the most variation, with an increase from 2012 to 2013, followed by a decline in 2014.

Table B3 shows the criteria used by CONAGUA in classifying each of the three surface water quality measures into five categories of cleanliness. In one of our robustness checks, we use this stratification to classify stations into those with good surface water quality ("excellent" and "good" categories) and bad quality (below "good"). Since it is unclear to us how CONAGUA chose these thresholds (or even why five categories and not just two), we conservatively use the "good" category as our cutoff to construct our binary measure of quality. Standards in developed settings suggest that the "acceptable" category put forward by CONAGUA tends to fall above established thresholds. For example, the Canadian Ministry

of the Environment establishes a threshold of 5.5-6.5 mg/L for BOD in warm-water ecosystems. Michigan's Department of Environmental Quality indicates water appears cloudy for TSS levels between 40-80 mg/L. Utah, according to the EPA, establishes a threshold of 90 mg/L for TSS.

B.3 Assignment of Water Variables to Public Outpatient Clinics

Figure B5 shows a map with the location of the public outpatient clinics in our sample, as well as the electoral precincts in which they fall.

Figure B6 shows the distribution of distance from each public outpatient clinic to the monitoring station assigned via the Thiessen polygons as described in the main text. The plot depicts the 90th, 95th, and 99th percentiles which are used in a robustness check.

Figure B7 shows the distribution of the continuous treatment index that we construct as part of our sensitivity analysis. The plot shows the distribution separately for clinics that were classified as treatment and control under the baseline definition of treatment. This plot shows that there is considerable overlap, although there is a larger mass towards the left for the original control group and a larger mass towards the right for the original treatment group.

Table B1:
Access to Piped Water at the Section Level

	Mean	Std. Dev.
Percentage of HH getting water from:		
Sources outside the home	37.1	36.8
Piped water from neighbors/communal tap	24.1	28.1
Water from vendors	1.2	6.3
Ground and Surface Water	11.8	24.6
Wells	8.6	19.9
Rivers, lakes and dams	3.3	12.0
Total observations	64,559	

Notes: This table shows precinct-level averages for access to piped water according to the 2010 census.

Table B2:
Summary Statistics of Surface Water Quality

	Mean	Std. Dev.	Median	Obs.
Biochemical oxygen demand (mg/L)				
2012	15.70	38.90	4.00	1,694
2013	17.29	115.44	4.90	1,876
2014	16.65	107.82	5.00	1,906
2012-2013 average	17.30	108.12	5.20	1,955
Three-year average	17.51	109.25	5.23	1,914
Chemical oxygen demand (mg/L)				
2012	54.99	107.71	22.00	1,696
2013	72.34	278.24	35.60	1,879
2014	70.07	238.75	37.20	1,905
2012-2013 average	68.76	247.71	33.88	1,956
Three-year average	69.06	249.59	34.07	1,915
Total suspended solids (mg/L)				
2012	42.82	104.76	18.00	1,843
2013	86.03	172.09	34.00	2,014
2014	72.16	231.73	27.00	2,071
2012-2013 average	69.20	144.38	31.33	2,119
Three-year average	69.49	144.95	32.00	2,071

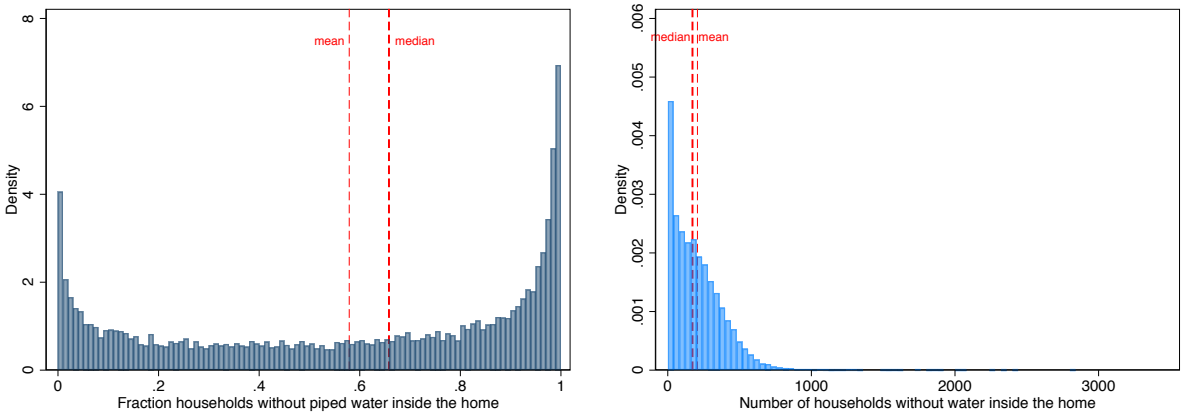
Notes: This table shows descriptives for the surface water quality measures across 2,071 monitoring stations used in the main sample.

