Association Between Seattle's Sweetened Beverage Tax and Change in BMI Among a Patient Population of Adults

THE EVALUATION OF SEATTLE'S SWEETENED BEVERAGE TAX

MARCH 2023

Jones-Smith, JC, Godwin J, Knox M, Mooney S, Walkinshaw LP, Arterburn D, Eavey J, Chan NL, Saelens BE.

Public Health
Seattle & King County

uwcphn

Seattle Children's®
HOSPITAL · RESEARCH · FOUNDATION
Research Institute

SUGGESTED CITATION

Jones-Smith, JC, Godwin J, Knox M, Mooney S, Walkinshaw LP, Arterburn D, Eavey J, Chan NL, Saelens BE. Association Between Seattle's Sweetened Beverage Tax and Change in BMI Among a Patient Population of Adults - The Evaluation of Seattle's Sweetened Beverage Tax. Report for City of Seattle and Seattle City Council. March 2023.

CONTACT INFORMATION

Jessica C. Jones-Smith, PhD
jjoness@uw.edu
415-613-6047

FUNDING SOURCE

This report is funded by City of Seattle Sweetened Beverage Tax Ordinance 125324.

CONTENTS

ASSOCIATION BETWEEN SEATTLE'S SWEETENED BEVERAGE TAX AND CHANGE IN BMI AMONG A PATIENT POPULATION OF ADULTS

EXECUTIVE SUMMARY

Objective.
This study used a quasi-experimental design to evaluate the association between Seattle's Sweetened Beverage Tax and changes in body mass index (BMI) - a measure of obesity, among a patient population of adults in Seattle and surrounding areas.

Methods.
We used data from electronic medical records for visits to Kaiser Permanente Washington Health System clinics from 2014-2019 to compare changes in BMI from before to after tax implementation for people aged 18-65 living in Seattle versus people living in three nearby Washington counties outside Seattle (King County (excluding Seattle), Pierce County, and Snohomish County).

We used statistical weighting in our analyses to account for measurable differences between people living in Seattle and in the surrounding areas that might also influence BMI change over time. We tested three main types of difference-in-difference statistical models (i.e., first difference, comparative interruptive time series, and event study), each with different strengths and weaknesses, to assess whether Seattle's Sweetened Beverage Tax was associated with a change in BMI for people living in Seattle compared to the matched sample of people living in the three nearby counties. All models compare the change in BMI among people living in Seattle from before to after the tax to the change in BMI among people living in the comparison areas over the same period. We hypothesized that those living in Seattle would have a lower increase in BMI than those living outside of Seattle from before to after the Sweetened Beverage Tax implementation.

Results.
The final sample consisted of 1,158,854 clinic visits from 98,787 unique Kaiser Permanente Washington Health System patients aged 18-65 years. Of these, 23,487 lived in Seattle and 75,300 lived outside of, but nearby Seattle.

Overall, BMI was increasing over time in this cohort of patients, in both Seattle and the comparison area. However, our results suggest that the tax was associated with modestly lower increases in BMI for people in Seattle. Specifically, in our first difference models, where each person contributes one BMI data point before (closest to but before December 31, 2017) and one taken furthest out from the tax implementation (closest to December 31, 2019) suggest an association between the tax implementation and lesser BMI increases in Seattle relative to those in the comparison area. This finding generally held up in the sensitivity analyses.

The other two models (which include all BMI data points available in the dataset) were inconclusive as the estimates were sensitive to the structure of the data, whereby some patients contributed more BMI data points than others. In these models, when we updated the models to limit the analyses for a dataset where each patient had an equal number of BMI data points in the model, we found a suggested association in the same direction, but these were not statistically significant.

INTRODUCTION

Sweetened beverage taxes have been implemented in seven US cities and over 50 countries around the world.[1] The goals of these taxes have typically been to reduce consumption of sugar-sweetened beverages and reduce prevalence of associated health outcomes such as obesity, diabetes, and cardiovascular disease, as well as raise revenue for governments, which in U.S. cities, has frequently been used to fund health promotion programs.

In U.S. cities, these taxes have been generally successful at increasing the price of taxed beverages and reducing *purchasing* of these beverages.[2–8] For instance, Philadelphia's sweetened beverage tax was associated with a 38% net reduction in volume sold of taxed beverages in both large and chain stores[6] and small, independent stores.[7] In Seattle, the Sweetened Beverage Tax has been associated with a net 22% reduction in volume sold of taxed beverages.[8] In Berkeley, sales of taxed beverages declined by roughly 10% after a sweetened beverage tax was implemented; though for non-Berkeley stores sales increased by 6.9%.[9] Similarly, purchases of taxed beverages in Oakland were estimated to decrease by 6-8% in association with Oakland's sweetened beverage tax.[10] An analysis of household receipt data from Philadelphia, Seattle, Oakland, and San Francisco found that every one cent per ounce in sweetened beverage tax was associated with a 12% decline in ounces of taxed beverages purchased per household.[11]

The evidence for whether sweetened beverage taxes are effective in reducing reported *consumption* have been mixed,[12–15] ranging from findings of substantial and sustained reductions among residents in lower-income neighborhoods in Berkeley,[14] to no evidence of reduced consumption (relative to comparison areas without a tax) among several studies spanning Philadelphia,[3,15] Seattle,[16] Oakland[13] and Berkeley.[9] Indeed, a recent meta-analysis found no significant impacts of sugar-sweetened beverages taxes on reported consumption of sugar-sweetened beverages.[17]

There are several reasons why studies might find that *purchasing* of taxed beverages may consistently decrease, but *consumption* of these beverages may not decrease. In particular, if consumers decrease purchasing at the retailers included in large-scale retail scanner databases (the most common way of assessing reductions in purchasing) but increase their consumption of the sugar-sweetened beverages from restaurants or small stores (both of which are either not captured or less likely to be captured in most purchasing studies), this could result in a decrease in purchasing from retailers, but no net change in consumption. However, a potentially more likely explanation for this apparent inconsistency in purchasing versus consumption impact lies in the difficulty in measuring dietary intake. For instance, Lawman et al[12] found inconsistencies in recall of beverage intake using the BEVQ administered to a longitudinal cohort compared to receipts collected from the same individuals.[12] In a cohort study of Seattle's Sweetened Beverage Tax, Saelens and colleagues also used a modified BEVQ in a longitudinal cohort and found decreases in volumes of all beverages, including water, in both Seattle and the comparison area.[16] The decreases in consumption of all beverages measured in the BEVQ are unexplained and could signal problems with the dietary measurement tool used. Additionally, qualitative interviews with the participants in this study found that participants perceived that the study itself and the relatively long questionnaire on beverage intake (a modified BEVQ) made them more aware of their beverage intake and prompted them to decrease consumption (see Understanding Behavior Changes in the SeaSAW Cohort: Qualitative Follow-up Interviews). Taken together, these findings suggest that consumption has proven difficult to measure in studies of sweetened beverage taxes.

Taking the purchasing and consumption evidence together, it is important for studies to evaluate whether these taxes are having the intended impacts on population health outcomes. Before taxes were adopted, simulation studies suggested that these taxes could result in decreases in average BMI and obesity prevalence.[18–20] Indeed, these simulation studies were an instrumental factor in these taxes gaining traction as a possible policy option among researchers and public health advocates.[21–23] To date, only two studies to our knowledge have examined the health outcome impacts of sweetened beverage taxes. In a longitudinal study utilizing health record data in Mexico, which passed a sweetened beverage tax in 2014, Gračner and colleagues found reductions in obesity prevalence among adolescent girls, but not boys, within Mexican jurisdictions that had higher versus lower increases in the price of sugar-sweetened beverages in association with the tax.[24] Using surveillance data from the Youth Behavioral Risk Factor Survey, Flynn found evidence that the sweetened beverage tax in Philadelphia, San Francisco, and Oakland reduced average BMI among youth, compared to youth in comparison areas.[25] Interestingly, decreases in sugar-sweetened beverage consumption were only detectable in Philadelphia in the Flynn study. A limitation of the Flynn study is that height and weight was self-reported, and the sample used repeated cross-sectional data, which could be subject to changes in the characteristics of the respondents over time.

We aimed to build on the recent work to identify the health impacts of sweetened beverage taxes by assessing whether Seattle's Sweetened Beverage Tax was associated with change in BMI and BMI trajectories over time among adults.

METHODS AND RESULTS

Below we describe key elements of our methods and results.

Study Design.

On January 1, 2018, Seattle implemented an excise tax of 1.75 cents per ounce, levied on sugar-sweetened beverages with added sugar distributed within the city, excluding sweetened milk and 100% juice. We used a quasi-experimental design to evaluate the association between the Sweetened Beverage Tax in Seattle and change in adult BMI. Using data from electronic medical records, we compare changes in BMI from before to after tax implementation for people living in Seattle to the analogous changes for people living in a comparison sample. The comparison sample consists of people accessing the same health system but living outside of Seattle in nearby cities and towns within three Washington counties.

We used fine stratification propensity score weighting to achieve balance on measurable confounders for people living in Seattle and the comparison area (people living outside of Seattle in urban areas of the three surrounding counties). Confounders are factors that are correlated with living in Seattle and BMI change over time and can include things like age or neighborhood factors. Balancing on these factors is a way of eliminating their effect on the outcome so that we can isolate the effect of the tax, instead of measured confounders, on BMI.

