Measuring the Impact of Calorie Labeling: The Mechanisms Behind Changes in Obesity*

Rodrigo Aranda Balcazar
Georgia State University

Michael E. Darden
Johns Hopkins University and NBER

Donald Rose
Tulane University

February, 2020

Abstract

Learning the true calorie content of fast food may induce consumers to change behavior, yet recent evidence is mixed on whether calorie labels cause consumers to order healthier meals. Especially for individuals for whom consumption of highly caloric fast-food is habitual, a rational response to calorie labeling may instead be to maintain consumption levels but increase physical activity. Using American Time Use Survey data from 2004 to 2012, we show that the 2008 New York City Calorie Labeling Mandate significantly improved several measures of physical activity, including overall metabolic equivalents of task units and minutes of sedentary activity. Our results translate to an average extra 28 calories burned per day or a 0.6kg weight decrease for the average person over one year. These results provide a plausible mechanism for calorie labeling mandates to lower obesity rates, which we demonstrate in the New York City context.

JEL Classification: D82; D83; I12; I18;

Keywords: Calorie Labeling; Information Asymmetry; Physical Activity

*Rodrigo Aranda Balcazar: rarandabalcazar@gsu.edu; Michael Darden: mdarden4@jhu.edu; Donald Rose: diego@tulane.edu. We thank Alan Barreca, Ali Enami, Jacqueline Fiore, Patrick Button, seminar participants at Iowa State University, Tulane University and 2018 American Society of Health Economists Conference.

Measuring the Impact of Calorie Labeling: The Mechanisms Behind Changes in Obesity

February, 2020

Abstract

Learning the true calorie content of fast food may induce consumers to change behavior, yet recent evidence is mixed on whether calorie labels cause consumers to order healthier meals. Especially for individuals for whom consumption of highly caloric fast-food is habitual, a rational response to calorie labeling may instead be to maintain consumption levels but increase physical activity. Using American Time Use Survey data from 2004 to 2012, we show that the 2008 New York City Calorie Labeling Mandate significantly improved several measures of physical activity, including overall metabolic equivalents of task units and minutes of sedentary activity. Our results translate to an average extra 28 calories burned per day or a 0.6kg weight decrease for the average person over one year. These results provide a plausible mechanism for calorie labeling mandates to lower obesity rates, which we demonstrate in the New York City context.

JEL Classification: D82; D83; I12; I18;

Keywords: Calorie Labeling; Information Asymmetry; Physical Activity

1 Introduction

The extent of the obesity epidemic in the United States is well known: between 2013 and 2014, 37.9% of adults and 17.2% of children were obese, and important differences in obesity rates exist by race, education, and income (Ogden *et al.*, 2016; Flegal *et al.*, 2016). The rise in obesity is not surprising given recent evidence that the persistent average daily calorie imbalance (intake vs. expenditure) required to explain the rise in weight is only 220 calories (Wang *et al.*, 2016; Hall *et al.*, 2011; Cawley, 2015). To put this imbalance in context, Burton *et al.* (2006) and Burton *et al.* (2014) show that fast food consumers (and nutrition professionals!) systematically underestimate the calorie content of fast food by an average of more than 600 calories.[1] Naturally, a focus of policy has been to inform individuals of the calorie content of restaurant food. For example, the Affordable Care Act (ACA) requires the disclosure of the nutritional content of fast food and vending machine menu items. Legal challenges delayed the implementation of nationwide calorie posting until May of 2018.[2] If calorie information causes consumers to update their expectations regarding the calorie content of food, then, consistent with the economic notion that body mass is produced via endogenous inputs that affect the consumption and expenditure of calories (Cawley, 2004; Lakdawalla *et al.*, 2005), rational consumers may adjust their investments in body mass, especially when the marginal disutility of weight gain is increasing in the distance from an individual's "ideal" weight (Lakdawalla *et al.*, 2005).

Yet recent evidence on the extent to which calorie labeling at restaurants causes consumers to order fewer calories is mixed at best. For example, while Cawley (2018) report a 3% reduction in calories ordered from a randomized calorie labeling field experience, studies of recent calorie labeling mandates in Philadelphia (Elbel *et al.*, 2013) and New York City (Dumanovsky *et al.*, 2010; Elbel *et al.*, 2009; Cantor *et al.*, 2015) have found no effect of purchased calories. Bleich *et al.* (2017), who reviewed 53 studies on the role of calorie labeling on both changes in consumer purchases as well as restaurant offerings, conclude that there is little evidence that such policies change behavior. The lack of a robust finding with respect to calorie consumption is particularly surprising given recent evidence from Restrepo (2017), who finds that county-level calorie labeling mandates in New York State reduced obesity rates by 12%.[3]

In this paper, we study an alternative response to calorie labels: physical activity. Physical activity (i.e., calorie expenditure) may be a means of adjustment for individuals with strong pref-

[1] Block *et al.* (2013) show that the degree of underestimation grows with the size of the meal. Powell *et al.* (2012) show that, in 2007 and 2008, fast food and food away from home accounted for 24% of adolescent and adult energy intake.

[2] See http://kff.org/interactive/implementation-timeline/

[3] Deb and Vargas (2016) find that the strongest effects from the New York State mandates were with respect to men.

erences for highly caloric fast food, perhaps because of habitual consumption or because of cultural norms. Learning of the nutritional content of fast food may induce more physical activity to the extent that healthier fast food is significantly more expensive, suggesting potentially important heterogeneity in the response to calorie labeling. Our point is that the production of body mass is a function of both the input and output of calories, and learning about the calorie content of food may induce behavioral responses on either dimension. Furthermore, the margin of adjustment matters with respect to policy design and evaluation. For example, if calorie labeling induces more physical activity, future public health efforts may be more effective by emphasizing the connection between physical activity and healthy body weight, as opposed to, for example, promoting approaches that make high-calorie items less prominent on menus.

To investigate empirically, we consider the 2008 New York City Calorie Labeling Mandate (NYC-CLM), which required food establishments that belong to a group of 15 or more establishments to post calorie information for standardized menu items. We use American Time Use Survey (ATUS) data from New York City and other metropolitan areas from 2004 to 2012 to estimate a series of physical activity entropy-weighted difference-in-differences models (Hainmueller, 2012). We compare trends in various physical activity measures in New York City to those in other metropolitan areas across the United States before and after the 2008 mandate.

We find a statistically significant 1 percent increase in metabolic equivalents of task, an effect that was primarily a result of a decrease in sedentary activity and an increase in light activities. Our results translate to an average extra 28.1 calories burned per day.[4] Our results are consistent across several alternative specifications and robustness checks. Furthermore, we demonstrate statistically similar pre-trends in conditional physical activity measures between New York City and both the weighted and unweighted aggregate of control metropolitan areas. Because we model changes in only one treatment group - residents of New York City - and because our research design includes a large number of control cities, we employ a cluster residual bootstrap with the heteroskedasticity correction proposed by Ferman and Pinto (2019) to conduct inference, and we show that the statistical significance of our findings is preserved under this adjustment.

Our physical activity results put into context evidence that the NYC-CLM caused a reduction in obesity, a finding which we revisit by studying the conditional obesity trends of adults in New York City relative to other urban areas in the Selected Metropolitan/Micropolitan Area Risk Trends of the Behavioral Risk Factor Surveillance System (SMART-BRFSS) from 2004 to 2012. One strength of our design is that our relatively large sample sizes are able to detect small effects precisely. We find that the labeling mandate reduced the proportion of obese adults by two

[4] The estimated 95% confidence interval is between $11.6 - 44.6$ calories.

percentage points, a roughly 9.5% decline. This result is confined almost entirely to transitions out of class I obesity, defined as BMI between 30 and 35. We find economically and statistically insignificant effects at higher levels of body mass index. Our main entropy-weighted difference-in-differences estimator passes tests for parallel pre-treatment trends, and our results are robust to a variety of alternative specifications and placebo tests. The result is also robust to the use of sample weights, entropy weights and a synthetic control method, which constructs counterfactual obesity and body mass index trends in the absence of NYC-CLM.[5][6] Additionally, alternative specifications that allow for commuters, alternative definitions of control and treatment groups, and controls for the 2007 trans-fat ban in New York City, all suggest that our results are robust. Finally, placebo analysis treating, for example, cigarette smoking as our dependent variable yields no economically or statistically significant effect.