Table B3:
Surface Water Quality Thresholds

	Biochemical oxygen demand	Chemical oxygen demand	Total suspended solids
Excellent	≤3 mg/L	≤10 mg/L	≤25 mg/L
Good	3-6	10-20	25-75
Acceptable	6-30	20-40	75-150
Polluted	30-120	40-200	150-400
Very polluted	>120	>200	>400

Notes: This table shows the thresholds established by CONAGUA in classifying each of the three surface water quality measures.

Figure B1:
Distribution of Piped Water Access

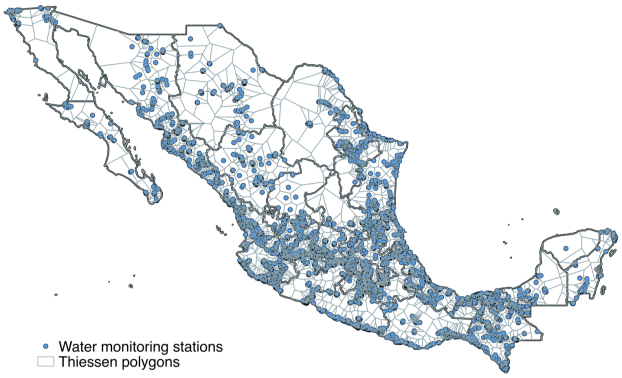


(a) Fraction of HHs without access

(b) Number of HHs without access

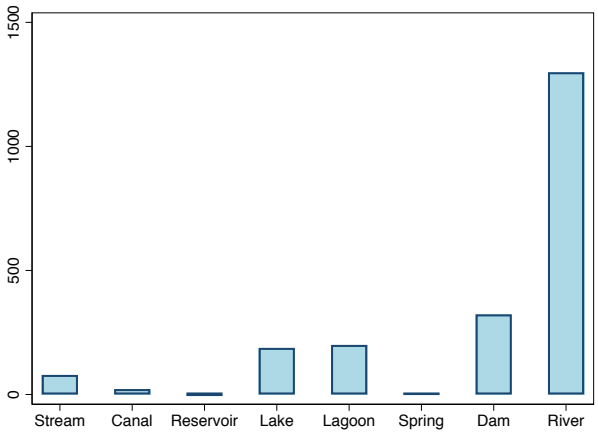
Notes: These graphs show histograms for piped water access based on data from the 2010 census. The graph on the left shows the distribution of the fraction of households in an electoral precinct that do not have access to piped water inside the home for all precincts. The graph on the right shows the number of households in a precinct without access to piped water inside the home (i.e., the fraction of households without access multiplied by the number of households in each electoral precinct). The thick dashed line represents the median of the distributions, and the thin dashed line corresponds to the mean.

Figure B2:
Water Monitoring Stations and Thiessen Polygons



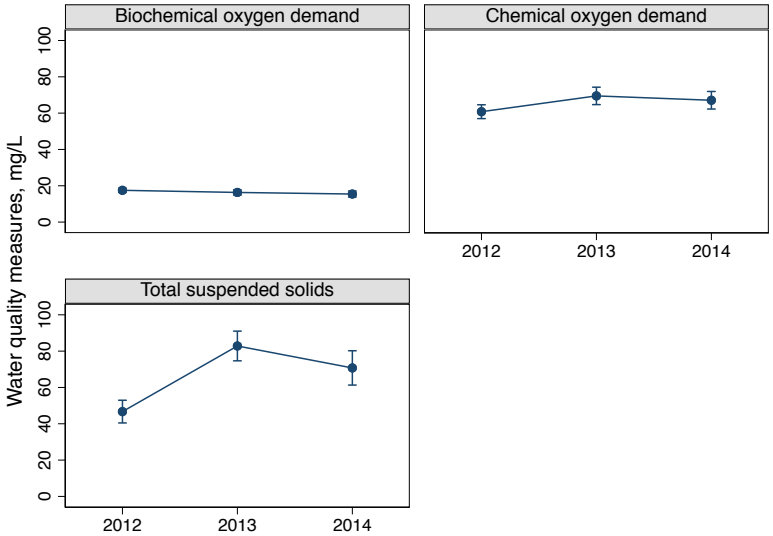
Notes: This map shows the 2,071 monitoring stations in the our sample, as well as their corresponding Thiessen polygons.