We assessed the relationship between the tax and BMI using three types of difference-in-difference statistical models--first difference, comparative interruptive time series, and event study. Generally, the difference-in-difference estimates assess whether the tax was associated with a change in BMI for people living in Seattle compared to a well-matched population living in three nearby counties, comparing the change in BMI outcomes in Seattle from before to after the tax to the change in BMI outcomes in the comparison area over the same period.

The <u>first difference model</u> uses exactly two BMI measurements from each person and precisely controls for the amount of time elapsed between measurements. In addition, we take the measurement closest to, but not after the beginning of the tax for the pre-tax period measurement, and the one furthest out from tax implementation for the post-tax period. We compare each individual to themselves over time, thus additionally controlling for individual characteristics that do not change (like genetics).

The <u>comparative interruptive time-series model</u> uses all of the available BMI measurements from each patient. The model estimates slopes of BMI for Seattle and the comparison area, pre-tax, and the change in these slopes in the post-tax period for both Seattle and the comparison area. This model imposes a linear relationship between the tax and change in BMI.

The <u>event study</u> also uses all the available BMI measurements from each person. Instead of modeling a linear relationship, this model estimates a difference-in-difference for BMI change over time for Seattle versus the comparison area, for both the pre-tax and post-tax years. It creates an estimate separately for each year, thus allowing for the association of the tax to vary over time.

More details on each of these models and their interpretation are provided alongside the results of each model below.

Exposure.

Our exposure of interest was Seattle's Sweetened Beverage Tax. We geocoded patient residential addresses at the time of each visit in the electronic health records to obtain each patient's home census tract.[1] We classified those living in census tracts that had any area inside Seattle as exposed to the tax and those living in all other areas as unexposed to the tax (our comparison area).

Outcome.

Our outcome of interest was body mass index (BMI, weight in kilograms/height in meters squared). Weight and height were obtained from Kaiser Permanente Washington Health System electronic health records.[2]

Study Sample.

We used data extracted from electronic health records from the Kaiser Permanente Washington Health system. The study sample was limited to people residing in three counties: King, Pierce, and Snohomish. Within these counties, we limited the sample to people residing in places classified by the 2010 Census as either an urbanized area (population of 50,000 or more) or an urban cluster (population between 20,000 and 50,000) who were between the ages of 18-65 years at any point during the period from January 1, 2014, to December 31, 2019, and who received primary care at the Kaiser Permanente Washington Health System. Additionally, we excluded individuals who had a cancer diagnosis a year prior to, or any time during our observation period (since many cancers and cancer treatments can result in unintentional weight loss), or bariatric surgery at any time during their history with the health system (since this also typically results in dramatic weight loss and may be differentially distributed between Seattle and surrounding areas). We excluded observations that occurred within nine months prior and three months after a live birth since these observations are likely not reflective of usual body weight. We excluded people who either moved out of our study area or moved between the taxed jurisdiction (Seattle) and non-taxed areas during the study period. Additionally, we limited the sample to people who had at least one height and weight measurement before and after Seattle's Sweetened Beverage Tax was implemented. Finally, we removed those individuals whose weight was less than 50 lbs. or greater than 700 lbs., whose modal observed height was below 4 ft or above 8 ft, and those whose BMI calculated from observed weight and modal observed height was outside the range of 15-90.

The final sample consisted of 1,158,854 observations from 98,787 unique patients. Of these, 23,487 patients lived in Seattle and 75,300 lived outside of Seattle but within one of the three surrounding counties (King County (excluding Seattle), Pierce County, and Snohomish County).

Table 1 displays sample characteristics of patients living in Seattle and the comparison area. We show the results prior to statistical weighting (unweighted) and those after the statistical weighting to balance demographic characteristics have been applied (called "FSATE weighted" in tables). The unweighted populations were modestly different on some individual-level (e.g., age) demographic characteristics and substantially different on some area-level characteristics (census tract-level racial/ethnic composition).

Specifically, modest differences were seen in individual-level age composition, with patients living in Seattle versus the comparison area more likely to be in the younger age categories and less likely to be in older age categories.

Additionally, the Seattle sample as compared to the comparison area had a lower proportion of the population who self-reported as Hispanic, American Indian/Alaskan Native, Asian, and Black/African American, and a higher proportion who self-reported as and White (66.0% vs. 61.9%). There were also modest differences in patient insurance type (commercial insurance: 80.2% vs. 84.1%; Medicare: 1.2% v. 1.9%; Medicaid: 2.9% vs. 3.6%; other: 38.1% vs. 36.3%).The numbers can

---

[1] Using boundaries from the 2010 US Census. A census tract typically contains about 4000 people.

[2] A modal height value was previously created for all patients and this value was used to avoid changes in height due to measurement error. In some of our analyses, we use pre-calculated BMI change per year to account for differences in time elapsed between patient visits, taking the post-tax BMI minus the pre-tax BMI, and dividing by months elapsed between measurements multiplied by 12.

add to more than 100% because the other insurance category includes insurance that is sometimes used in combination with commercial insurance.

Census tract racial composition indicators, which reflect the neighborhoods where people live, rather than the racial composition of the people in our sample, were also modestly different between Seattle versus the comparison area for the proportion who were Hispanic (6.7% vs. 10.4%), Black (7.9% vs. 5.6%), American Indian/Alaskan Native (0.5% vs. 0.8%), Asian (14.2% vs. 12.4%), and Native Hawaiian/Other Pacific Islander (0.5% vs. 1.0%). There were also modest differences in the average tract-level percent of the population who moved in that last year (21.2% vs. 16.7%) and in the average tract-level percent of the population living in poverty.

There were substantial differences between Seattle versus the comparison area in the average tract-level percent of the population with a college degree or higher (57.8% vs. 33.6%) and in the tract-level population density (11.2 vs. 4.0).

TABLE 1. CHARACTERISTICS OF PATIENTS RESIDING IN SEATTLE AND THE COMPARISON AREA, UNWEIGHTED AND WEIGHTED (N=98,787)				
CHARACTERISTICS	UNWEIGHTED		FSATE WEIGHTED	
	SEATTLE	COMPARISON	SEATTLE	COMPARISON
	PERCENT (SE)			
SEX				
FEMALE	60.7% (0.3%)	61.3% (0.2%)	60.5% (0.9%)	60.7% (0.5%)
MALE	39.3% (0.3%)	38.7% (0.2%)	39.5% (0.9%)	39.3% (0.5%)
AGE				
18-29	19.9% (0.3%)	18.4% (0.1%)	16.8% (0.7%)	19.3% (0.5%)
30-39	24.4% (0.3%)	18.9% (0.1%)	22.8% (0.8%)	20.8% (0.5%)
40-49	22.2% (0.3%)	23.4% (0.2%)	23.8% (0.8%)	22.7% (0.4%)
50-59	25.6% (0.3%)	31.2% (0.2%)	28.8% (0.8%)	29.3% (0.4%)
60-65	7.9% (0.2%)	8.1% (0.1%)	8.0% (0.5%)	8.0% (0.3%)
RACIAL AND ETHNIC IDENTITY				
HISPANIC	5.8% (0.2%)	7.0% (0.1%)	7.6% (0.5%)	6.7% (0.2%)
NON-HISPANIC, AMERICAN INDIAN/ALASKAN NATIVE	0.5% (0.0%)	0.6% (0.0%)	0.7% (0.2%)	0.6% (0.1%)
NON-HISPANIC, ASIAN	12.4% (0.2%)	13.9% (0.1%)	17.1% (0.7%)	15.5% (0.4%)
NON-HISPANIC, BLACK, OR AFRICAN AMERICAN	6.8% (0.2%)	7.0% (0.1%)	7.2% (0.5%)	6.8% (0.2%)
NON-HISPANIC, MORE THAN ONE RACE	3.0% (0.1%)	3.3% (0.1%)	3.1% (0.3%)	3.1% (0.2%)
NON-HISPANIC, NATIVE HAWAIIAN, OR OTHER PACIFIC ISLANDER	0.5% (0.0%)	1.4% (0.0%)	1.2% (0.2%)	1.2% (0.1%)
NON-HISPANIC, OTHER	1.7% (0.1%)	1.7% (0.0%)	2.0% (0.3%)	1.9% (0.2%)
NON-HISPANIC, UNKNOWN, OR NOT REPORTED	3.3% (0.1%)	3.1% (0.1%)	3.5% (0.3%)	3.3% (0.2%)
NON-HISPANIC, WHITE	66.0% (0.3%)	61.9% (0.2%)	57.6% (0.9%)	60.9% (0.5%)
INSURANCE TYPE				
COMMERCIAL INSURANCE	80.2% (0.3%)	84.1% (0.1%)	82.1% (0.7%)	81.7% (0.5%)
MEDICARE	1.2% (0.1%)	1.9% (0.1%)	1.7% (0.3%)	1.7% (0.1%)
MEDICAID	2.9% (0.1%)	3.6% (0.1%)	3.8% (0.4%)	3.8% (0.2%)
OTHER INSURANCE	38.1% (0.3%)	36.3 (0.2%)	38.0% (0.9%)	39.5% (0.6%)
CENSUS TRACT RACE/ETHNICITY				
TRACT % HISPANIC	6.7% (0.0%)	10.4% (0.0%)	11.5% (0.2%)	9.3% (0.1%)
TRACT % AMERICAN INDIAN/ALASKAN NATIVE	0.5% (0.0%)	0.8% (0.0%)	0.7% (0.0%)	0.6% (0.0%)
TRACT % ASIAN	14.2% (0.1%)	12.4% (0.0%)	16.9% (0.3%)	16.4% (0.3%)
TRACT % BLACK	7.9% (0.1%)	5.6% (0.0%)	7.5% (0.1%)	5.7% (0.1%)
TRACT % TWO OR MORE RACIAL GROUPS	5.2% (0.0%)	5.3% (0.0%)	6.2% (0.1%)	5.1% (0.0%)
TRACT % NATIVE HAWAIIAN/OTHER PACIFIC ISLANDER	0.5% (0.0%)	1.0% (0.0%)	1.2% (0.0%)	0.8% (0.0%)
TRACT % OTHER RACIAL GROUP	0.2% (0.0%)	0.2% (0.0%)	0.1% (0.0%)	0.1% (0.0%)
TRACT % WHITE	64.8% (0.1%)	64.3% (0.1%)	56.0% (0.4%)	61.9% (0.2%)
CENSUS TRACT CHARACTERISTICS				
TRACT DENSITY (PEOPLE PER SQUARE MILE DIVIDED BY 1000)	11.2 (0.1)	4.0 (0.0)	6.1 (0.1)	6.1 (0.1)
TRACT % LIVING IN POVERTY	12.9% (0.1%)	11.2% (0.0%)	13.3% (0.2%)	11.4% (0.0%)
TRACT % OF PEOPLE WHO MOVED IN LAST YEAR	21.2% (0.1%)	16.7% (0.0%)	19.0% (0.1%)	20.0% (0.2%)
TRACT % OF PEOPLE WITH COLLEGE OR MORE	57.8% (0.1%)	33.6% (0.1%)	43.4% (0.3%)	40.5% (0.4%)