Our results provide some clarity to a mixed literature on the effects of calorie labeling, and we demonstrate that physical activity is a potentially important mechanism for calorie labeling to translate into body mass reduction. Indeed, the total calorie deficit required to generate our obesity result is about 43 calories per day over a year on average. The increased physical activity we report accounts for 28 calories per day, or about two thirds of this deficit. Moreover, as a policy matter, there is evidence that physical exercise can become habitual following relatively short-term incentive schemes (Charness and Gneezy, 2009).[7] For completeness, we also return to the question of whether the NYC-CLM changed trends in fast food consumption. While our SMART-BRFSS data do not include measures of fast food consumption, we estimate models of fast food frequency and expenditure in the Consumer Expenditure Survey. We show a slight decrease in frequency, but we reject the null hypothesis that pre-trends are parallel. Conditional upon visiting a fast food restaurant, the CEX data suggest a statistically significant, but again very small in magnitude, increase in expenditure per visit. This increase may be consistent with substitution towards healthier, more expensive items. Overall, our results on fast food frequency and expenditure are consistent with several studies of the NYC-CLM that have found a negligible effect on purchasing behavior conditional on visiting a fast food restaurant (Dumanovsky *et al.*, 2010; Elbel *et al.*, 2009; Cantor *et al.*, 2015).

Our paper proceeds as follows. Section 2 provides background on the NYC-CLM and calorie labeling initiatives more generally, and it develops a brief conceptual framework for the production

[5] See Table 3, Figure 2 and Figure 3

[6] Our result is roughly consistent with Restrepo (2017), who finds that the NYC-CLM reduced BMI by 1.5% and reduced obesity by 12%. That paper uses variation in the timing of calorie label mandates over several New York State counties, relative to unaffected counties, while our estimator focuses on variation New York City relative to trends in the entropy-weighted average of similar cities.

[7] For reference, a tablespoon of ketchup has approximately 20 calories.

of body mass that is consistent the theoretical literature (Grossman, 1972; Cawley, 2004). Section 3 demonstrates summary statistics; documents our empirical methods; presents our results; and displays the robustness of our findings. Finally, Section 4 discusses the implications of our finds and concludes.

2 Background

2.1 Conceptual Framework

Conceptually, Cawley (2004) provides a positive economic framework to understand trends in body mass and obesity. The model decomposes a 24-hour day into five categories of time: S for hours of sleep, L for hours of leisure, O for hours of paid work, T for hours of transportation, and H for hours of unpaid home production. For time spent in each activity, individuals gain utility (or disutility) both directly and indirectly, as time spent on leisure, for example, also influences body weight, which enters the utility function. The key equation is that which dictates how body weight evolves:

$$\Delta W = c(F) - f(S, L, O, T, H) - \delta(G)W. \tag{1}$$

Here, the change in weight is a function of food consumption F, which is transformed into body weight by the biological function $c(.)$; the function $f(.)$ represents the biological transformation of time spent in each activity category into body weight; and $\delta(G)W$ represents the baseline metabolic rate of calorie expenditure. Cawley (2004) argues that the absence of perfect information on the part of individuals about the functional transformations and the inherent levels in arguments in Equation 1 is justification for policy that better informs individuals.

Among other things, the framework is useful for understanding how individuals will react to the exogenous release of calorie information. Assuming that calorie labels at fast food restaurants reveal higher than expected calorie counts, the extent to which a rational individual will re-optimize depends on several factors. First, the extent to which her current weight deviates from her ideal weight, as well as the functional relationship between this deviation and utility - if the marginal disutility of weight deviating from ideal weight is increasing in the deviation, for example, then individuals further from their ideal weight would be expected to make a more dramatic reallocation, all else equal. Second, the utility cost of reallocating towards either lower food consumption or higher activity levels may retard reallocation (i.e., individuals with strong preferences for food or sedentary behavior will reallocate less). Finally, if the price of healthier food is significantly higher, we would expect less substitution away from unhealthy food, especially for those with low-income.

2.2 Evidence on Calorie Labeling

A sizable literature on the effect of calorie labeling on both obesity rates and behavior has found mixed results. Table 1 summarizes this literature, providing for each paper the relevant population, a summary of the results, and the data used. In an important study, Bollinger and Sorensen (2011) estimate a difference-in-differences estimator of individuals in New York City relative to Boston and Philadelphia with internal data from Starbucks on over 100 million transactions. Those authors found that calorie labels reduced calories purchased by 5.8% per transaction, which was mainly driven by their choice of supplemental or side items.[8] On the other hand, Finkelstein *et al.* (2011) studies the effect of the 2009 King County, Washington fast food calorie labeling law on fast food frequency and expenditure using a difference-in-differences estimator. Focusing on one specific fast food chain, the authors found no effect of the calorie labeling mandate on calorie content of food ordered. A fair assessment of the recent literature is that calorie labeling can change calorie purchasing behavior, but the context of the labeling is important, and there is little evidence of persistent differences in behavior.

Table 1

Panel B of Table 1 summarizes the brief literature on the effect of calorie labels on obesity rates and body mass index. While Deb and Vargas (2016) find no effect overall, they find significant reductions in obesity for men who were exposed to a variety of different calorie labeling laws at the county level. Alternatively, Restrepo (2017) finds a large reduction in overall obesity in counties in New York state, including New York City, that implemented calorie labeling laws.

2.3 New York City Calorie Labeling Mandate

In 2006, New York City, though its Department of Health and Mental Hygiene Board of Health (DHMH), amended Article 81 of the New York City Health Code to require food service establishments to disclose calorie information for standardized menu items on menu boards and menus next to each menu item. The requirement was implemented for food items that are standardized with regard to portion size, formulation, and ingredients.[9] The mandate was required for fast food chains that belong to a group of 15 or more establishments and that operate under common ownership or are individually franchised, whether it is locally or nationally, or they do business under the same name.[10] While the calorie labeling mandate became effective on March 31, 2008,

[8] Longitudinal analysis on a subset of loyalty customers showed that calories purchased for those usually buying more than 250 calories dropped by 26%.

[9] Notice of adoption of an amendment (81.50) to article 81 of the New York City Health Code.

[10] Approximately 10% of restaurants in New York City met this requirement.

implementation was delayed until May 5th, 2008 due to litigation, and full enforcement did not occur until July 18th, 2008.

During the final three months of 2008, New York City's Health Department launched a Calorie Education Campaign called "Read 'em before you eat 'em", which consisted of five different advertisements appearing in over 100 New York City subway trains. The campaign was designed both to help individuals realize how quickly calories consumed in fast-food establishments can accumulate and to inform individuals of the average calorie recommendation for an adult.[11] To the extent that the advertising campaign registered with subway riders, it potentially reached a wider audience than just those that frequent fast-food restaurants. By the end of 2008, compliance with the calorie-posting mandate was 85 percent[12].

Evidence on the NYC-CLM suggests that a.) the ordering behavior of consumers at fast food restaurants did not significantly change, and b.) body mass index and obesity fell significantly. For example, Cantor *et al.* (2015) studies survey evidence and cash register receipts from four large restaurant chains before and after the 2008 mandate. Using a difference-in-differences design relative to nearby Newark, New Jersey, they show no systematic change in calories ordered or the nutrition content of food ordered, despite the fact that treated consumers reported seeing and using the information more both in 2008 and in a later follow-up in 2013 and 2014. Elbel *et al.* (2009) find similar results using a similar research design in low-income, minority communities, and Elbel *et al.* (2011) find similar results for adolescents. Dumanovsky *et al.* (2010) find significant heterogeneity in the NYC-CLM, with significant reductions in calories purchased at McDonalds, KFC, and Au Bon Pain, but no effect overall. Finally, Vadiveloo *et al.* (2011) note that, among the group of consumers who noticed the labels, visits to fast food restaurants fell significantly.

Seemingly inconsistent with the large literature that finds no effect on calories purchased, Restrepo (2017) analysed the impact of local mandatory calorie labeling laws implemented in several New York counties, including New York City, on body weight. Using the timing of calorie label mandates across counties (but within New York State), he finds that calorie information reduced the rate of obesity by 12%, an effect which was concentrated in lower income minority consumers. Using data from the Behavioral Risk Factor Surveillance Survey between 2004-2012, Restrepo (2017) does not examine the ordering behavior of consumers at affected restaurants, but he finds negligible effects with respect to smoking, alcohol consumption, and the consumption of healthy alternative foods. Importantly, Restrepo (2017) uses self-reported participation in physical activities during the past month and finds no effect on the frequency or intensity of exercise. Restrepo (2017) concludes that "other margins of adjustment drive the body-weight impacts estimated here."

[11] Press Release # 066-08 New York City Health Department.

[12] New York City Department of Health and Mental Hygiene, Agency Biennial Report 2007-2008

Our contribution to this literature is two-fold. First, we revisit the obesity results that are specific to New York City by studying trends in obesity in New York relative to other large cities, where we largely confirm the results from Restrepo (2017). Second, in contrast to Restrepo (2017), we find a biologically significant increase in physical activity that explains the reduction in obesity in the absence of any significant change in fast food behavior.