Figure B3:
Water Sources for Monitoring Stations in Sample



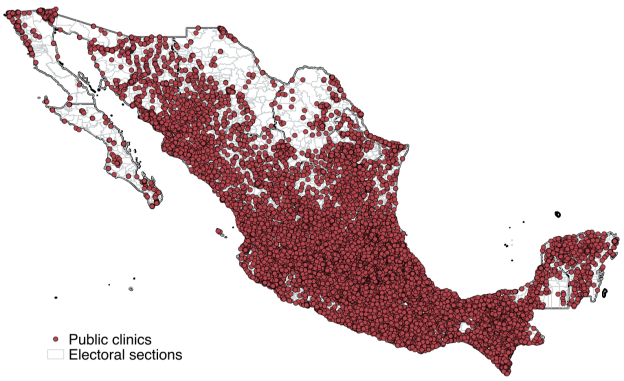
Notes: This graph shows the number of monitoring stations that correspond to each water source where quality is being measured.

Figure B4:
Evolution of Water Quality Measures over Time



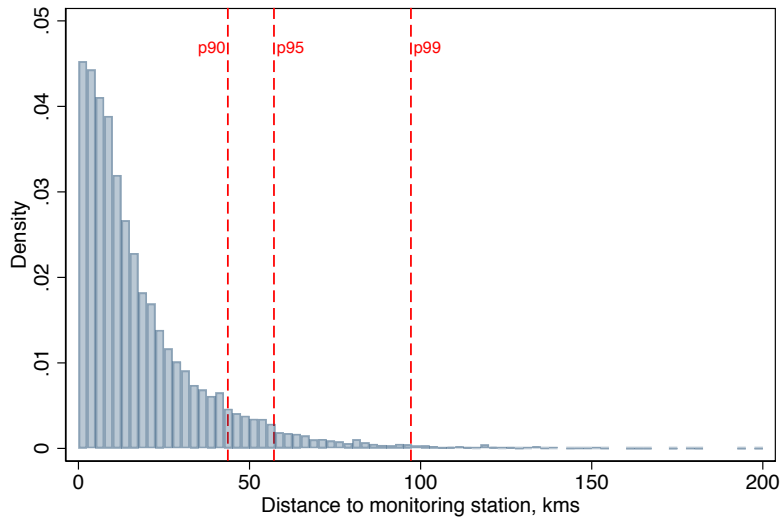
Notes: These graphs show how the water quality measures change over time for the three years from 2012 to 2014. Each plot shows the point estimates from a regression of the quality measure on indicators for each year and monitoring station fixed effects. Error bars show 95% confidence intervals using robust standard errors.

Figure B5:
Electoral Precincts and Public Outpatient Clinics



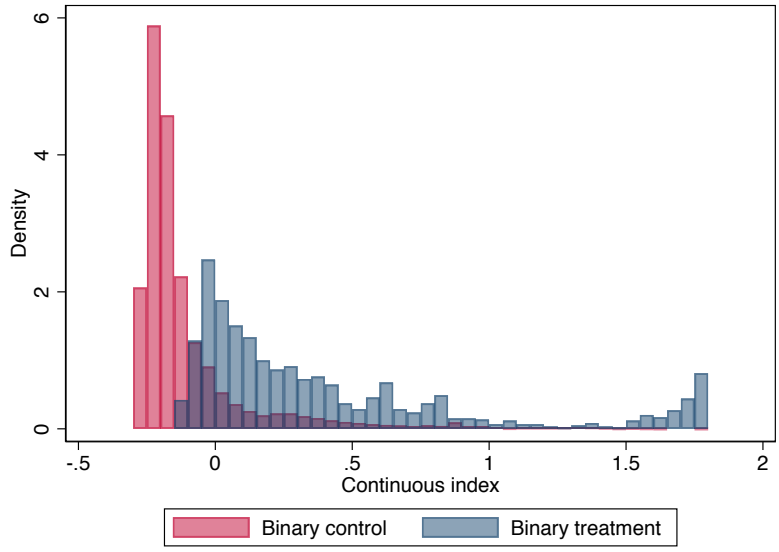
Notes: This map shows the 15,634 public outpatient clinics in our sample, as well as their corresponding electoral precinct.