Some items are not mutually exclusive, and may sum to >100%

FSATE stands for 'Fine Stratification Average Treatment Effect' weight; characteristics included are all of the characteristics listed in Table 1.

Table 1 also displays the sample characteristics after applying the fine stratification average treatment effect (FSATE) weights, which are explained in detail in the next section. Weighting the sample using the FSATE weights results in the characteristics of Seattle and the comparison area samples being well balanced on observed variables, in which mean levels of some variables, particularly the ones that were highly unbalanced (tract density and tract education levels) fall in between the unweighted values of the two places.

STATISTICAL ANALYSES FOR IDENTIFYING THE IMPACT OF THE SBT ON BODY MASS INDEX

Treatment models

Our primary models use the FSATE weights described above to balance sample characteristics and thereby control for confounding by the individual and area level demographic characteristics used to construct the weights. We then use three different types of statistical models to estimate the impact of the tax on change in body mass index between Seattle and the comparison area from before to after the tax went into effect in Seattle.

Model 1: First Difference Model

The first model was a simple “first difference” model. Specifically, for each individual we took their BMI closest to but not after December 31, 2017 (last day before the tax went into effect), for the pre-period BMI and subtracted this from their BMI measurement closest to December 31, 2019 (the day after the last day in the post period) and not before tax implementation. We then divided this by the number of months between the measurement dates and multiplied by 12 to create our annualized BMI change measure that is used as the outcome in this first difference model. We then modeled annualized BMI change as a function of residence in Seattle or the comparison area. This model includes time-varying insurance type (e.g., to account for whether an individual changed type of insurance from one point to the next, which may be a proxy of socioeconomic resources) as a confounder in the model (this is the only time-varying individual level variable available).

The result from this model can be interpreted as a difference-in-difference estimate--the BMI change from before to after the tax in Seattle, beyond the difference seen in the comparison area.

Table 2 displays the results from the difference-in-difference estimator from the first difference models using the FSATE weights to control for confounding.

The results suggest a statistically significantly lower average annual change in BMI for patients living in Seattle as compared to those living in the comparison area (FSATE: -0.031 (95% CI: -0.063, -0.0002); p-value 0.049). This would suggest the tax was associated with a small but statistically significant reduction in BMI gains (-0.031 kg/m^2 per year).

The changes in BMI in the comparison area over this time period are shown in **Supplemental Table 3**; BMI increased on average at a pace of 0.166 BMI units per year. Said differently, the model suggests tax was associated with preventing about 19% of the expected BMI gain per year ((-0.031/0.166)*100=19%). For an individual with a BMI of 27 who weighed 167 pounds and was 5 foot 6 inches tall this would equate to reduction in BMI gains of 0.19 pounds per year.

Given that this is a population-level analysis not limited to people who necessarily consumed sugar-sweetened beverages before the tax was implemented, and therefore, many for whom the tax would have no impact on consumption and ultimately BMI, this is a reasonable level of attenuation of BMI gains at the population level.

To probe the degree to which the tax may have been more effective in tempering BMI increases for some populations compared to others, we ran the weighted model above separately for different population groups. Specifically, we ran models separately by age, sex, racial and ethnic identity, neighborhood level poverty, and neighborhood level education level. These results for the FSATE weighted model are also shown in **Table 3**, which displays the population group-specific estimates from separate stratified models.

The stratified models suggest that for many demographic groups the direction of the estimate of the association of the tax with BMI was negative such that those in Seattle gained less BMI than those in the comparison area from before to after the tax, with the relationship strongest for women, those living in neighborhoods with higher overall education status, and those living in neighborhoods with lower poverty levels. Of note as well is that the coefficients were negative and of largest magnitude for people identifying as American Indian/Alaskan Native and Black, however these did not reach statistical significance.

Finally, patients who did not have particularly high frequency of visits in the health care system (<10 per year) had results consistent with the overall estimate, however for those who visited very frequently, the effect estimate was in the opposite direction and not statistically significant. We performed additional sensitivity tests on these models, which are described in the Appendix and shown in **Supplemental Table 6**.

TABLE 2. DIFFERENCE-IN-DIFFERENCE ESTIMATES FROM FIRST DIFFERENCE MODELS OF ANNUALIZED BMI FOR SEATTLE VERSUS THE COMPARISON AREA, OVERALL AND STRATIFIED BY POPULATION GROUPS

		DIFFERENCE-IN-DIFFERENCE		
		COEFFICIENT (SE)	P-VALUE	N
OVERALL		**-0.031* (-0.063, -0.0002)**	**0.049**	**98,787**
STRATIFIED ESTIMATES				
SEX	FEMALE	-0.043* (-0.085, -0.002)	0.041	60,417
	MALE	0.031 (-0.032, 0.093)	0.338	38,370
RACE/ETHNICITY	HISPANIC	0.013 (-0.105, 0.132)	0.824	6,632
	NON-HISPANIC, AMERICAN INDIAN/ALASKAN NATIVE	-0.269 (-0.623, 0.085)	0.136	597
	NON-HISPANIC, ASIAN	-0.005 (-0.141, 0.132)	0.947	13,369
	NON-HISPANIC, BLACK, OR AFRICAN AMERICAN	-0.083 (-0.253, 0.087)	0.338	6,896
	NON-HISPANIC, MORE THAN ONE RACE	0.002 (-0.225, 0.228)	0.989	3,217
	NON-HISPANIC, NATIVE HAWAIIAN, OR OTHER PACIFIC ISLANDER	-0.064 (-0.663, 0.535)	0.834	1,209
	NON-HISPANIC, OTHER	-0.03 (-0.242, 0.181)	0.778	1,679
	NON-HISPANIC, UNKNOWN, OR NOT REPORTED	-0.015 (-0.199, 0.169)	0.875	3,083
	NON-HISPANIC, WHITE	-0.059 (-0.184, 0.066)	0.357	62,105
AGE GROUP	18-29	-0.057 (-0.183, 0.069)	0.377	12,421
	30-49	0.005 (-0.132, 0.142)	0.941	38,635
	50-66	0.057 (-0.074, 0.188)	0.394	47,731
POVERTY	LOW POVERTY	-0.036* (-0.068, -0.005)	0.024	92,129
	HIGH POVERTY	0.027 (-0.086, 0.14)	0.638	6,658
EDUCATION	HIGH EDUCATION	-0.045* (-0.08, -0.01)	0.012	70,403
	LOW EDUCATION	0.028 (-0.04, 0.095)	0.420	28,384
FREQUENCY OF VISITING HEALTH CARE PROVIDER	NON-HIGH FREQUENCY	-0.034+ (-0.07, 0.0001)	0.050	79,198
	HIGH FREQUENCY	0.0322 (-0.05, 0.12)	0.452	19,589

*Notes: * indicates statistical significance at 0.05 level; + indicates statistical significance at 0.10 level*

Model 2: Comparative interruptive time-series Model
Using a different model that incorporates as many observations as possible and takes advantage of the multiple measurements both before and after the tax for a large proportion of the sample in the electronic health record, we conducted an additional analysis. In essence, the model controls for each person's starting BMI and the estimates are of the change in BMI from that starting level over the time period of the study. More details of the model equation can be found in the Appendix. In addition to utilizing the individual-level fixed effects and the FSATE-weighted model to balance confounders, we also include time-varying insurance status as a covariate in these models as in the models above.

Table 3 displays the results of the comparative interruptive time-series model with individual fixed effects, which utilize all BMI observations from individuals who had at least one BMI measurement pre-tax and one BMI post-tax, and also incorporates all other observations from each individual included in the sample. Because our health outcomes are from electronic health records, patients have their BMI measured almost every time they see a health care provider. Thus, the model is unbalanced, meaning that individuals contribute different quantities of BMI observations and BMI is not measured at the same calendar time for everyone.

The difference-in-difference estimator from this model suggests no association between the tax and change in BMI.