3 Empirical Evidence

To study the impact of the NYC-CLM on physical activity, we begin by re-evaluating the extent to which the mandate lowered obesity rates using the Selected Metropolitan/Micropolitan Area Risk Trends subset of the Behavioral Risk Factor Surveillance Survey (SMART-BRFSS). SMART-BRFSS uses cross-sectional telephone survey questions in the BRFSS in select cities to gauge the prevalence of several health behaviors, including cigarette smoking and alcohol abuse, and the basic health status of resident individuals. Included in the survey are self-reported measures of height and weight, from which we are able to construct body mass index (BMI).[13] These data are ideal for our setting because we are able to identify individual respondents in New York City, as well as respondents in counties surrounding New York City and other metropolitan and micropolitan statistical areas (MMSAs). The data are representative of the areas in which the surveys are administered.

Using repeated cross-sections of the SMART-BRFSS from 2004 to 2012, our data include 1,739,240 observations from the continental United States. To arrive at our estimation sample, we drop pregnant women (13,438 observations deleted). Next, we restrict our sample to those individuals with a calculated body mass index of more than 10 or less than 60 (80,953 observations deleted). To avoid contaminating our control group with individuals who may plausibly be affected by the NYC-CLM, we exclude individuals who live in counties near New York City. These potential commuters are individuals who work, study, or have other relevant activities in New York City on a regular basis, and who could also be affected by the policy change. We identify and drop those counties in which more than nine thousand individuals work in New York City (73,515 observations deleted).[14] We drop individuals from other counties or states that enacted calorie labeling mandates during this period of time (297,629 observations deleted). Finally, we drop

[13] See Cawley (2015) for an overarching critique of using self-reported height and weight information to study obesity. In the context of our study, misreporting will only cause a bias (beyond traditional measurement bias) if the NYC-CLM caused an individual to report her height/weight differently.

[14] As measured by the Census Bureau for the Residence County to Workplace County Commuting Flows from 2009 to 2013. Our results are not sensitive to this definition. We conduct sensitivity analysis in which we keep respondents living in counties adjacent to New York City, separately treating these potential commuters in our treatment and control groups. Our results are not sensitive to the way in which we classify commuters. See below.

individuals with missing values on height or weight, county, time of the survey, age, gender, race, education, no physical activity, and income (223,666 person/year observations deleted).[15] Our final sample consists of 1,104,495 observations.

Within our final sample, we are able to identify individuals at the county/month/year level. Because the month of interview is recorded in our data, we define the "post" NYC-CLM treatment period to be all observations after July 31st, 2008. Treatment individuals are those living in any New York City county, while our control group consists of all respondents living in metropolitan/micropolitan statistical areas in the 370 other counties surveyed by BRFSS that do not have an existing calorie labeling mandate. Table 2 presents summary statistics of key variables in New York City and control cities before and after the NYC-CLM.[16] To control for unobserved, supply-side heterogeneity, we merged to our SMART-BRFSS data, information on number of fast food restaurants, full-service restaurants, and fitness and recreation centers per 1,000 residents from the U.S. Census Bureau County Business Patterns.[17] Table 2 shows the importance of examining trends in outcomes rather than comparing levels: prior to the NYC-CLM policy, New York City has statistically larger fractions of African-Americans, Hispanics, females, and unemployed.

TABLE 2

To identify the extent to which the NYC-CLM changed BMI and Obesity rates, we compare the conditional trends in weight outcomes in New York City to those in other metropolitan areas which are concentrated in 370 counties across the United States.[18] Our main estimating equation is given as:

$$y_{ict} = \delta_o + \gamma d_t + \phi NYC_c + \delta NYC_c * d_t + \beta' X_{ict} + \epsilon_{ict}. \tag{2}$$

Where y_{ict} is the weight outcome of individual i living in county c in month/year t, d_t is a binary variable for observations recorded after July 18th, 2008, NYC_c is a binary variable for New York City residents, and X_{ict} is a vector of exogenous control variables listed in Table 2. Our empirical model also includes county and month/year fixed effects, and we control for supply-side fast-food characteristics such as the number of fast food restaurants, full-service restaurants, and fitness and recreation centers per 1,000 residents in a county. To measure the impact on weight outcomes,

[15] Our results are not sensitive to including these individuals in our sample and estimating our preferred model with missing dummy variables.

[16] All descriptive statistics are weighted using SMART-BRFSS, ATUS and CES weights to see comparability before and after the policy for treated and control units.

[17] Currie *et al.* (2010) show that the density of fast-food restaurants may cause weight gain in young adults.

[18] Tables 7 and 8 of the Appendix list each of these areas and the frequency at which we observe them.

we separately estimate equation 2 for both the logarithm of body mass index and a dichotomous variable that states if individual i is obese. The key identifying assumption is that the conditional time trends in our outcome variables would be the same in New York City as those trends in other cities in the absence of the NYC-CLM; that is, the mandate introduced a deviation from the New York City trend. Thus, our parameter of interest is δ, the difference-in-differences parameter on the interaction term between NYC_c and d_t. We estimate our difference-in-differences models using sample weights and entropy balanced weights separately. To create entropy balanced weights, we follow Deb and Vargas (2016) and use the weighting scheme developed by (Hainmueller, 2012), which uses an entropy reweighting method that calibrates unit weights between treatment and control units to satisfy conditions of covariate balance associated with sample moments. The main advantage of using these weights is that the new estimating weights perfectly balance the covariates and outcome pre-treatment trends. Furthermore, we show that our entropy-weighted results are similar to those from a difference-in-differences estimator using sample weights, which is suggestive of trend balance prior to treatment.

To address issues with inference when there are a limited number of treated units under a difference in differences approach, we estimate Ferman and Pinto (2019) p-values. Under the assumption that the distribution of the linear combination of errors is independent to treatment status, Ferman and Pinto (2019) correct for heteroskedasticity by re-scaling the pre-post difference in average residuals of the control groups so that they inform about the distribution of the pre-post difference in average errors of the treated group(s). This method applies a cluster residual bootstrap with the aforementioned heteroskedasticity correction. We use this approach because we have a large number of control groups and we get consistency when this number goes to infinity.[19]

Panel A of Table 3 presents estimates of δ for both the log of body mass index (BMI) and from a linear probability model for obesity using only sample weights. For each outcome, we present three estimates that come from models which increasingly control for various fixed effects. Consistent across columns 1-3 and 4-6, our results suggest a statistically significant and large decrease in BMI and obesity in New York City following the calorie labeling mandate. In our preferred specification in columns 3 and 6, we find a 1.4% reduction in BMI and a 2.3 percentage point reduction in obesity, respectively, in New York City relative to our control cities.

Table 3

[19] Methods to address inference estimations with limited treated units with difference in differences methods has been mainly addressed in the literature by Ferman and Pinto (2019), Conley and Taber (2011) and MacKinnon and Webb (2018). However, we use Ferman and Pinto (2019) to be able to correct for heteroskedasticity and to take into account potential complex structures on the errors.

While the fundamental identification assumption of any difference-in-differences strategy is untestable, parallel pre-trends in our outcome variables provide suggestive evidence that our assumption is reasonable. In Panel A of Table 3 the P-Value on statistically similar pre-trends is 0.681 and 0.765 for BMI and obesity, respectively, which we interpret as suggestive evidence that our results in Panel A may be significantly biased. To proceed, we estimate the same model using entropy balanced weights, which balance the pre-trend means between treatment and control units. To highlight the difference between using sample and entropy weights, Figure 1 demonstrates trends in log BMI and obesity in New York City relative to control cities under both schemes. Figures b. and d., which include entropy-weights, clearly demonstrate a reduction in pre-trend variation by treatment status. We also estimate a synthetic control method model following from Abadie *et al.* (2010), which uses sample information in the pre-period to construct a counterfactual New York City.[20] Figure 2 and Figure 3 show the pre-treatment balance between treatment and the selected synthetic control group for BMI and obesity, respectively.

FIGURE 1

FIGURE 2

FIGURE 3

Panels B and C of Table 3 present estimates from our entropy weighted and synthetic control method estimators. Results are consistent with those in Panel A: both BMI and obesity fell in New York City relative to control cities and counties. In our preferred set of controls in columns 3 and 6, we find between a 1.3% and 0.8% reduction in BMI for entropy-weights and synthetic control, respectively. The respective reductions in obesity are 2 and 1.96 percentage points. When addressing issues in inference due to having only one treated unit, the Ferman-Pinto P-Value corroborate that these estimates are statistically significant.

To explore the effect of NYC-CLM at different points in the BMI distribution, we estimate linear probability models for overweight (BMI$\in [25, 30)$), class one obesity (BMI$\in [30, 35)$), class two obesity (BMI$\in [35, 40)$), and class three obesity (BMI> 40). Table 4 presents results for each

[20] We discuss the details of both entropy weighting and the synthetic control method in the Appendix.

category using sample and entropy weights. Table 4 shows that most of the observed decrease in obesity can be explained by individuals transitioning from class 1 obesity with a 2.6 percentage point reduction and a statistically significant increase of 3.1 percentage points in the probability of having normal weight (BMI$\in [20, 25)$). These results indicate a plausible shift in the body weight distribution. Results are consistent between using sample (Panel A) and entropy balanced weights (Panel B).