Figure B6:
Distance from Clinics to Monitoring Stations



Notes: This graph shows the distribution of distance from the public outpatient clinics to their assigned monitoring station. The dashed lines show the 90th, 95th, and 99th percentile of the distribution.

Figure B7:
Distribution of the Continuous Treatment Index



Notes: This graph shows the distribution of the continuous index for treatment separately for control and treatment areas, based on the main binary definition of treatment. The index has been standardized to have a mean of zero and a standard deviation of one. The values of the index in the graph are capped at the 99th percentile for clarity.

C Clinic Staffing and Infrastructure

An important potential confounder for our results is that the soda tax coincided with other policies that improved public clinic staffing and infrastructure, with a larger improvement in our treatment areas. In order to show that this is not the case, we obtain clinic-level data on staffing and infrastructure at a yearly level for both 2013 and 2014. Unfortunately, data for previous years are not available. We estimate a regression akin to our main specification.

Since we cannot test for pre-trends given that we only observe one year pre-tax, we include flexible controls that are meant to capture the possibility of differential trends. For this, we construct quartiles of the variables measuring the fraction of households without access to electricity, without a bathroom inside the house, and without access to the sewerage system. We then interact an indicator for each of these quartiles with an indicator for 2014. This is analogous to the controls in our main specification.

Formally, our estimating equation for this exercise, analogous to the main text equation 2, is given by:

$$y_{ct} = \beta(Treat_c \times \mathbb{1}_{[t=2014]}) + \lambda_c + \theta_t + \sum_{q=1}^{4} \gamma_q (X_c^q \times \mathbb{1}_{[t=2014]}) + \nu_{ct} \tag{C1}$$

where y_{ct} is a resource outcome at clinic c in year t, λ_c are clinic fixed effects, θ_t is an indicator for each year, X_c^q is a vector of indicators for whether household characteristics (i.e., access to electricity, indoor bathroom, and sewerage) surrounding a clinic c fall in quantile q, and ν_{ct} is the error term. We transform the outcome variables using the inverse hyperbolic sine function. Standard errors are clustered at the clinic level. The coefficient of interest is given by β. This estimate measures whether resources at a treated clinic changed differentially over time relative to control clinics.

We include twelve variables measuring clinic resources as our outcomes of interest. These measures can be divided in two parts. First, we consider variables that refer to infrastructure and material resources. Here we include the following available measures: number of

examination rooms at the clinic, number of x-ray machines, and number of mammogram machines. Note that we exclude very specialized machinery that are reported in the data such as medical linear accelerators and hyperbaric oxygen beds. We also exclude less specialized material resources that are not widely used in our treatment clinics such as MRI machines and CAT scanners.

The second set of resources refer to medical staffing at the clinic. We include all the variables reported in the data as follows: total number of doctors, total general practitioners, total family medicine doctors, total number of interns, total number of pediatricians, total OB/GYN doctors, total number of nurses, total general nurses, and total specialized nurses.

Table C1 shows the estimates for this exercise. All coefficients are small relative to the pre-tax mean in treated clinics, and statistically indistinguishable from zero. Panel A indicates that there is no differential change in material resources in treatment clinics over time relative to the control. The first variable, measuring the number of examination rooms, captures infrastructure that may increase or decrease clinic capacity. The next two variables refer to physical resources in the form of diagnostic machinery, which may correlate with other additional funding impacting the clinic's capacity to diagnose and treat GID cases. The null results reassure us that changes in material resources are not the drivers of our main results.

Panel B shows that medical staffing does not change differentially in treated clinics over time relative to the control. We find no significant effects on the total number of doctors, nor any compositional effects in terms of the doctors' field of medicine. Likewise, there is no evidence of differential changes in staffing of nurses, either for the total nor by type of nurse. As above, these estimates indicate that our main results are not driven by differential changes in the human resources at the treated clinics.

Overall, this exercise provides convincing evidence that we are not confounding the effect of the soda tax in areas where people lack access to safe drinking water with any differential supply-side changes.