TABLE 3. ESTIMATED ASSOCIATION OF SEATTLE'S SWEETENED BEVERAGE TAX WITH CHANGE IN BMI USING A COMPARATIVE INTERRUPTIVE TIME SERIES MODEL, FSATE WEIGHTED AND INTERPOLATED RESULTS

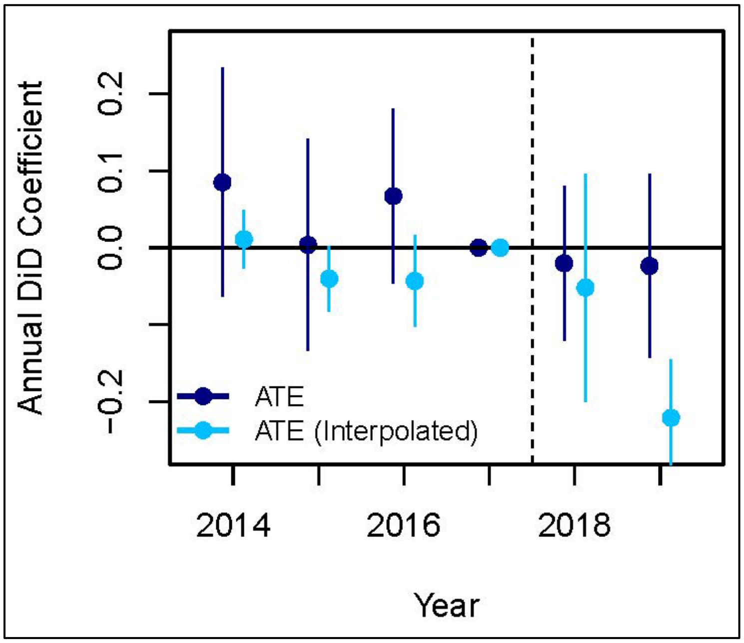
DIFFERENCE-IN-DIFFERENCE	FSATE		INTERPOLATED FSATE	
	COEFFICIENT (95% CI) | P-VALUE			
DIFFERENCE IN THE CHANGE IN THE TIME TREND IN BMI FOR SEATTLE COMPARED TO THE CHANGE IN THE TIME TREND IN THE COMPARISON AREA FROM BEFORE TO AFTER THE TAX	-0.001 (0.025, -0.027)	0.929	-0.005 (-0.015, 0.005)	0.31

Due to the highly unbalanced nature of the data (meaning that people contributed differing numbers of observations with different time-spacing between observations), we re-ran these models using a dataset with BMI linearly interpolated to create a quarterly BMI measurement between every two measurements. This linear interpolation allows for a balanced panel with measurements taken at the same time for all individuals. This resulted in every patient having six measurements. In these models, the estimate for the tax effect is of larger magnitude and the confidence interval substantially smaller (Interpolated model: -0.005 (95% CI: -0.015, 0.005); p-value 0.31). The fact that the interpolated results are substantively different (five times larger magnitude) suggests that perhaps the unbalanced nature of the health records data is influential in the results. That is, it may be that patients who have more ambulatory care visits (and thus contribute more observations to the data) are less affected by the tax than those with fewer visits. This is consistent with an interpretation that older patients, who on average seek clinical care more frequently, are less affected by the tax, whereas younger people and those living in neighborhoods with less poverty and with a higher percentage of people with a college degree seek care less frequently. Indeed, these are the populations whose BMI seems to be more affected by the tax according to the stratified results presented in **Table 2**.

Model 3: Event Study Model
Finally, we fitted an "event study" model, which instead of imposing a linear time trend pre-tax and post-tax, allows the trend in BMI to vary each year. It computes a difference-in-difference estimate comparing the change in BMI in Seattle to the change in the comparison area using the year before the tax as the referent year. It is different from the comparative interruptive time series model because it allows the tax association to vary in each year after the tax implementation, but it is similar in that it includes all BMI observations. We ran the event study models on the interpolated data as well due to the unbalanced nature of the panel. These results are shown in Figure 1, which display the difference-in-difference estimates and confidence intervals for both the pre-tax and post-tax periods. The estimates

are close to zero in all pre-tax periods, which indicates similar trends in the BMI between Seattle and the comparison area before the tax implementation. The estimates are negative in the post-tax period for the FSATE models, but confidence intervals are wide and cross zero. In the interpolated data, the estimates are still negative and are statistically significant in the second year after the tax. This suggests that the tax is associated with a tendency towards a lesser increase in BMI in Seattle versus the comparison area, particularly in the second year after the tax, but this is only apparent in the interpolated, balanced data set.

Figure 1. Event study difference-in-difference estimates for BMI in Seattle compared to the comparison area, using the difference-in-difference in 2017, the year before the tax as the reference period.



CONCLUSIONS

In summary, we tested three main types of statistical models, each with different strengths and weaknesses, to assess whether Seattle's Sweetened Beverage Tax was associated with a change in BMI for people living in Seattle compared to a well-matched population living outside of Seattle within the three nearby counties.

We find suggestive evidence that Seattle's Sweetened Beverage Tax was associated with reductions in BMI gains over time in our first difference models. However, the results from our two other models, the comparative interrupted time series and the event study models, were inconclusive and sensitive to the nature of medical record data, where some patients have many observations and others have far fewer.

Overall, one set of models (the first difference models) suggest a small beneficial association of the tax with BMI in that those living in Seattle had lesser increases in BMI from before to after the tax, which was robust to most sensitivity analyses, in that all but one found difference-in-difference estimates similar to the original analysis. The second and third set of models allowed for multiple observations before and after the tax and find no association between the tax and BMI. However, we found evidence in both of these multiple-observation models that the unbalanced nature of the data (i.e. that some people have many BMI measurements and some have few, and that people are not measured at the same time) appears to be influential in the results. In particular, when we use a statistical technique to linearly interpolate between every two observed BMI values so that all participants have an equal number of BMI measurements in the model, the estimates from these models substantively change, particularly for the event study models. In the case of the event study, the model with interpolated values suggests statistically significant beneficial association between the tax and BMI, two years after the tax. In the case of the comparative interruptive time series models, the estimated association is more negative (in the expected direction), but still not statistically significant.

Due to the fact that the multi-observation models are sensitive to the interpolation process (i.e. the results change), our preferred set of models are the first difference models. The benefit of the first difference model is that each person contributes exactly two BMI observations to the model, and, for each individual, we are able to preference the observation that is closest to immediately before the tax is implemented in the pre-tax period and the observation that is closest to the end of our study period. For these reasons, the model is less complicated than the other two model types. In addition, it best accounts for the fact that BMI change is often small, but cumulative, and the effect of the tax is likely demonstrated best at the most distal time after the tax to capture the accumulative effect of consumption on change in BMI.

Based on this model, Seattle's Sweetened Beverage Tax appears to have been associated with less weight gain among Seattle residents. Specifically, there was some evidence that after using statistical techniques to attain a well-matched sample of people in the comparison area who have similar characteristics to those in Seattle, when comparing individuals to themselves before and after the tax, that people living in Seattle experienced lower BMI increases compared to those living in the comparison area. The size of the association was small, amounting to a reduction in the annual rate of BMI increase by 20%, or amounting to 0.2 pounds per year for someone with a BMI of 27 weighing 167 pounds and 5 feet 6 inches tall, population-wide. This size of an effect is consistent with a recent study of post-tax data in four cities that estimated that the decrease in purchasing seen should result in a population-level BMI reduction of about 0.5 pounds a year,[11] and is generally consistent with an early simulation study that suggested a sweetened beverage tax that reduced consumption by 20% would be expected to lower mean BMI by 0.08 BMI units among adults.[18]

We additionally performed stratified analyses to assess whether the associations observed in the primary model appeared to be driven by any particular demographic groups. The beneficial association between the tax and BMI was strongest for women, people living in low poverty neighborhoods, and people living in neighborhoods with a higher proportion of residents who have a college degree or higher. While not statistically significant, there was a large magnitude of the association for American Indian/Alaskan Native and Black populations. This would suggest that these populations were more responsive to the tax. On the flipside, it is concerning that the tax did not have the same or greater impact for people living in higher poverty neighborhoods or neighborhoods with lower average education levels.

The one specification to which the first difference model results were not robust was using the FSATT weighting instead of the FSATE weighting. The FSATT weighting results had a markedly lower difference-in-difference estimate (although still negative) and did not indicate an impact of the SSB tax on BMI change among Seattle versus comparison area residents. This may be an artifact of the less robust approach – FSATT estimates are less precise than the FSATE results and may be more affected by statistical outliers because the individual statistical weights are much larger.

It still should be noted that we found no evidence of an association between the tax and BMI trajectories in a second and third set of models that utilized all the BMI measurements that a patient had recorded in their medical record (the comparative interruptive time series and the event study). It might be the case that the results from these may be more reflective of the experience of people who visit the doctor more frequently (those who are older and with more medical issues). The comparative interruptive times series model also may not be adequately flexible to pick up non-linear time trends. When we use interpolation to impute a BMI for all patients quarterly, the estimate of the tax effect on BMI is larger and in the direction of the first differences model, yet still not statistically significant. Whereas for the event study model, after interpolation, the estimated effect in the second year after the tax is statistically significant. Taken together, it seems that the unbalanced nature of these models is influential in the results. These combined results (from the first difference models and comparative interruptive time series models and event study models with the interpolation sensitivity analyses) could be consistent with the tax having no effect on BMI for people who are older and less healthy and visit the doctor frequently, but still potentially having an impact on the BMIs those who are younger, healthier and go to the doctor less frequently.