Table 4

To summarize, using a variety of difference-in-differences estimators, we demonstrate that obesity fell in New York City following the CLM. Table 5 demonstrates that our obesity results are robust. Panel A of Table 5 varies how we define treatment and control groups and what we control for. For example, commuters to New York City may be exposed to the CLM but not enter our treatment group because they live outside of the city. We treat this group as treated, our results are qualitatively similar. Furthermore, when we restrict our sample to just northeastern states, and even just northeastern cities, our results are similar. Panels B Table 5 show that our results are unique to 2008, and Panel C of Table 5 shows no effects in placebo tests for rates of smoking and asthma.

Table 5

3.1 NYC Mandate on Physical Activity

How the obesity results demonstrated above could occur is the focus of the rest of this paper. As noted above, there is little evidence of the NYC-CLM inducing a difference in fast food behavior, and Restrepo (2017) finds no evidence that physical activity in counties in the State of New York changed significantly. To investigate, we again focus on New York City, and we use data from the American Time Use Survey (ATUS) from the Bureau of Labor Statistics from 2004 to 2012. ATUS measures the amount of time people spend in different activities during a 24-hour recall period that ends on 4 AM of the interview day, which provides a detailed calculation of the impact measured in this analysis. It is a nationally representative survey done via telephone interviews. The sample comes from a sub-sample of Current Population Survey (CPS) respondents.

An advantage of using ATUS is the disaggregation in the activities which an individual reports, with whom they do these activities, and where they were done. For each activity, Tudor-Locke *et al.* (2009), demonstrate how to construct metabolic equivalents of task (MET), which measures the ratio of the metabolic rate for a given activity to a standard resting metabolic rate. One MET is approximately equal to 1 calorie expended per kilogram of body mass per hour. To construct our MET variable of interest we add up the number of METs for each individual given the activities and duration of activities of that person. For example, in the extreme condition that a person only does extremely sedentary activities (MET=1) all day, their total number of METs would be 24. The MET measure allows the researcher to categorize activities listed in ATUS by their level of strenuousness, and we define the following three types of activities:

- Sedentary Activities ($MET \in [0; 1.5)$), such as watching TV, relaxing, waiting, reading, riding in a car, sleeping, etc.
- Light Activities ($MET \in [1.5; 3)$), such as socializing, attending sports events, computer use, shopping, walking at a slow pace, grocery shopping, etc.
- Moderate to Vigorous Activities ($MET \geq 3$), such as dancing, doing sports, hiking, playing with household children

To create these variables, we add up for every individual the number of minutes a day in which they do that specific type of activity. For example, if an individual spends 2 hours watching TV on the reference day and 1 hour reading, they would have 180 minutes of sedentary activities.

Table 6 presents results from estimation of Equation 2 in which we consider various physical activity dependent variables. The first column uses only sample weights, whereas the second column uses our preferred entropy-weighting scheme. Our results suggest a statistically significant increase in Metabolic Equivalents of Task of 1.1 percent, which translates to an extra 28.1 calories burned per day. Point estimates in Column two indicate that the overall increase in activity is explained mostly by a decline in minutes of sedentary activities of 1.9 percent and an increase in minutes of light activities of 3.1 percent although these estimates are not statistically significant as the Ferman-Pinto P-Value shows. We find no statistically significant effect for minutes of moderate to vigorous activities. The magnitude and direction of these results are similar between sample and entropy weights. For both the sample and entropy weighted estimates, the P-Values presented in Table 6 imply that we fail to reject similar pre-period trends in any of the outcomes. Additionally, we find a 23.3 percent statistically significant increase in minutes of recreational activities and sports which underscores the overall increase in activity levels shown by metabolic equivalents of task and the sign and direction of estimates for the level of strenuousness of individual activities.

TABLE 6

To highlight the difference between using sample and entropy weights for physical activity, Figure 4 shows trends in log of metabolic equivalents of task and minutes of recreational activities and sports for New York City and control cities under both weighting schemes. Figure 5 does the same for minutes of sedentary and light activities. Both figures show that the pre-treatment variation between treated and control units is reduced when using entropy weights.

Figure 4

Figure 5

Following results of Deb and Vargas (2016), who find significant heterogeneity in the effects of calorie labeling, Table 7 presents results that investigate heterogeneity by demographic and socioeconomic characteristics. We estimate Equation 2 separately for each group, and we report the pre-trend P-Value and the number of observations. Focusing on the entropy-weighted results in column two, our MET results are driven by white NYC residents, who saw a 2.0% increase in MET following the CLM although the result is not statistically significant as shown by the Ferman-Pinto P-Value.

Table 7

Table 8 demonstrates that our activity results are not sensitive to how we treat commuters into New York City, the inclusion or exclusion of similar Northeastern cities, or controls for the 2007 Trans-Fat Ban. Additionally, placebo tests for minutes spent listening to the radio or consuming tobacco or drugs show no significant effects. Of course, listening to the radio may be a sedentary activity, but the sign of δ in this regression is negative, suggesting that it contributes to the significant reduction in sedentary activity found above. Our results suggest robust evidence of an increase in physical activity - mainly through a reduction in sedentary activity - that serves a plausible mechanism for a reduction in obesity.

Table 8

3.2 Expenditure and Frequency of Visits to Fast Food Restaurants

Finally, following Cawley (2004), we investigate the extent to which the NYC-CLM changed both the frequency of and expenditure at fast food ordering. Individual-level expenditure data come from the Diary Survey of the Consumer Expenditure Survey (CEX) from the Bureau of Labor Statistics from 2006 to 2012.[21]. CEX data treat New York City as a stratum, which allows us to measure the impact of the NYC-CLM on frequency of visits to fast food establishments, expenditure per visit and overall expenditure. The diary survey asks individuals how many times and how much money they spent at fast food restaurants.

Table 9 presents results for expenditure and frequency of visits to fast food establishments using sampling and entropy weights. We find a positive and significant increase in expenditures of approximately 6.2% of a standard deviation when we focus on expenditure per fast food visit. We find a statistically significant decline in expenditure per consumer unit (household) equivalent to 4.2% of a standard deviation. Finally, our results suggest statistically significant declines in the frequency of fast food visits per month of approximately 8.2% of a standard deviation. Unfortunately, the P-Value on the null hypothesis that pre-NYC-CLM trends in fast food frequency and expenditure per visit are parallel between New York City and other areas is 0.000. However, results using entropy weights where pre-trends are identical have similar point estimates and levels of significance. As with mixed previous results on calorie labeling, we find that the calorie labeling mandate in NYC may have changed fast food frequency and expenditure behavior, though perhaps not enough by itself to drive observed changes in obesity.

Table 9

4 Discussion and Conclusion

The decline in weight that would result from the 1.3 percent decrease in BMI is equivalent to 1.16 kg. The necessary energy imbalance gap (between intake and expenditure) to attain this decrease needs to be around 43 calories per day for a year using the calculations from Hall *et al.* (2011)[22]. In our preferred estimation for physical activity, we find that calorie labeling in NYC led to a 1 percent increase in metabolic equivalents of task. This increase translates to 28.1 daily extra calories burned on average (95% CI 12-45 calories). To put these results in context, if the average

[21] 2006 is the first year in which we can identify observations that come from New York City

[22] `https://www.niddk.nih.gov/bwp`

individual changes their energy imbalance gap by 28 calories per day for a year, it would translate to a 0.6 kg decrease in weight. For BMI results to be attainable, individuals would have to consume 15 less calories on average to achieve the necessary energy imbalance gap (a tablespoon of ketchup has approximately 20 calories). The decrease in BMI would be attainable given the observed increase in activity if we have decreases in calorie consumption after calorie labeling equivalent to those found by Bollinger and Sorensen (2011), VanEpps *et al.* (2016) and Wisdom *et al.* (2010) .

Results in this paper provide some hope for calorie labeling. While much attention has been paid to modest behavioral responses at fast food restaurants, we argue that calorie labels may have promoted an overall awareness of obesity and its determinants that motivated behavioral changes mostly on levels of physical activity. Given our results in physical activity after exposure to calorie labeling, it will be relevant to test whether physical activity equivalents on calorie labels have a differential impact on consumption and levels of activity.

References

Abadie, A., Diamond, A. and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, **105** (490), 493–506.

— and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, **93** (1), 113–132.

Bleich, S. N., Economos, C. D., Spiker, M. L. and et al. (2017). A systematic review of calorie labeling and modified calorie labeling interventions: Impact on consumer and restaurant behavior. *Obesity*, **25** (12), 2018–2044.

Block, J. P., Condon, S. K., Kleinman, K. and et al. (2013). Consumers' estimation of calorie content at fast food restaurants: cross sectional observational study. *BMJ*, **346** (2907), 1–10.

Bollinger, P. L., Bryan and Sorensen, A. (2011). Calorie posting in chain restaurants. *American Economic Journal: Economic Policy*, **3**, 91–128.