Table C1:
DD Estimates on Public Outpatient Clinic Resources

	Treatment × 2014	Observations	R-squared	Mean dep. var.
Panel A: Material Resources				
Number of examination rooms	-0.008	31,268	0.976	1.391
	(0.006)			
Number of x-ray machines	-0.000	31,268	0.974	0.002
	(0.000)			
Number of mammogram machines	-0.000	31,268	0.975	0.001
	(0.000)			
Panel B: Human Resources				
Total number of doctors	0.016	31,268	0.958	1.697
	(0.011)			
Number of general practitioner doctors	0.003	31,268	0.938	0.883
	(0.012)			
Number of family medicine doctors	-0.001	31,268	0.996	0.050
	(0.002)			
Number of medical interns	0.000	31,268	0.916	0.003
	(0.002)			
Number of pediatricians	-0.000	31,268	0.902	0.004
	(0.000)			
Number of obstetricians (OB/GYN)	-0.000	31,268	0.923	0.001
	(0.000)			
Total number of nurses	0.007	31,268	0.950	1.810
	(0.012)			
Number of general nurses	-0.011	31,268	0.927	0.536
	(0.012)			
Number of specialized nurses	-0.002	31,268	0.967	0.018
	(0.002)			

Notes: This table shows DD estimates on public clinic resources for the balanced panel of clinic infrastructure data (15,634 clinics × 2 years). Each row corresponds to a different regression based on equation C1 for the outcome variable listed. All outcomes transformed using the inverse hyperbolic sine function. The first column reports the coefficient of interest, measuring the differential change from 2013 to 2014 in treatment clinics relative to the control. The last column shows the mean of the dependent variable in treated clinics in 2013, prior to the soda tax. Regressions include clinic fixed effects, an indicator for the post-tax year, and flexible household SES controls as described in equation C1. Standard errors are clustered at the clinic level.
*** p<0.01, ** p<0.05, * p<0.1

D Prevalence of GIDs and Likelihood of Seeking Outpatient Care

Our estimates rely on using outpatient GID rates as our dependent variable. Evidently, this may not be an accurate representation of the overall prevalence of diarrheal disease in a clinic's catchment area. This means that we are identifying a lower bound on all possible new GID cases, since not everyone who is sick seeks medical attention.

In terms of the validity of our strategy, this is only a concern if individuals around treatment and control clinics differentially changed their likelihood of seeking medical care when sick. The question then is whether the mapping of unobserved GID prevalence to observed GID visits at public clinics is effectively changing over time, specifically as GIDs become more prevalent.

To shed light on this potential issue, we turn to survey data from the National Health Survey (ENSANUT). This is a nationally representative survey, usually carried out every six years. We explore data from the 2006 and 2012 rounds. Unfortunately, the 2016 round only focused on nutrition and chronic diseases, excluding questions on disease and healthcare utilization. Nevertheless, we believe that this exercise is informative.

We estimate the following equation using the individual-level data for both rounds:

$$y_{imr} = \beta_1 sick_{imr} + \beta_2 rate_{mr} + \beta_3 (sick \times rate)_{imr} + X'_{imr}\gamma + \lambda_m + \theta_r + \eta_{imr} \tag{D1}$$

where y_{imr} is an indicator for whether individual i in municipality m in survey round r sought medical care at a public clinic, $sick_{imr}$ is an indicator for being sick with a GID, $rate_{mr}$ is the GID rate excluding individual i, $\frac{1}{N_{mr}} \sum_{j \neq i} sick_{jmr}$, X'_{imr} is a vector of controls, λ_m are municipality fixed effects, θ_r are indicators for each round, and η_{imr} is the error term. Note that this is a repeated cross-section, where we cannot track the same individuals over time.

We recognize that these estimates only allow us to identify correlations within the data. However, these simple relationships may be very informative. The coefficient β_1 indicates by how much the observed probability of going to the public clinic increases when an individual is sick with a GID. The coefficient β_2 measures changes in the likelihood of seeking care as the prevalence of GIDs increases. Lastly, the coefficient β_3 indicates whether this probability changes differentially for individuals that are sick with a GID in areas with varying prevalence of GIDs.

We are especially interested in β_3. If we find a positive and significant coefficient, this would mean that the probability of seeking care when sick with a GID increases with the overall prevalence of GIDs in an individual's municipality. This would then suggest that clinic reports of GIDs increase mechanically whenever the prevalence of GIDs increases. If instead we find a statistical zero, then an individual's decision of seeking care when sick is independent of the overall GID rate, regardless of the general effect of GID rates on this likelihood. This would be reassuring, since it would indicate that the mapping of GID prevalence to our clinic reports does not change with changes in GID rates.

Panel A in Table D1 shows the results from estimating equation D1, with an indicator for seeking medical care at a public clinic as the dependent variable. We begin in column 1 by simply showing the correlation between the likelihood of seeking care at a public clinic and being sick with a GID. In columns 2 and 3, we successively add the GID rate and base controls, as well as municipality and survey round fixed effects. These three columns show a positive and significant link between being sick with a GID and seeking care at a public clinic. The magnitude is relatively stable, increasing the probability of care by 30 percentage points. Columns 2 and 3 also indicate that an additional GID case per 1,000 individuals in a given municipality-year is associated with a small but significant increase in the likelihood of seeking care for any kind of disease.