LIMITATIONS

We rely entirely on the timing of the tax and the limited taxed jurisdiction (Seattle) to identify the association of the tax with BMI. It is possible if another policy changed at the same time, affecting only Seattle, or Seattle more than the surrounding areas, and impacting BMI, we could be misattributing the effect to the effect of the tax. We are not aware of a policy or other change that might meet these criteria. Height and weight were measured during routine health care visits, not on research protocol, so there is likely more measurement error compared to if these heights and weights had been taken in duplicate by trained research staff. Our sample only examines those patients who seek health care frequently enough to be included in the sample and who were with the same health system long enough to be measured before and after the tax. The Kaiser Permanente Washington Health System patient sample is large and is representative of commercially insured adults and Medicare patients in Seattle and King County, but not representative of people with Medicaid insurance or without insurance.

Overall, we find some evidence that Seattle's Sugar-Sweetened Beverage tax may have had beneficial impacts on the BMI of adults in Seattle. This is fairly robust in the model that uses only two observations from every patient in the sample. This association was strongest for women, and people living in low poverty neighborhoods and those living in neighborhoods with a high percentage of highly educated people.

CITATIONS

1. Global Food Research Program UNC. *Sugary Drink Taxes around the World*. Chapel Hill; 2022. https://www.globalfoodresearchprogram.org/wp-content/uploads/2022/11/Sugary_Drink_Tax_maps_2022_11.pdf.
2. Jones-Smith JC, Pinero Walkinshaw L, Oddo VM, et al. Impact of a sweetened beverage tax on beverage prices in Seattle, WA. *Econ Hum Biol*. 2020;39:100917. doi:10.1016/j.ehb.2020.100917
3. Cawley J, Frisvold D, Hill A, Jones D. The impact of the Philadelphia beverage tax on purchases and consumption by adults and children. *J Health Econ*. 2019;67:102225. doi:10.1016/j.jhealeco.2019.102225
4. Falbe J, Rojas N, Grummon AH, Madsen KA. Higher retail prices of sugar-sweetened beverages 3 months after implementation of an excise tax in Berkeley, California. *Am J Public Health*. 2015;105(11):2194-2201. doi:10.2105/AJPH.2015.302881
5. Cawley J, Frisvold DE. The Pass-Through of Taxes on Sugar-Sweetened Beverages to Retail Prices: The Case of Berkeley, California. *J Policy Anal Manag*. 2017;36(2):303-326. doi:10.1002/pam.21960
6. Roberto CA, Lawman HG, Levasseur MT, et al. Association of a Beverage Tax on Sugar-Sweetened and Artificially Sweetened Beverages with Changes in Beverage Prices and Sales at Chain Retailers in a Large Urban Setting. *JAMA - J Am Med Assoc*. 2019;321(18):1799-1810. doi:10.1001/jama.2019.4249
7. Bleich SN, Lawman HG, LeVasseur MT, et al. The Association Of A Sweetened Beverage Tax With Changes In Beverage Prices And Purchases At Independent Stores. *Heal Aff*. 2020;39(7):1130-1139. doi:10.1377/hlthaff.2019.01058
8. Powell LM, Leider J. The impact of Seattle's Sweetened Beverage Tax on beverage prices and volume sold. *Econ Hum Biol*. 2020;37:100856. doi:10.1016/j.ehb.2020.100856
9. Silver LD, Ng SW, Ryan-Ibarra S, et al. Changes in prices, sales, consumer spending, and beverage consumption one year after a tax on sugar-sweetened beverages in Berkeley, California, US: A before-and-after study. *PLoS Med*. 2017;14(4):e1002283. doi:10.1371/journal.pmed.1002283
10. Leider J, Powell LM. Longer-term impacts of the Oakland, California, sugar-sweetened beverage tax on prices and volume sold at two-years post-tax. *Soc Sci Med*. 292:114537. doi:10.1016/j.socscimed.2021.114537
11. Cawley J, Frisvold D, Jones D. The impact of sugar-sweetened beverage taxes on purchases: Evidence from four city-level taxes in the United States. *Heal Econ*. 2020;29(10):1289-1306. doi:10.1002/hec.4141
12. Lawman HG, Bleich SN, Yan J, et al. One-year changes in sugar-sweetened beverage consumers' purchases following implementation of a beverage tax: a longitudinal quasi-experiment. *Am J Clin Nutr*. 2020;112(3):644-651. doi:10.1093/ajcn/nqaa158
13. Cawley J, Frisvold D, Hill A, Jones D. Oakland's sugar-sweetened beverage tax: Impacts on prices, purchases and consumption by adults and children. *Econ Hum Biol*. 2020;37:100865. doi:10.1016/j.ehb.2020.100865
14. Lee MM, Falbe J, Schillinger D, Basu S, Mcculloch CE, Kristine A. Sugar-Sweetened Beverage Consumption 3 Years After the Berkeley , California , Sugar-Sweetened Beverage Tax. *Am J Public Health*. 2019:1-3. doi:10.2105/AJPH.2019.304971
15. Zhong Y, Auchincloss AH, Lee BK, McKenna RM, Langellier BA. Sugar-Sweetened and Diet Beverage Consumption in Philadelphia One Year after the Beverage Tax. *Int J Env Res Public Heal*. 2020;17(4). doi:10.3390/ijerph17041336
16. Saelens B, Rowland M, Qu P, et al. *The Evaluation of Seattle's Sweetened Beverage Tax 12 Month Report: Store Audits & Child Cohort*. Seattle; 2020.
17. Andreyeva T, Marple K, Marinello S, Moore TE, Powell LM. Outcomes Following Taxation of Sugar-

Sweetened Beverages: A Systematic Review and Meta-analysis. *JAMA Netw open*. 2022;5(6):e2215276. doi:10.1001/jamanetworkopen.2022.15276
18. Long MW, Gortmaker SL, Ward ZJ, et al. Cost Effectiveness of a Sugar-Sweetened Beverage Excise Tax in the U.S. *Am J Prev Med*. 2015;49(1):112-123. doi:10.1016/j.amepre.2015.03.004
19. Vecino-Ortiz AI, Arroyo-Ariza D. A tax on sugar sweetened beverages in Colombia: Estimating the impact on overweight and obesity prevalence across socio economic levels. *Soc Sci Med*. 2018;209:111-116. doi:10.1016/J.SOCSCIMED.2018.05.043
20. Grummon AH, Golden SD. Sugar-Sweetened Beverage Pricing Policies: Simulation of Minimum Price Laws and Taxes in New York City. *Am J Prev Med*. 62(3):e159-e168. doi:10.1016/j.amepre.2021.08.029
21. Krieger J, Bleich SN, Scarmo S, Ng SW. Sugar-Sweetened Beverage Reduction Policies: Progress and Promise. *Annu Rev Public Health*. 2020;42:439-461. doi:10.1146/annurev-publhealth-090419-103005
22. Arantxa Colchero M, Rivera JA, Popkin BM, Ng SW, M. Arantxa Colchero, Juan A. Rivera, Barry M. Popkin SWN. Sustained consumer response: evidence from two-years after implementing the sugar sweetened beverage tax in Mexico. *Health Aff*. 2017;36(3):564-571. doi:10.1016/j.physbeh.2017.03.040
23. Backholer K, Sarink D, Beauchamp A, et al. The impact of a tax on sugar-sweetened beverages according to socio-economic position: a systematic review of the evidence. *Public Health Nutr*. 2016;19(17):3070-3084. doi:10.1017/S136898001600104X
24. Gračner T, Marquez-Padilla F, Hernandez-Cortes D, Gracner T, Marquez-Padilla F, Hernandez-Cortes D. Changes in Weight-Related Outcomes Among Adolescents Following Consumer Price Increases of Taxed Sugar-Sweetened Beverages. *JAMA Pediatr*. 2022;176(2):150-158. https://jamanetwork.com/journals/jamapediatrics/fullarticle/2786784. Accessed December 17, 2021.
25. Flynn J. Do sugar-sweetened beverage taxes improve public health for high school aged adolescents? *Health Econ*. 2022;(October 2021):47-64. doi:10.1002/hec.4609

APPENDIX

STATISTICAL WEIGHTING TO ACHIEVE A WELL-MATCHED COMPARISON GROUP

We used a weighting approach to attempt to balance potential confounders between the treatment groups (Seattle versus comparison). Specifically, we used a weighting method called “fine stratification weighting” since the distribution of the propensity scores was substantially skewed and the treated group was smaller than comparison group.[28] The fine stratification weights, which lower the likelihood of any individual having a weight close to 0 or 1, are created by first estimating the propensity or probability of being in the treatment group (live in a census tract in Seattle) using a logistic regression model, based on the characteristics in **Table 1**.[3] We then stratify the propensity scores into the following bins separately within Seattle and within the comparison area: [(0, 0.1), (0.1, 0.2), …, (0.9, 1)]. Every individual within a specific stratum receives the same weight based on whether they live in Seattle or the comparison area, which is the ratio of the proportion of all individuals whose propensity scores fall in that strata to the proportion of individuals in Seattle or the comparison area, respectively, in that strata.

We explored two weighting schemes in order to figure out which one would provide the best balancing and would match our desired theoretical treatment effect (average treatment effect or average effect of treatment on the treated, explained futher below). The first, FSATE weights (used to estimate the average treatment effect (ATE)), aim to balance the sample characteristics between Seattle (“treatment” group) and the comparison area (outside of Seattle) and resultingly make the weighted pseudo-population sample look more like the overall averages of the sample on covariates used to construct weights. The second, FSATT weights (used to estimate the average effect of treatment on the treated (ATT)), reweight only the comparison group, aiming to make the weighted comparison group most similar to the treatment group on covariates used to compute weights.