Burton, S., Creyer, E. H., Kees, J. and Huggins, K. (2006). Attacking the obesity epidemic: the potential health benefits of providing nutrition information in restaurants. *American Journal of Public Health*, **96** (9), 1669–1675.

—, Tangari, A. H., Howlett, E. and Turri, A. M. (2014). How the Perceived Healthfulness of Restaurant Menu Items Influences Sodium and Calorie Misperceptions: Implications for Nutrition Disclosures in Chain Restaurants. *Journal of Consumer Affairs*, **Spring**, 62–95.

Cantor, J., Torres, A., Abrams, C. and Elbel, B. (2015). Five years later: Awareness of new york city's calorie labels declined, with no changes in calories purchased. *Health Affairs*, **34** (11), 1893–1900.

Cawley, J. (2004). An economic framework for understanding physical activity and eating behaviors. *American journal of preventive medicine*, **27** (3), 117–125.

— (2015). An economy of scales: A selective review of obesity's economic causes, consequences, and solutions. *Journal of Health Economics*, **43**, 244–268.

— (2018). The impact of information disclosure on consumer behavior: Evidence from a randomized field experiment of calorie labels on restaurant menus. *NBER Working Paper*, **24889**.

Charness, G. and Gneezy, U. (2009). Incentives to exercise. *Econometrica*, **77** (43), 909–931.

Conley, T. G. and Taber, C. R. (2011). Inference with "difference in differences" with a small number of policy changes. *The Review of Economics and Statistics*, **93** (1), 113–125.

Currie, J., DellaVigna, S., Moretti, E. and Pathania, V. (2010). The effect of fast food restaurants on obesity and weight gain. *American Economic Journal: Economic Policy*, **2** (August), 32–63.

Deb, P. and Vargas, C. (2016). Who benefits from calorie labeling? an analysis of its effects on body mass. *NBER Working Paper*, (21992).

Dumanovsky, T., Huang, C. Y., Bassett, M. T. and Silver, L. D. (2010). Consumer awareness of fast-food calorie information in new york city after implementation of a menu labeling regulation. *American Journal of Public Health*, **100** (12), 2520–2525.

Elbel, B., Gyamfi, J. and Kersh, R. (2011). Child and adolescent fast-food choice and the influence of calorie labeling: a natural experiment. *International Journal of Obesity*, **35** (4), 493–500.

—, Kersh, R., Brescoll, V. L. and Dixon, L. B. (2009). Calorie labeling and food choices: a first look at the effects on low-income people in new york city. *Health Affairs*, **28** (6), w1110–w1121.

—, Mijanovich, T., Dixon, B., Abrams, C., Weitzman, B., Kersh, R., Auchincloss, A. and Ogedegbe, G. (2013). Calorie labeling, fast food purchasing and restaurant visits. *Obesity*, **21** (11), 2172–2179.

Ferman, B. and Pinto, C. (2017). Placebo tests for synthetic controls.

— and — (2019). Inference in differences-in-differences with few treated groups and heteroskedasticity. *Review of Economics and Statistics*, **101** (3), 452–467.

Finkelstein, E. A., Strombotne, K. L., Chan, N. L. and Krieger, J. (2011). Mandatory menu labeling in one fast-food chain in king county, washington. *American journal of preventive medicine*, **40** (2), 122–127.

Flegal, K., Kruszon-Moran, D., Carroll, M., Fryar, C. and Ogden, C. (2016). Trends in obesity among adults in the united states, 2005 to 2014. *Journal of the American Medical Association*, **315** (21), 2284–2291.

for Disease Control, C. and (CDC), P. (2004-2012). Behavioral risk factor surveillance system survey data. atlanta, georgia: U.s. department of health and human services, centers for disease control and prevention.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, **80** (2), 223–255.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, **20**, 25–46.

Hall, K., Sacks, G., Chandramohan, D., Chow, C., Wang, C., Gortmaker, S. and Swinburn, B. (2011). Quantification of the effect of energy imbalance on bodyweight. *The Lancet*, **9793** (378), 826–837.

Hofferth, S., Flood, S. and Sobek, M. (2004-2012). American time use survey data extract builder: Version 2.7.[dataset], college park, md: University of maryland and minneapolis, mn: Ipums.

Krieger, J. W., Chan, N. L., Saelens, B. E., Ta, M. L., Solet, D. and Fleming, D. W. (2013). Menu labeling regulations and calories purchased at chain restaurants. *American journal of preventive medicine*, **44** (6), 595–604.

Lakdawalla, D., Philipson, T. and Bhattacharya, J. (2005). Welfare-Enhancing Technological Change and the Growth of Obesity. *American Economic Review*, **95** (2), 253–257.

MacKinnon, J. G. and Webb, M. D. (2018). The wild bootstrap for few (treated) clusters. *The Econometrics Journal*, **21** (2), 114–135.

of Labor Statistics, U. S. B. (2006-2012). Consumer expenditure survey.

Ogden, C. L., Carroll, M. D., Lawman, H., Fryar, C., Kruszon-Moran, D., Kit, B. and Flegal, K. (2016). Trends in obesity prevalence among children and adolescents in the united states, 1988-1994 through 2013-2014. *Journal of the American Medical Association*, **315** (21), 2292–2299.

Powell, L. M., Nguyen, B. T. and Han, E. (2012). Energy intake from restaurants: demographics and socioeconomics, 2003–2008. *American journal of preventive medicine*, **43** (5), 498–504.

Restrepo, B. (2017). Calorie labeling in chain restaurants and body weight: Evidence from new york. **26**, 1191–1209.

Roberto, C. A., Larsen, P. D., Agnew, H., Baik, J. and Brownell, K. D. (2010). Evaluating the impact of menu labeling on food choices and intake. *American Journal of Public Health*, **100** (2), 312–318.

Tudor-Locke, C., Washington, T. L., Ainsworth, B. E. and Troiano, R. P. (2009). Linking the american time use survey (atus) and the compendium of physical activities: methods and rationale. *Journal of Physical Activity and Health*, **6** (3), 347–353.

Vadiveloo, M., Dixon, B. and Elbel, B. (2011). Consumer purchasing patterns in response to calorie labeling legislation in new york city. *International Journal of Behavioral Nutrition and Physical Activity*, **8** (51).

VanEpps, E. M., Downs, J. S. and Loewenstein, G. (2016). Calorie label formats: using numeric and traffic light calorie labels to reduce lunch calories. *Journal of Public Policy & Marketing*, **35** (1), 26–36.

Wang, Y., Gortmaker, S., Sobol, A. and KM, K. (2016). Estimating the energy gap among us children: a counterfactual approach. *Pediatrics*, **118** (6), 1721–1733.

Wisdom, J., Downs, J. S. and Loewenstein, G. (2010). Promoting healthy choices: Information versus convenience. *American Economic Journal: Applied Economics*, **April, 2**, 164–178.

5 Tables

Table 1: Summary of Literature on Calorie Labeling

Source	Population	Results	Data
Panel A: Impact on Obesity and Body Mass Index (BMI)			
1. Restrepo (2017)	Adults in NYC	12% decrease in Obesity,1.5% decline in BMI	BRFSS 2004-2012
2. Deb and Vargas (2016)	Women and Men exposed to labeling in United States	-.0309 decrease in men, non-sig. results for women	BRFSS 2004-2012
Panel B: Calories Purchasing Behavior			
1. Bollinger and Sorensen (2011)	Starbucks clients in NYC	Calories per transaction fall by 6%	Starbucks transaction data
2. Roberto *et al.* (2010)	Participants in a study diner	Calories consumed decreased 14%per transaction	Randomized diner participants
3. Flegal *et al.* (2016)	Children from low-income areas in NYC and Newark, NJ	No effect on calories consumed	Chain restaurant clients data
4. Finkelstein *et al.* (2011)	Clients in 7 King County restaurants and 7 control locations	No effect on calories consumed	Chain restaurant clients data
5. Wisdom *et al.* (2010)	Clients in fast food restaurants	Calories consumed decreased (60 calories)	Diners at fast food restaurant
6. Elbel *et al.* (2009)	Fast food clients from low-income areas in NYC and NJ	No effect on calories consumed	Receipts and survey responses
7. Dumanovsky *et al.* (2010)	Fast food clients in NYC	No overall effect on calories, decrease in some fast food locations	Receipts from 11 fast food chains before and after intervention
8. Krieger *et al.* (2013)	Clients in chain restaurants from King County, WA	No effect on calories	Receipts from before and after the intervention
9. VanEpps *et al.* (2016)	Employees from a corporation ordering lunch on-line	60-78 calorie decline	Randomized control trial

For a thorough review of calorie labeling impact on consumers and restaurants, see Bleich *et al.* (2017).