Column 4 adds the interaction between the indicator for whether the individual is sick with a GID and the local GID rate. The estimate is small, negative and statistically indis-

tinguishable from zero. Column 5 includes additional individual-level controls. The results remain unchanged. The fact that we do not find a significant coefficient, and that the point estimate is negative and not positive, suggests that there is no differential change in the likelihood of seeking care at a public clinic when sick with a GID as the local prevalence of GIDs increases. As such, this provides reassurance that the fact that we observe GID visits at public clinics, instead of the full prevalence of GIDs, does not introduce an important bias in our results.

A related concern would be that individuals alter their decisions with respect to seeking private care. Panel B in Table D1 shows similar results, using an indicator for seeking care at a private clinic as the dependent variable. We again find that being sick with a GID increases the probability of seeking private care, as does the local GID rate. We do not find significant effects of the interaction term at the conventional 95% confidence level. This suggests that there are no differential changes in seeking private care when the local prevalence of disease changes.

Table D1:
Relationship Between Seeking Medical Attention
and Being Sick with a GID

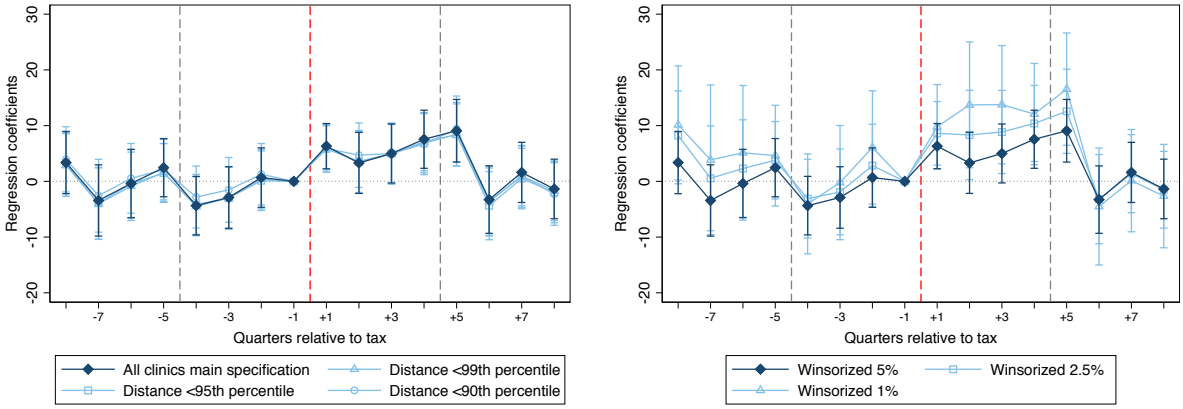
	(1)	(2)	(3)	(4)	(5)
Panel A: Seeking Attention at a Public Clinic					
Sick with GID	0.3071***	0.3039***	0.3042***	0.3239***	0.3094***
	(0.0108)	(0.0108)	(0.0107)	(0.0175)	(0.0175)
GID rate per 1,000		0.0004***	0.0005***	0.0005***	0.0004***
		(0.0001)	(0.0002)	(0.0002)	(0.0001)
Sick with GID × GID rate				-0.0023	-0.0027
				(0.0019)	(0.0018)
Observations	401,450	401,450	401,450	401,450	363,074
R-squared	0.0153	0.0185	0.0273	0.0274	0.0374
Panel B: Seeking Attention at a Private Clinic					
Sick with GID	0.1852***	0.1844***	0.1848***	0.1881***	0.1913***
	(0.0082)	(0.0082)	(0.0082)	(0.0116)	(0.0122)
GID rate per 1,000		0.0003***	0.0004***	0.0004***	0.0003***
		(0.0001)	(0.0001)	(0.0001)	(0.0001)
Sick with GID × GID rate				-0.0004	-0.0019*
				(0.0010)	(0.0010)
Observations	401,450	401,450	401,450	401,450	363,074
R-squared	0.0115	0.0133	0.0200	0.0200	0.0219
Municipality FE			X	X	X
Year FE			X	X	X
Base controls		X	X	X	X
Additional controls					X