The ATE represents the average effect that would be expected if everyone in the population were exposed to the treatment (i.e. the tax). Broadly, the ATE is considered more informative in scenarios where everyone in the population could theoretically be exposed to the treatment, because it represents the most generalizable estimate least prone to error due to large weights. By contrast, the ATT is recommended in scenarios where some people could not be exposed to the treatment, or where one wants to assess the treatment effect on the population being analyzed without extending results more generally.

In practice, weighting the sample using the FSATT weights created a pseudo-population that looks more similar to the baseline characteristics in Seattle, but also did not achieve good balance on some characteristics, such as racial/ethnic categories at the individual and census tract level (% NH-Asian and % NH-White) (**Supplemental Table 2**). Therefore, based on the better theoretical match of the FSATE to our purposes and the better performance of the FSATE weight (based on the balance metrics presented in **Supplemental Table 2** and having less extreme weights (shown in **Supplemental Table 1**), the FSATE became our preferred weight and we present the statistical models using this weight.[4]

[3] The propensity score model included variables we hypothesized to be associated with both the propensity to live in Seattle and BMI outcomes. This included, at the individual level, age categories, race/ethnicity categories, insurance type, sex. We also included census tract characteristics: percent of the population who are tract who are AIAN, NHOPI, Black, Asian, and White (each as separate variables), proportion of the tract who moved in the previous year, proportion of tract with a college degree or higher, and population density of the tract. We additionally included several interaction terms: proportion of the tract with college degree or higher interacted with tract racial composition variables (as described above but also including tract proportion multiracial and tract proportion with a race not listed and tract proportion with unknown race), proportion with a college degree or higher interacted with tract population density, tract density interacted with racial composition variables.
[4] Supplemental Table 1 displays the propensity scores binned into the 10 strata used in the calculation of the fine stratification weights and shows the number of patients in Seattle and the comparison area whose propensity scores lie within that bin, as well as the corresponding FSATT and FSATE weights for each level for Seattle and the comparison area. Of note, the majority (81%) of people in the comparison area had propensity scores in the lowest level (0 to .1) whereas the majority (54%) of people in Seattle had propensity scores in the highest category (.9 to 1). However, there is still common support (people from both Seattle and comparison area in each propensity score category) across the range scores, despite the high level of skew. Additionally it can be seen that the FSATT weights are more extreme at the low and high ends.

SUPPLEMENTAL TABLE 1. NUMBER OF PATIENTS IN EACH PROPENSITY SCORE BIN BY SEATTLE RESIDENCY							
PROPENSITY	COMPARISON	SEATTLE	TOTAL	SEATTLE FSATE WEIGHTED	COMPARISON FSATE WEIGHTED	SEATTLE FSATT WEIGHTED	COMPARISON FSATT WEIGHTED
0,0.1	61,311	1,384	62,695	10.77	0.78	1	0.07
1.0,0.2	6,084	851	6,935	1.94	0.87	1	0.45
0.2,0.3	3,134	556	3,690	1.58	0.90	1	0.57
0.3,0.4	1,516	858	2,374	0.66	1.19	1	1.81
0.4,0.5	1,050	790	1,840	0.55	1.34	1	2.41
0.5,0.6	1,077	939	2,016	0.51	1.43	1	2.80
0.6,0.7	450	1,306	1,756	0.32	2.97	1	9.30
0.7,0.8	318	1,979	2,297	0.28	5.51	1	19.95
0.8,0.9	143	2,206	2,349	0.25	12.52	1	49.46
0.9,1.0	208	12,618	12,826	0.24	47.00	1	194.49

FSATE stands for 'Fine Stratification Average Treatment Effect'
FSATT stands for 'Fine Stratification Average Treatment Effect Among the Treated'

Supplemental Table 2 displays the standardized differences in levels of potential confounders (the variables in **Supplemental Table 1**) between Seattle and the comparison area. This is a metric by which the weighting approaches can be judged for the degree to which they balanced potential confounders. Based on this metric, both the FSATE and the FSATT have substantially lower standardized differences than the unweighted sample, but the FSATE has the lowest total summed standardized differences across all variables. For this reason, the FSATE weighting approach is our preferred weighting approach.

SUPPLEMENTAL TABLE 2. BALANCE METRICS: STANDARDIZED DIFFERENCES BETWEEN SEATTLE AND COMPARISON AREA CHARACTERISTICS COMPARING UNWEIGHTED, FSATE-WEIGHTED AND FSATT-WEIGHTED			
VARIABLE	UNWEIGHTED	FSATE WEIGHTED	FSATT WEIGHTED
FEMALE	-2.17	-0.33	1.24
MALE	2.17	0.33	-1.24
SUM OF ABSOLUTE VALUE OF STANDARDIZED DIFFERENCES	**4.34**	**0.66**	**2.48**
18-29	7.33	-4.42	-1.73
30-39	24.87	3.06	-1.82
40-49	-5.5	1.76	1.64
50-59	-24.04	-0.81	2.05
60-65	-1.6	0.03	0.64
SUM OF ABSOLUTE VALUE OF STANDARDIZED DIFFERENCES	**63.34**	**10.08**	**7.88**
HISPANIC	-9.41	2.35	0.31

NON-HISPANIC, AMERICAN INDIAN/ALASKAN NATIVE	-5.06	0.54	-0.03
NON-HISPANIC, ASIAN	-8.25	2.75	-6.54
NON-HISPANIC, BLACK OR AFRICAN AMERICAN	-2.12	1.29	1.53
NON-HISPANIC, MORE THAN ONE RACE	-3.41	-0.26	1.2
NON-HISPANIC, NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	-21.1	-0.51	-0.46
NON-HISPANIC, OTHER	-0.92	0.72	-1.47
NON-HISPANIC, UNKNOWN OR NOT REPORTED	3.15	0.7	-1.41
NON-HISPANIC, WHITE	16.44	-4.45	5.57
SUM OF ABSOLUTE VALUE OF STANDARDIZED DIFFERENCES	**69.86**	**13.57**	**18.52**
COMMERCIAL INSURANCE	-18.82	0.57	4.27
MEDICAID	-8.16	-0.13	-2.29
MEDICARE	-12.08	0.11	1.52
OTHER INSURANCE	6.95	-2.02	-7.77
SUM OF ABSOLUTE VALUE OF STANDARDIZED DIFFERENCES	**46.01**	**2.83**	**15.85**
TRACT % WHITE	4.31	-20.64	33.82
TRACT % BLACK	52.48	20.82	14.27
TRACT % AMERICAN INDIAN/ALASKAN NATIVE	-53.55	2.31	38.49
TRACT % ASIAN	29.94	2.04	-36.38
TRACT % NATIVE HAWAIIAN/OTHER PACIFIC ISLANDER	-78.68	13.28	5.27
TRACT % OTHER RACIAL GROUP	16.05	-1.66	43.6
TRACT % HISPANIC	-125.37	15.5	9.35
TRACT % TWO OR MORE RACIAL GROUPS	-7.35	21.88	24.59
TRACT DENSITY	185.06	0.27	-7.8
TRACT % LIVING IN POVERTY	37.49	14.6	6.08
TRACT % OF PEOPLE WHO MOVED IN LAST YEAR	91.29	-6.98	-33.75
TRACT % OF PEOPLE WITH COLLEGE OR MORE	267.8	8.29	-11.81
SUM OF ABSOLUTE VALUE OF STANDARDIZED DIFFERENCES	**949.37**	**128.27**	**265.21**
SUM OF ABSOLUTE VALUE OF STANDARDIZED DIFFERENCES	**1132.92**	**155.41**	**309.94**

Note: *Standardized Differences >10 are considered imbalanced*
Standardized differences were calculated according to these formulae:

$$Std\ Diff\ = \frac{\overline{x}_{Sea} - \overline{x}_{Not}}{\sqrt{0.5(s^2_{Sea} + s^2_{Not})}} \quad Std\ Diff\ = \frac{\overline{p}_{Sea} - \overline{p}_{Not}}{\sqrt{0.5\left(\frac{\overline{p}_{Sea}(1-\overline{p}_{Sea})}{n_{Sea}} + \frac{\overline{p}_{Not}(1-\overline{p}_{Not})}{n_{Not}}\right)}}$$