Table 2: Descriptive Statistics for BRFSS, ATUS and CES

	2004-2008		2008-2012		
	NYC	Control Cities	NYC	Control Cities	DID P-Value
	(1)	(2)	(3)	(4)	(5)
Panel A: BRFSS					
Body Mass Index	26.582	27.202	26.595	27.575	0.001
Obesity	0.210	0.250	0.213	0.277	0.000
Women	0.508	0.490	0.507	0.492	0.109
African-American	0.215	0.130	0.249	0.142	0.033
Hispanic	0.240	0.126	0.203	0.132	0.030
White	0.406	0.689	0.418	0.667	0.063
Ed. Highschool or Less	0.346	0.332	0.342	0.336	0.248
Ed. More than Highschool	0.654	0.668	0.658	0.664	0.248
Panel B: American Time Use Survey					
Met. Equiv. of Task (METs)	36.878	38.291	36.998	37.872	0.000
Mins. of Recreational Act.	12.011	16.142	15.117	16.038	0.000
Mins. of Seden. Act.	393.493	351.174	378.541	356.369	0.000
Mins. of Mod. to Vig. Act.	71.001	95.892	69.851	94.607	0.691
Women	0.496	0.519	0.528	0.517	0.000
African-American	0.327	0.140	0.318	0.150	0.000
Hispanic	0.289	0.119	0.267	0.137	0.000
White	0.355	0.703	0.335	0.669	0.873
Ed. Highschool or Less	0.508	0.462	0.435	0.441	0.000
Ed. More than Highschool	0.492	0.538	0.565	0.559	0.000
Panel C: Consumer Expenditure Survey					
Expenditure per Visit	5.411	6.351	5.606	6.284	0.001
Freq.of Fast Food Visits	3.647	2.682	3.206	2.527	0.000
Expenditure per Consumer Unit	19.732	17.375	17.975	16.135	0.191
Women	0.569	0.521	0.585	0.543	0.535
African-American	0.244	0.127	0.255	0.133	0.176
Hispanic	0.302	0.103	0.261	0.107	0.000
White	0.346	0.728	0.371	0.717	0.000
Ed. Highschool or Less	0.496	0.382	0.437	0.362	0.000
Ed. More than Highschool	0.492	0.616	0.559	0.636	0.000

Notes: Authors estimations using BRFSS-SMART (total sample: 1,104,495), ATUS (total sample: 56,459) and CES (total sample: 78,579) data from 2004 to 2012. Treatment individuals are those living in one of the five New York City counties/boroughs. Column 5 shows the resulting P-Value of regressing each row variable under a difference in differences estimation model.

Table 3: Estimation Results for BMI and Obesity in NYC

	BMI Logarithm, s.d.:0.189			Obesity, mean: 0.210		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Difference in Differences						
NYC*Post	-0.016***	-0.015***	-0.014***	-0.027***	-0.024***	-0.023***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.007)	(0.008)
Pre-trend P-Value			0.681			0.765
Ferman-Pinto P-Value			0.000^{+++}			0.005^{+++}
Panel B: Entropy Weights						
NYC*Post	-0.022***	-0.012***	-0.013***	-0.039***	-0.017	-0.020*
	(0.005)	(0.004)	(0.005)	(0.006)	(0.011)	(0.010)
Ferman-Pinto P-Value			0.000^{+++}			0.012^{++}
Panel C: Synthetic Control Method (Average Effect)						
NYC*Post			-0.008**			-0.0196**
Year-Month Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
County controls	No	Yes	Yes	No	Yes	Yes
Sociodem. Charact.	No	No	Yes	No	No	Yes

Notes: N=1,104,495. Authors estimations using BRFSS-SMART data from 2004 to 2012. Treatment individuals are those living in one of the five New York City counties/boroughs. Control individuals are those living in any of 370 other counties within MMSAs in the United States. Obesity results are from linear probability models. Robust standard errors clustered by county in parentheses. The P-Value on pre-trends is with respect to the null hypothesis that trends in the conditional expectation are parallel prior to treatment. The Ferman and Pinto (2019) P-Value uses their bootstrapping procedure to address having only one treatment unit (NYC) and heteroskedastic errors. For more information on the methods behind entropy weighting and synthetic control method, see the Appendix. Statistical inference for Synthetic Control Method was obtained using a permutation test with 49 other control units which consist of the 50 most populated cities in the United States. * p<0.10, ** p<0.05, *** p<0.01 with clustered robust standard errors. $^{+}$ p<0.10, $^{++}$ p<0.05, $^{+++}$p<0.01 with Ferman-Pinto (2019) inference method.

Table 4: Results by BMI Category using Sample and Entropy Weights

BMI Category	[20-25)	[25-30)	[30-35)	[35-40)	[>40]
	(1)	(2)	(3)	(4)	(5)
Panel A: Sample Weights					
NYC*Post	0.024***	-0.003	-0.020***	0.003	-0.005*
	(0.007)	(0.011)	(0.007)	(0.005)	(0.003)
Mean	0.413	0.360	0.145	0.039	0.025
Pre-trend P-Value	0.122	0.560	0.261	0.295	0.024
Ferman-Pinto P-Value	0.007+++	0.538	0.003+++	0.241	0.029++
R-squared	0.047	0.011	0.017	0.012	0.016
Panel B: Entropy Weights					
NYC*Post	0.031***	-0.003	-0.026***	0.007	-0.001
	(0.009)	(0.012)	(0.009)	(0.005)	(0.005)
Mean	0.413	0.360	0.145	0.039	0.025
Ferman-Pinto P-Value	0.014++	0.628	0.000+++	0.037++	0.526
R-squared	0.080	0.023	0.031	0.022	0.023

Notes: N=1,104,495. Authors estimations using SMART-BRFSS from 2004 to 2012. Treatment individuals are those living in one of the five New York City counties/boroughs. Control individuals are those living in any of 370 other counties within MMSAs in the United States. Results come from linear probability models. Robust standard errors clustered by county in parentheses. The P-Value on pre-trends is with respect to the null hypothesis that trends in the conditional expectation are parallel prior to treatment. The Ferman and Pinto (2019) P-Value uses their bootstrapping procedure to address having only one treatment unit (NYC) and heteroskedastic errors.
* p<0.10, ** p<0.05, *** p<0.01 with clustered robust standard errors. + p<0.10, ++ p<0.05, +++p<0.01 with Ferman-Pinto (2019) inference method.

Table 5: Estimates from Alternative Specifications and Dependent Variables using BRFSS

	Log(BMI) (1)	Obesity (2)
Panel A: Alternative Control/Treatment Definitions		
Northeast States	-0.013***	-0.024**
	(0.003)	(0.010)
Northeast Megapolis	-0.012***	-0.017**
	(0.003)	(0.008)
With NYC Commuters as Treated	-0.010***	-0.021***
	(0.002)	(0.005)
Without Control for Trans-Fat Ban	-0.013***	-0.024***
	(0.004)	(0.007)
Without Time to Comply	-0.017***	-0.031***
	(0.006)	(0.011)
Panel B: Placebo Policy Timing		
Intervention in 2007	-0.006	-0.005
	(0.004)	(0.012)
Intervention in 2006	-0.003	0.001
	(0.006)	(0.011)
Panel C: Unrelated Dependent Variables		
Smokes	0.014	
	(0.012)	
Asthma	0.009	
	(0.009)	

Notes: Authors estimations using SMART-BRFSS from 2004 to 2012. Each difference in-differences estimate is generated from a separate regression. Robust standard errors clustered by county in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Northeast States results use as controls Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New Jersey, New York, and Pennsylvania. Northeast Megalopolis use as controls District of Columbia, Virginia, Maryland, Delaware, Pennsylvania, New Jersey, New York, Connecticut, Rhode Island, Massachusetts, New Hampshire, and Maine. With Commuters as Treated results include commuter counties to NYC as part of the treatment group. Without Control for Trans-Fat Ban results excludes a control for the timing and counties who banned cooking with oils that have trans fats. Without Time to Comply results does not include observations from May to December 2008.