Notes: This table shows the correlation between seeking medical attention and being sick with a GID, using data from the 2006 and 2012 ENSANUT survey rounds and estimating equation D1. Panel A focuses on public clinics, and Panel B on private care. Observations are individuals in a given municipality-year. The dependent variable is an indicator for seeking medical attention at a public clinic (or private clinic) for any symptoms, unconditional on reporting being sick. GID rate per 1,000 is the prevalence of GID rates in a given municipality-year. Base controls include age, gender, whether the individual lives in a house with a dirt floor, electricity, piped water, and sewerage, as well as municipality-year level averages of these last four characteristics. Additional controls, for which a few missing values are recorded, include education indicators and indicators for health insurance status. Robust standard errors clustered at the municipality level.
*** p<0.01, ** p<0.05, * p<0.1

E Graphical Estimates of the Robustness Checks and Additional Results

This appendix shows a graphical version of the robustness checks and additional results reported as point estimates in table format in the main text.

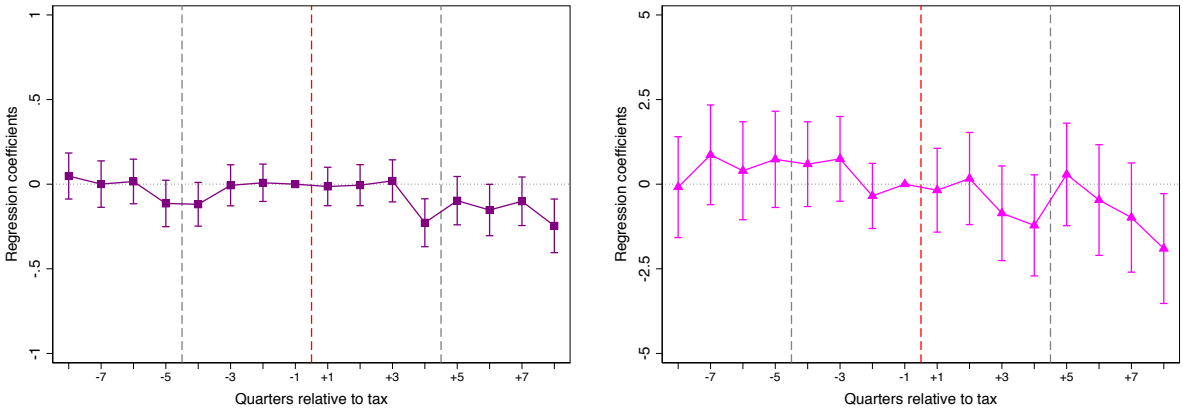
Figure E1:
DD Effect of the Soda Tax on GID Rates: Robustness Checks on Distance to the Monitoring Stations and Winsorization Levels



(a) Excluding large distances (b) Alternative winsorization levels

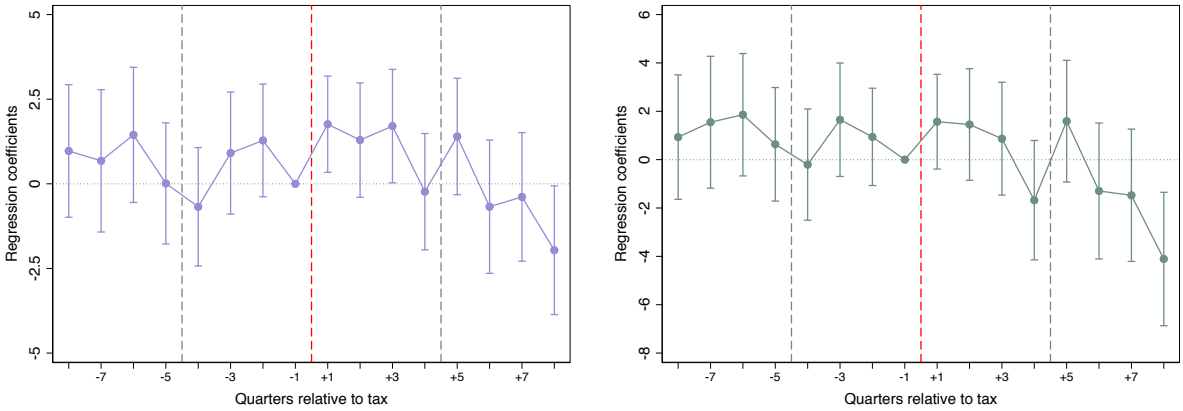
Notes: These graphs show robustness checks on the main results. The graph on the left plots estimates from equation 3 on a balanced panel of outpatient clinic-quarters, excluding clinics assigned to far away monitoring stations. The dark coefficient series corresponds to the main result including all clinics. The lighter-colored series exclude clinics by distance to the assigned monitoring station: top 1% (distance greater than 97 km; 156 clinics excluded), top 5% (distance greater than 57 km; 781 clinics excluded), and top 10% (distance greater than 44 km; 1,563 clinics excluded). The plot on the right considers winsorizing the outcome variable (clinic GID rate per 100,000) at different levels. The dark coefficient series corresponds to the outcome in the main specification, winsorized at the 5%. The lighter-colored series winsorize the outcome at more conservative levels: 2.5% and 1%. Coefficients for the interaction of the treatment indicator and each quarter for two years before and after the tax was introduced are shown, with quarter 4 of 2013 as the excluded period. Robust standard errors clustered at the clinic level. Error bars show 95% confidence intervals. The mean of the dependent variable for the treatment clinics prior to the tax for the baseline specification is 110.