SUPPLEMENTAL TABLE 3. SAMPLE CHARACTERISTICS, INCLUDING FSATT WEIGHTED				
	UNWEIGHTED		FSATT	
	SEATTLE	COMPARISON	SEATTLE	COMPARISON
VARIABLE	PERCENT (SE)			
SEX				
FEMALE	60.7% (0.3%)	61.3% (0.2%)	60.7% (0.3%)	58.9% (2.1%)
MALE	39.3% (0.3%)	38.7% (0.2%)	39.3% (0.3%)	41.1% (2.1%)
AGE				
18-29	19.9% (0.3%)	18.4% (0.1%)	19.9% (0.3%)	22.2% (1.9%)
30-39	24.4% (0.3%)	18.9% (0.1%)	24.4% (0.3%)	26.9% (1.9%)
40-49	22.2% (0.3%)	23.4% (0.2%)	22.2% (0.3%)	20.3% (1.6%)
50-59	25.6% (0.3%)	31.2% (0.2%)	25.6% (0.3%)	23.1% (1.7%)
60-65	7.9% (0.2%)	8.1% (0.1%)	7.9% (0.2%)	7.4% (1.0%)
RACIAL AND ETHNIC IDENTITY				
HISPANIC	5.8% (0.2%)	7.0% (0.1%)	5.8% (0.2%)	5.6% (0.8%)
NON-HISPANIC, AMERICAN INDIAN/ALASKAN NATIVE	0.5% (0.0%)	0.6% (0.0%)	0.5% (0.0%)	0.5% (0.3%)
NON-HISPANIC, ASIAN	12.4% (0.2%)	13.9% (0.1%)	12.4% (0.2%)	20.6% (1.7%)
NON-HISPANIC, BLACK OR AFRICAN AMERICAN	6.8% (0.2%)	7.0% (0.1%)	6.8% (0.2%)	5.8% (0.9%)
NON-HISPANIC, MORE THAN ONE RACE	3.0% (0.1%)	3.3% (0.1%)	3.0% (0.1%)	2.5% (0.6%)
NON-HISPANIC, NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	0.5% (0.0%)	1.4% (0.0%)	0.5% (0.0%)	0.6% (0.3%)
NON-HISPANIC, OTHER	1.7% (0.1%)	1.7% (0.0%)	1.7% (0.1%)	2.4% (0.7%)
NON-HISPANIC, UNKNOWN OR NOT REPORTED	3.3% (0.1%)	3.1% (0.1%)	3.3% (0.1%)	4.3% (0.9%)
NON-HISPANIC, WHITE	66.0% (0.3%)	61.9% (0.2%)	66.0% (0.3%)	57.8% (2.1%)
CENSUS TRACT RACE / ETHNICITY				
TRACT % WHITE	64.8% (0.1%)	64.3% (0.1%)	64.8% (0.1%)	54.1% (0.4%)
TRACT % BLACK	7.9% (0.1%)	5.6% (0.0%)	7.9% (0.1%)	5.8% (0.2%)
TRACT % AMERICAN INDIAN/ALASKAN NATIVE	0.5% (0.0%)	0.8% (0.0%)	0.5% (0.0%)	0.2% (0.0%)
TRACT % ASIAN	14.2% (0.1%)	12.4% (0.0%)	14.2% (0.1%)	29.2% (0.6%)
TRACT % NATIVE HAWAIIAN/OTHER PACIFIC IASLANDER	0.5% (0.0%)	1.0% (0.0%)	0.5% (0.0%)	0.3% (0.0%)
TRACT % OTHER RACIAL GROUP	0.2% (0.0%)	0.2% (0.0%)	0.2% (0.0%)	0.1% (0.0%)
TRACT % HISPANIC	6.7% (0.0%)	10.4% (0.0%)	6.7% (0.0%)	5.8% (0.1%)

CENSUS TRACT CHARACTERISTICS				
TRACT DENSITY	11.2 (0.1)	4.0 (0.0)	11.2 (0.1)	12.9 (0.3)
TRACT % TWO OR MORE RACIAL GROUPS	5.2% (0.0%)	5.3% (0.0%)	5.2% (0.0%)	4.5% (0.0%)
TRACT % LIVING IN POVERTY	12.9% (0.1%)	11.2% (0.0%)	12.9% (0.1%)	12.1% (0.2%)
TRACT % OF PEOPLE WHO MOVED IN LAST YEAR	21.2% (0.1%)	16.7% (0.0%)	21.2% (0.1%)	30.6% (0.4%)
TRACT % OF PEOPLE WITH COLLEGE OR MORE	57.8% (0.1%)	33.6% (0.1%)	57.8% (0.1%)	62.8% (0.6%)
INSURANCE TYPE				
COMMERCIAL INSURANCE	80.2% (0.3%)	84.1% (0.1%)	80.2% (0.3%)	74.2% (2.0%)
MEDICARE	1.2% (0.1%)	1.9% (0.1%)	1.2% (0.1%)	0.9% (0.3%)
MEDICAID	2.9% (0.1%)	3.6% (0.1%)	2.9% (0.1%)	4.4% (0.9%)
OTHER INSURANCE	38.1% (0.3%)	36.3 (0.2%)	38.1% (0.3%)	49.7% (2.1%)

FSATT stands for 'Fine Stratification Average Treatment Effect Among the Treated'

SUPPLEMENTAL TABLE 4. FIRST DIFFERENCE MODELS OF ANNUALIZED BMI CHANGE FROM BEFORE TO AFTER THE TAX IMPLEMENTATION FOR SEATTLE VERSUS THE COMPARISON AREA (N=98,787)

STRATA		CHANGE IN COMPARISON AREA		DIFFERENCE-IN-DIFFERENCE ESTIMATE (DIFFERENCE IN CHANGE IN SEATTLE COMPARED TO THE COMPARISON AREA)	
		FSATE WEIGHTED			
ALL	**ALL**	0.166* (0.148, 0.185)	<0.001	-0.031* (-0.063, -0.0002)	0.049
SEX	**FEMALE**	0.204* (0.181, 0.227)	<0.001	-0.043* (-0.085, -0.002)	0.041
	MALE	-0.095* (-0.131, -0.059)	<0.001	0.031 (-0.032, 0.093)	0.338
RACE / ETHNICITY	**HISPANIC**	0.181* (0.123, 0.24)	<0.001	0.013 (-0.105, 0.132)	0.824
	NON-HISPANIC, AMERICAN INDIAN/ALASKAN NATIVE	0.076 (-0.172, 0.324)	0.549	-0.269 (-0.623, 0.085)	0.136
	NON-HISPANIC, ASIAN	-0.054 (-0.132, 0.024)	0.172	-0.005 (-0.141, 0.132)	0.947
	NON-HISPANIC, BLACK OR AFRICAN AMERICAN	-0.002 (-0.107, 0.104)	0.975	-0.083 (-0.253, 0.087)	0.338
	NON-HISPANIC, MORE THAN ONE RACE	-0.006 (-0.106, 0.093)	0.901	0.002 (-0.225, 0.228)	0.989
	NON-HISPANIC, NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER	0.013 (-0.099, 0.125)	0.820	-0.064 (-0.663, 0.535)	0.834
	NON-HISPANIC, OTHER	-0.084 (-0.218, 0.05)	0.220	-0.03 (-0.242, 0.181)	0.778

	NON-HISPANIC, UNKNOWN OR NOT REPORTED	0.04 (-0.066, 0.146)	0.456	-0.015 (-0.199, 0.169)	0.875
	NON-HISPANIC, WHITE	-0.011 (-0.074, 0.051)	0.721	-0.059 (-0.184, 0.066)	0.357
AGE GROUP	**18-29**	0.418* (0.355, 0.481)	<0.001	-0.057 (-0.183, 0.069)	0.377
	30-49	-0.193* (-0.265, -0.122)	<0.001	0.005 (-0.132, 0.142)	0.941
	50-66	-0.385* (-0.45, -0.319)	<0.001	0.057 (-0.074, 0.188)	0.394
POVERTY	**LOW POVERTY**	0.165* (0.145, 0.184)	<0.001	-0.036* (-0.068, -0.005)	0.024
	HIGH POVERTY	0.028 (-0.015, 0.071)	0.200	0.027 (-0.086, 0.14)	0.638
EDUCATION	**HIGH EDUCATION**	0.162* (0.14, 0.184)	<0.001	-0.045* (-0.08, -0.01)	0.012
	LOW EDUCATION	0.015 (-0.022, 0.052)	0.426	0.028 (-0.04, 0.095)	0.420

*Notes: * indicates statistical significance at 0.05 level; + indicates statistical significance at 0.10 level*

Model 3: Comparative interrupted time-series details
Specifically, we fit this model:

$$
\begin{aligned}
Y_{it} = [\beta_0] + [\beta_1 Seattle_i] &+ \beta_2 Time\_since_{it} + \beta_3 Posttax_t \\
&+ \beta_4 Seattle_i * time\_since_{it} + \beta_5 time\_since_{it} * Posttax_t \\
&+ + \beta_6 Seattle_i * time\_since_{it} * Posttax_t + \theta \boldsymbol{X}_{it} + \theta \boldsymbol{C}_{ict} + t + \varepsilon_{it}
\end{aligned}
$$

Where Y_{it} is BMI for individual *i* at time *t, Seattle* is an indicator variable that equal 1 if patient lives in Seattle and 0 if they live outside of Seattle. *Time_since* is our main time variable and it is the time that has elapsed since the study start date, operationalized by subtracting the study start date from the visit date.

Posttax is an indicator variable equal to 1 if a visit occurred after the date of tax implementation (Jan 1 2018) and 0 for before tax implementation. t are for person-level fixed effects, which are included to compare people to themselves over time, thus providing a within-person estimate of effects and controlling for all observed and unobserved time-fixed confounders (i.e. genetics, gender, time-fixed preferences), **X** is a vector of time-varying covariates, in this case, health insurance status was the only time-varying measured confounder we identified. ***C*** is a vector of time-varying community level variables that we add in sensitivity analyses since, except in the cases where people move, they change slowly over time and their baseline values are comparative for in the propensity score matching and through the inclusion of person-level fixed effects.