Table 6: Estimation Results for Activity Outcomes

	(1)	(2)
Metabolic Equivalents of Task Logarithm, Std. Dev.: 0.173		
NYC*Post	0.008***	0.011***
	(0.002)	(0.003)
Pre-trend P-Value	0.296	0.368
Ferman-Pinto P-Value	0.112	0.021++
Minutes of Recreational Activities and Sports, Mean: 12.011		
NYC*Post	2.171***	2.799***
	(0.665)	(0.650)
Pre-trend P-Value	0.493	0.453
Ferman-Pinto P-Value	0.119	0.040++
Minutes of Sedentary Activities, Mean: 921.575		
NYC*Post	-18.853***	-17.204***
	(2.657)	(3.614)
Pre-trend P-Value	0.003	0.012
Ferman-Pinto P-Value	0.024++	0.420
Minutes of Light Activities, Mean: 437.194		
NYC*Post	17.841***	13.539***
	(3.072)	(4.139)
Pre-trend P-Value	0.812	0.640
Ferman-Pinto P-Value	0.021++	0.509
Minutes of Moderate to Vigorous Activities, Mean: 71.001		
NYC*Post	0.839	4.321*
	(2.113)	(2.504)
Pre-trend P-Value	0.000	0.000
Ferman-Pinto P-Value	0.746	0.213
Year-Month and CBSA Fixed Effects	Yes	Yes
Sociodem. Charact.	Yes	Yes
Sample Weights	Yes	No
Entropy Weights	No	Yes

Notes: N= 56,459. Robust standard errors clustered at the CBSA level. * p<0.10, ** p<0.05, *** p<0.01 with clustered robust standard errors. + p<0.10, ++ p<0.05, +++p<0.01 with Ferman-Pinto (2019) inference method. Column 1 shows results using sample weights, Column 2 results with entropy weights. The P-Value on pre-trends is with respect to the null hypothesis that trends in the conditional expectation are parallel prior to treatment. The Ferman and Pinto (2019) P-Value uses their bootstrapping procedure to address having only one treatment unit (NYC) and heteroskedastic errors. Authors estimations using ATUS data from 2004 to 2012.

Table 7: Estimation Results for Metabolic Equivalents of Task Logarithm for Selected Subpopulations

	(1)	(2)
Panel A: Women		
NYC*Post	0.015***	0.018***
	(0.002)	(0.003)
Pre-trend P-Value	0.032	0.643
Ferman-Pinto P-Value	0.000^{+++}	0.356
Observations	32,154	32,154
Panel B: Men		
NYC*Post	-0.001	0.002
	(0.004)	(0.005)
Pre-trend P-Value	0.028	0.135
Ferman-Pinto P-Value	0.847	0.952
Observations	24,305	24,305
Panel C: Black population		
NYC*Post	-0.010*	-0.012**
	(0.005)	(0.005)
Pre-trend P-Value	0.000	0.000
Ferman-Pinto P-Value	0.055^{+}	0.000^{+++}
Observations	9,845	9,845
Panel D: Hispanic population		
NYC*Post	0.002	0.007
	(0.007)	(0.009)
Pre-trend P-Value	0.000	0.000
Ferman-Pinto P-Value	0.508	0.081^{+}
Observations	7,498	7,498
Panel E: White population		
NYC*Post	0.016***	0.020***
	(0.003)	(0.004)
Pre-trend P-Value	0.000	0.001
Ferman-Pinto P-Value	0.098^{+}	0.397
Observations	36,979	36,979
Year-Month and CBSA Fixed Effects	Yes	Yes
Sociodem. Charact.	Yes	Yes
Sample Weights	Yes	No
Entropy Weights	No	Yes

Notes: Robust standard errors clustered at the CBSA level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$ with clustered robust standard errors. $^{+}$ $p<0.10$, $^{++}$ $p<0.05$, $^{+++}$$p<0.01$ with Ferman-Pinto (2019) inference method. The Ferman and Pinto (2019) P-Value uses their bootstrapping procedure to address having only one treatment unit (NYC) and heteroskedastic errors. Authors estimations using ATUS data from 2004 to 2012.

Table 8: Estimates from Alternative Specifications and Dependent Variables using ATUS

	Log(MET) (1)	Log(MET) (2)
Panel A: Alternative Control/Treatment Definitions		
Northeast Megapolis	0.011*	0.012*
	(0.006)	(0.006)
With NYC Commuters as Treated	0.009***	0.010***
	(0.002)	(0.003)
Without Control for Trans-Fat Ban	0.012***	0.014***
	(0.002)	(0.003)
Without Time to Comply	0.010***	0.013***
	(0.002)	(0.003)
Panel B: Placebo Policy Timing and Weights		
No Weights	0.005***	0.005***
	(0.002)	(0.002)
Intervention in 2007	0.004*	0.005
	(0.002)	(0.003)
Intervention in 2006	0.016***	0.015***
	(0.003)	(0.004)
Panel C: Unrelated Dependent Variables		
Minutes Listening to the Radio	-0.298	-0.281
	(0.197)	(0.317)
Minutes Consuming Tobacco or Drugs	0.052	0.095
	(0.060)	(0.059)
Sample Weights	Yes	No
Entropy Balanced Weights	No	Yes

Notes: * .10 ** .05 *** .01 sig. levels. Robust standard errors clustered at the CBSA level. Authors estimations using ATUS data from 2004 to 2012. Northeast Megalopolis use as controls District of Columbia, Virginia, Maryland, Delaware, Pennsylvania, New Jersey, New York, Connecticut, Rhode Island, Massachusetts, New Hampshire, and Maine. With Commuters as Treated results include commuter counties to NYC as part of the treatment group. Without Control for Trans-Fat Ban results excludes a control for the timing and counties who banned cooking with oils that have trans fats. Without Time to Comply results does not include observations from May to December 2008.

Table 9: Estimation Results for Expenditure and Frequency of Visits to Fast Food Restaurants

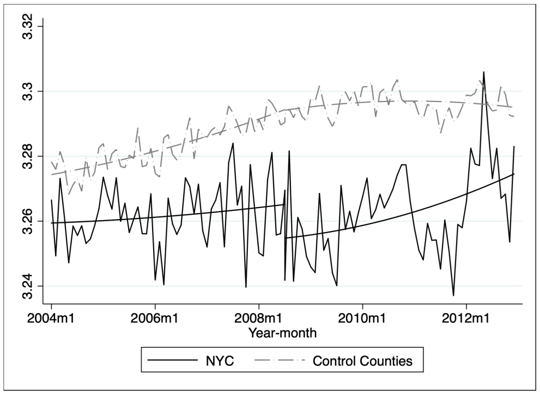
	(1)	(2)	(3)
Expenditure per Visit, Std. Dev.: 5.411	0.421***	0.335**	
	(0.107)	(0.139)	0.000
Observations	207,570	207,570	
Expenditure per Consumer Unit, Std. Dev.: 31.869	-1.664***	-1.346***	
	(0.286)	(0.474)	0.818
Observations	78,579	78,579	
Frequency of Fast Food Visits, Std. Dev.: 5.922	-0.554***	-0.488***	
	(0.054)	(0.079)	0.000
Observations	78,579	78,579	
Year-Month Fixed Effects	Yes	Yes	
State Fixed Effects	Yes	Yes	
Sociodem. Charact.	Yes	Yes	
Sample Weights	Yes	No	
Entropy Weights	No	Yes	
Pre-trend P-Value	No	No	Yes

Notes: * .10 ** .05 *** .01 sig. levels. Robust standard errors clustered at the State level in parentheses. Column 1 shows results using sample weights, Column 2 results with entropy weights and Column 3 the the pre-trend P-Values for the estimations with sample weights. The P-Value on pre-trends is with respect to the null hypothesis that trends in the conditional expectation are parallel prior to treatment.
Authors estimations using CES from 2006 to 2012. Expenditures are in January 2008 prices.

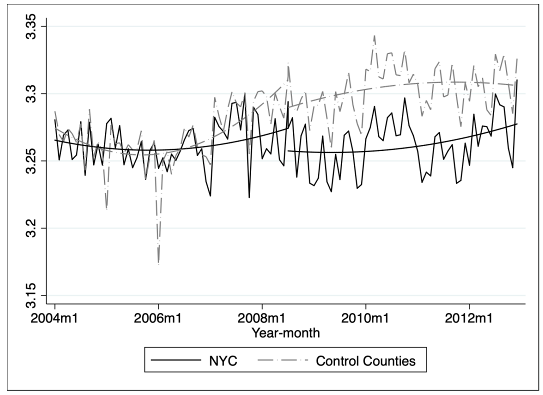
6 Figures

Figure 1: **Average BMI and Obesity Trends with Entropy Weights**

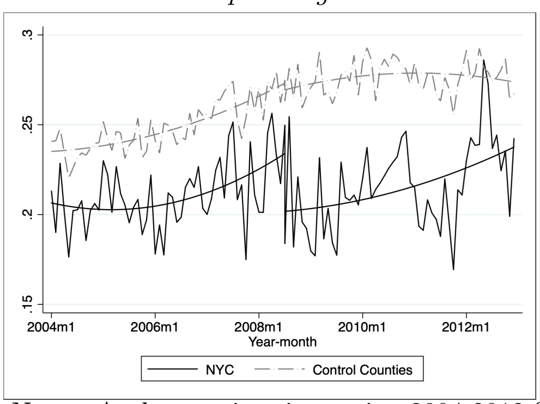
a) ln(BMI):
Sample weights



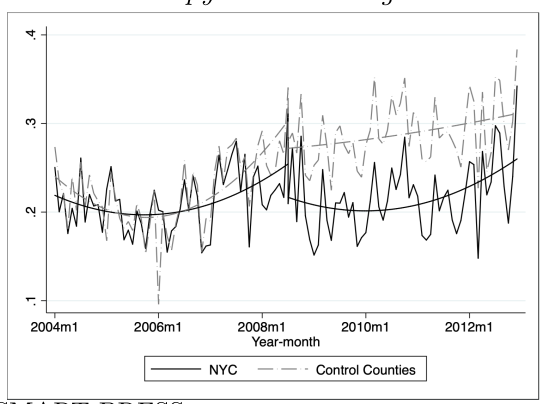
b) ln(BMI):
Entropy balanced weights



c) Obesity:
Sample weights



d) Obesity:
Entropy balanced weights



Notes: Authors estimations using 2004-2012 SMART-BRFSS

Figure 2: **Trends in BMI Logarithm of Adults: New York MSA versus Synthetic New York MSA**