Figure E2:
DD Effect of the Soda Tax on GID Rates: Placebo Check on Unrelated Diseases and Conditions



(a) Accidents and external injury

(b) Sexually transmitted diseases

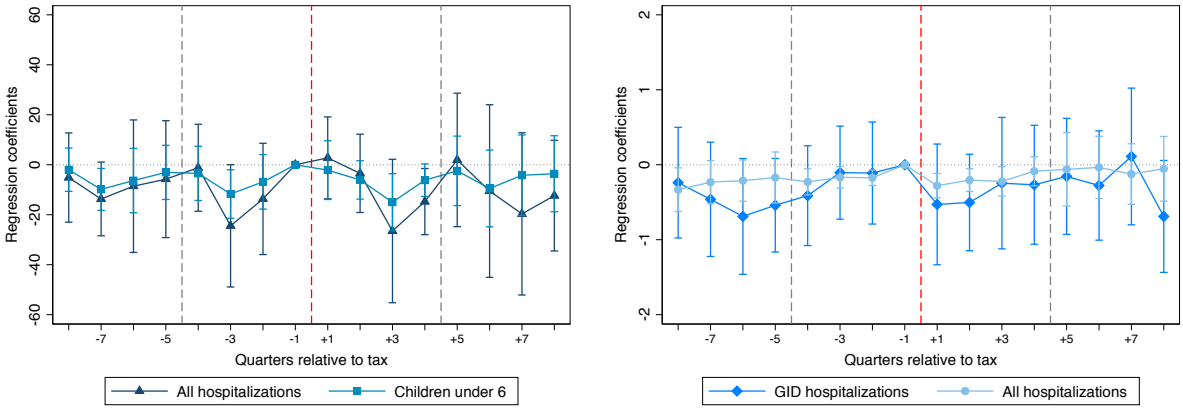


(c) Chronic diseases

(d) Any of the above

Notes: These graphs show placebo tests of the main results from estimating equation 3 on a balanced panel of outpatient clinic-quarters (15,634 clinics $\times$ 28 quarters) for unrelated conditions. The outcomes considered are accidents and external injuries, sexually transmitted diseases, chronic diseases, and the sum of all three. All outcomes are measured as rates per 100,000, winsorized at the 5%. Coefficients for the interaction of the treatment indicator and each quarter for two years before and after the tax was introduced are shown, with quarter 4 of 2013 as the excluded period. Robust standard errors clustered at the clinic level. Error bars show 95% confidence intervals. The mean of the dependent variable for the treatment clinics prior to the tax is 6 for accidents, 15 for sexually transmitted diseases, 56 for chronic conditions, and 77 for the sum.

Figure E3:
DD Effect of the Soda Tax on GID Hospitalization Rates at SSA Hospitals



(a) GID hospitalization rates

(b) Length of hospital stays

Notes: These graphs show results on SSA hospitalizations from estimating equation 3 on a balanced panel of SSA hospital-quarters (766 hospitals × 28 quarters). The outcomes for the graph on the left are GID rates per 100,000 for the whole population and for children under 6 years old, winsorized at the 5%. The outcomes for the graph on the right are average lengths of stay for GID hospitalizations and all hospitalizations. Coefficients for the interaction of the treatment indicator and each quarter for two years before and after the tax was introduced are shown, with quarter 4 of 2013 as the excluded period. Robust standard errors clustered at the hospital level. Error bars show 95% confidence intervals. The mean of the dependent variable for the treatment hospitals prior to the tax is 72 for hospitalization rates, 35 for hospitalization rates of children under 6, 1.4 for length of hospital stays for a GID, and 2.6 for length of all hospital stays.