Comparative interrupted time-series results
The estimates indicate that on average BMI increased in the comparison area over the pre-tax period at a rate of 0.13 kg/m2 per year (Row 1, **Supplemental Table 5**). Additionally, the estimate for the change in the rate of BMI gain from pre- to post-tax in the comparison area (Row 2) is positive but not statistically significant, indicating that for the comparison area, the rate of BMI gain remained stable post-tax ($FSATE_{PostTax-Time}$: 0.009 (95% CI: 0.023, -0.005); p-value 0.225)). The estimate for the difference in time trends in BMI gains in the pre-tax period for Seattle versus the comparison area is negative, but not statistically significant (Row 3; $FSATE_{Seattle-Time}$: -0.018 (95% CI: 0.03, -0.065); p-value 0.461). This indicates that in the pre-tax period, in weighted models, the rates of BMI gain were not different for patients in Seattle compared to the comparison area. The difference-in-difference estimate for the effect of the tax, or the difference in the change in the time trend from pre- to post-tax for Seattle versus the comparison area (Row 4; $FSATE_{Seattle-PostTax-Time}$: -0.001 (95% CI: 0.025, -0.027); p-value 0.929) is not statistically significant, indicating no difference in the change in BMI slope from pre to post-tax for Seattle as compared to the comparison area.

Due to the highly unbalanced nature of the data, we re-ran these models using a dataset with BMI linearly interpolated to create a quarterly BMI measurement between every two measurements. This linear interpolation allows for a balanced panel with measurements taken at the same time for all individuals. For measurements before the first measurement and after the last measurement, rather than linearly extrapolate to fill in missing data, we pulled the first observation back and the last forward. This resulted in every patient having 6 BMI measurements In these models, the estimate for the tax effect is of larger magnitude and the confidence interval substantially smaller (Interpolated model $_{\text{Seattle-PostTax-Time}}$: -0.005 (95% CI: -0.015, 0.005); p-value 0.31). The fact that the interpolated results are substantively different (5 times larger magnitude) suggests that perhaps the unbalanced nature of the health records data is influential in the results. That is, it may be that patients who have more ambulatory care visits (and thus contribute more observations to the data) are less affected by the tax than those with fewer visits. This is consistent with an interpretation that older patients, who on average seek clinical care more frequently, are less affected by the tax, whereas younger people and those living in neighborhoods with less poverty and with a higher percentage of people with a college degree seek care less frequently. Indeed, these are the populations whose BMI seems to be more affected by the tax according to the stratified results presented in **Supplemental Table 4**.

SUPPLEMENTAL TABLE 5. ESTIMATED ASSOCIATION OF SEATTLE'S SBT WITH CHANGE IN BMI USING AN COMPARATIVE INTERRUPTIVE TIME SERIES MODEL, FSATE WEIGHTED AND INTERPOLATED RESULTS

		FSATE		INTERPOLATED FSATE	
		COEFFICIENT (95% CI) | P-VALUE			
ROW 1	TIME TREND IN BMI GAINS IN THE COMPARISON AREA PRE-TAX	0.132* (0.153, 0.112)	<0.001	0.118* (0.105, 0.131)	<0.001
ROW 2	CHANGE IN TIME TREND IN BMI FROM BEFORE TO AFTER THE TAX IN THE COMPARISON AREA	0.009 (0.023, -0.005)	0.225	-0.002 (-0.006, 0.003)	0.512
	SEATTLE-TIME				
ROW 3	DIFFERENCE IN TIME TREND IN BMI BETWEEN THE COMPARISON AREA AND SEATTLE, PRE-TAX	-0.018 (0.03, -0.065)	0.461	-0.01 (-0.037, 0.016)	0.451
	SEATTLE-POSTTAX-TIME				
ROW 4	DIFFERENCE IN THE CHANGE IN THE TIME TREND IN BMI FOR SEATTLE COMPARED TO THE CHANGE IN THE TIME TREND IN THE COMPARISON AREA FROM BEFORE TO AFTER THE TAX (THE DID ESTIMATE)	-0.001 (0.025, -0.027)	0.929	-0.005 (-0.015, 0.005)	0.31

*Notes: * indicates statistical significance at 0.05 level*

We performed several sensitivity analyses using the first difference model. These sensitivity analyses are displayed below in **Supplemental Table 6**. First, we removed people with very high statistical weights from the model to see if these observations were driving the estimates of the impact of the tax or the wide confidence intervals around these estimates. The results were substantively similar after this removal. We also removed people with a change in BMI in one year of more than 5 BMI units. While such a change is biologically plausible, it could also arise due to data errors (e.g., mis-recorded weight or height in the medical record), and we wanted to ensure these large changes were not driving the results. Again, the results were substantially similar after the removal. We tested whether results would change if we additionally included time-varying neighborhood-level census tract characteristics [in the regression model or in the FSATE weight?]. In these models, the estimate of effect (the coefficient in the table below) is unchanged but the confidence intervals become wider and therefore the results are no longer statistically significant. The combination of an unchanged point estimate (coefficient) and a wider confidence interval suggests that adjusting for time-varying characteristics decreases precision without reducing bias. Finally, we fitted this model with the interpolated data set. In this case the point estimate for the estimated impact of the tax was substantially smaller and no longer statistically significant. This could be explained by carrying observations forward and backwards in time rather than extrapolating beyond observed BMI values, which could lead to an attenuated effect. We also repeated the analyses, but removed census tracts that had any area outside of the Seattle boundary. Additionally, while not sensitivity analyses per se, when instead of using the FSATE weights, we use the FSATT weights, there is no evidence of an impact of the tax on BMI as indicated by the confidence interval straddling zero and the large p-value.

SUPPLEMENTAL TABLE 6. SENSITIVITY ANALYSES FOR THE DIFFERENCE-IN-DIFFERENCE ESTIMATES FOR FIRST DIFFERENCE MODEL

VARIABLE	DIFFERENCE-IN-DIFFERENCE ESTIMATE	
	COEFFICIENT (95% CI)	P-VALUE
PRIMARY MODEL (FSATE WEIGHTED)	-0.031* (-0.063, 0.0)	0.049
REMOVED PEOPLE WITH VERY HIGH STATISTICAL WEIGHTS	-0.030 (-0.056, -0.004)	0.025
INTERPOLATED BMI	-0.014 (-0.044, 0.015)	0.349
ADD TIME VARYING TRACT CHARACTERISTICS AS CONTROLS	-0.031^{+} (-0.068, 0.005)	0.094
EXCLUDE PEOPLE WITH >5BMI UNIT CHANGE IN A YEAR	-0.031 (-0.063, -0.0002)	0.049
REMOVE TRACTS THAT HAVE ANY AREA OUTSIDE OF SEATTLE BOUNDARY	-0.047 (-0.096, 0.002)	0.061
FSATT WEIGHTED	-0.011 (-0.079, 0.057)	0.752

** indicates statistical significance at 0.05 level; $^{+}$ indicates statistical significance at 0.10 level*

We performed additional sensitivity analyses on the comparative interrupted time-series model. The null results of this model were also unchanged when applying each of the same sensitivity analyses described above (**Supplemental Table 7**).

SUPPLEMENTAL TABLE 7. INTERUPTED TIME-SERIES TYPE MODEL SENSITIVITY ANALYSES							
VARIABLE	FSATE WEIGHTED NO TRACT	TRIMMED WEIGHTS	TRACT TIME VARYING VARS	NO BMI	REMOVE TRACTS THAT TOUCH BORDERS	EXCLUDE MOVERS	UNWEIGHTED NO TRACT
	COEFFICIENT (95% CI); P-VALUE						
TIME (TIME TREND IN BMI IN THE COMPARISON AREA PRE-TAX)	0.132* (0.153,0.112) <0.001	0.134 (0.123, 0.145) <0.001	0.135 (0.156, 0.115) <0.001	0.130 (0.116, 0.144) <0.001	0.131 (0.153,0.11) <0.001	0.103 (0.115,0.09) <0.001	0.131* (0.141,0.121) <0.001
POSTTAX-TIME (CHANGE IN TIME TREND IN BMI FROM BEFORE TO AFTER THE TAX IN THE COMPARISON AREA)	0.009 (0.023, -0.005) 0.225	0.004 (-0.003, 0.011) 0.776	0.008 (0.022, -0.007) 0.291	0.008 (-0.001, 0.016) 0.273	0.01 (0.025,-0.005) 0.192	0.003 (0.011,-0.006) 0.564	0.004 (0.009, -0.002) 0.172
SEATTLE-TIME (DIFFERENCE IN TIME TREND IN BMI BETWEEN THE COMPARISON AREA AND SEATTLE, PRE-TAX)	-0.018 (0.03, -0.065) 0.461	-0.020 (-0.062, 0.022) 0.004	-0.022 (0.026, -0.07) 0.372	-0.016 (-0.059, 0.026) 0.699	-0.017 (0.038,-0.073) 0.540	-0.026 (0.022,-0.075) 0.28713	-0.025* (-0.007, -0.043) 0.007
SEATTLE-POSTTAX-TIME (DIFFERENCE IN THE CHANGE IN THE TIME TREND IN BMI FOR SEATTLE COMPARED TO THE CHANGE IN THE TIME TREND IN THE COMPARISON AREA FROM BEFORE TO AFTER THE TAX (THE DID ESTIMATE))	-0.001 (0.025, -0.027) 0.929	0.004 (-0.018, 0.026) 0.173	0.0 (0.026, -0.026) 0.993	0.001 (-0.021, 0.023) 0.617	-0.004 (0.029,-0.036) 0.820	0.02 (0.045,-0.005) 0.121	0.008 (0.02, -0.003) 0.144

**indicates statistical significance at 0.05 level*
FSATT stands for 'Fine Stratification Average Treatment Effect Among the Treated'