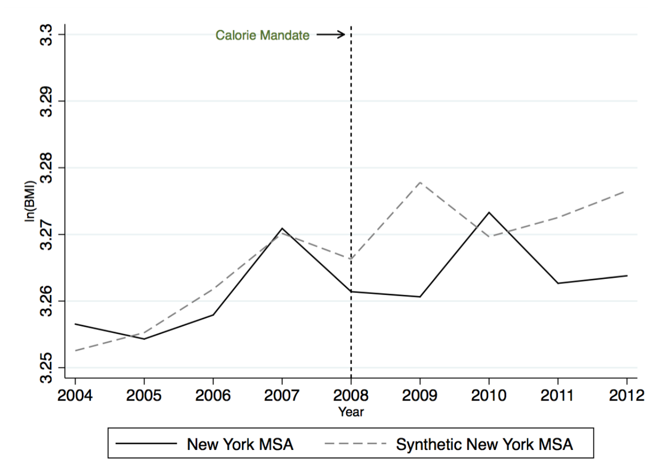


Figure 3: **Trends in Percentage of Obese Adults: New York MSA versus Synthetic New York MSA**

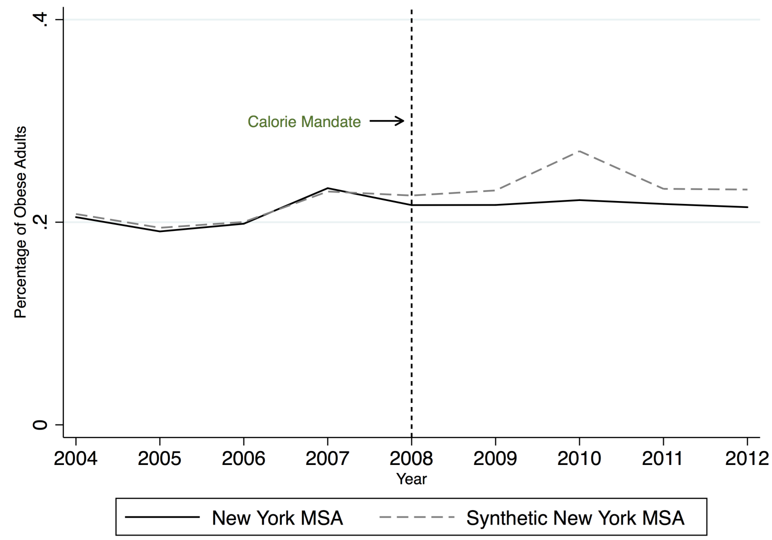
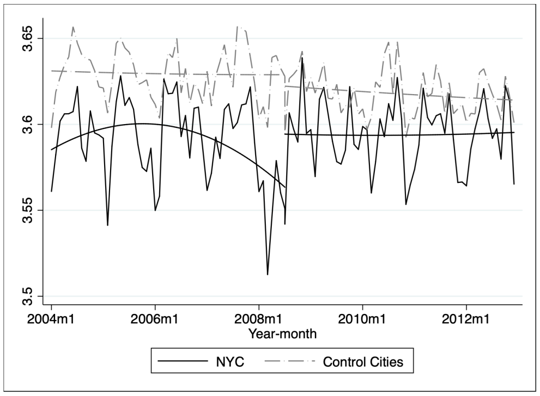
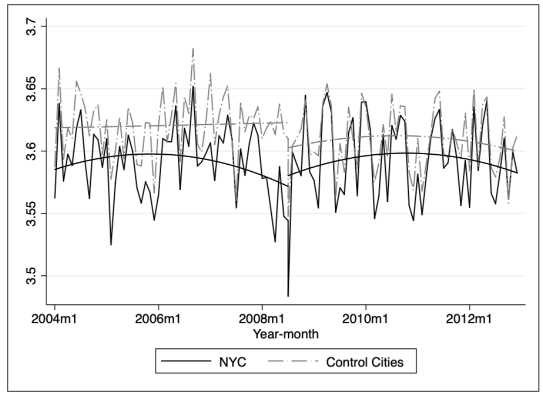


Figure 4: **Average MET and Recreational Activities Trends with Entropy Weights**

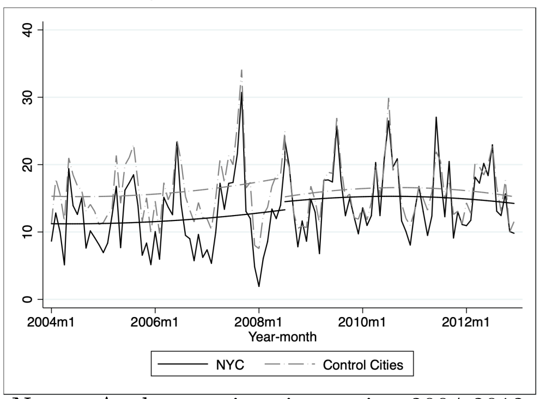
a) ln(MET):
Sample weights



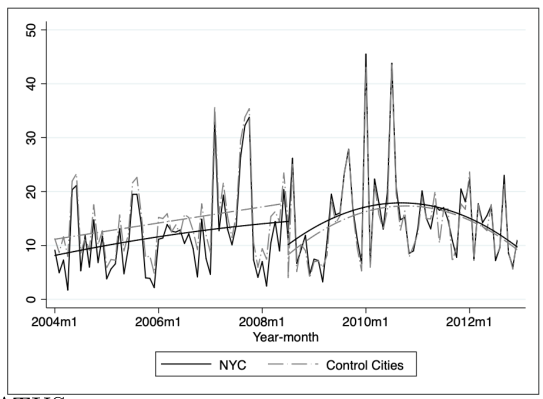
b) ln(MET):
Entropy balanced weights



c) Mins. of Recreational Act. and Sports:
Sample weights



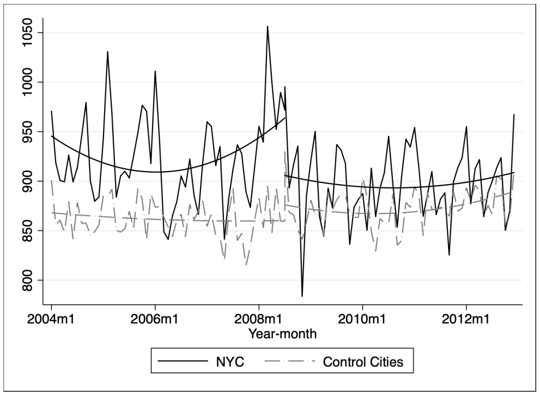
d) Mins. of Recreational Act. and Sports:
Entropy balanced weights



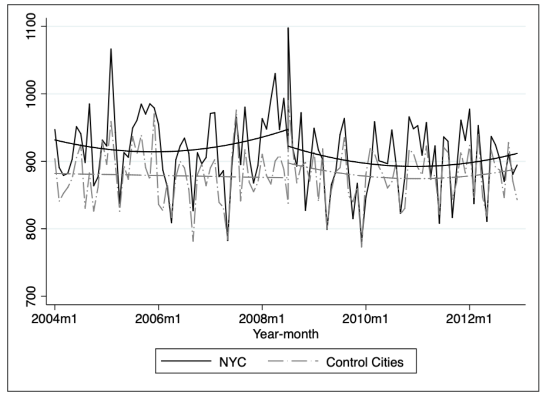
Notes: Authors estimations using 2004-2012 ATUS

Figure 5: **Average Minutes of Sedentary and Light Activities Trends with Entropy Weights**

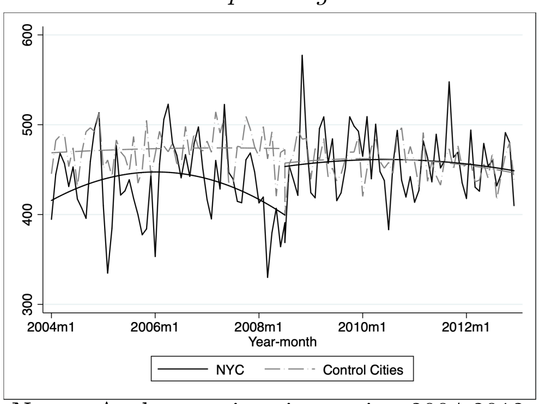
a) Mins. of Sedentary Act.:
Sample weights



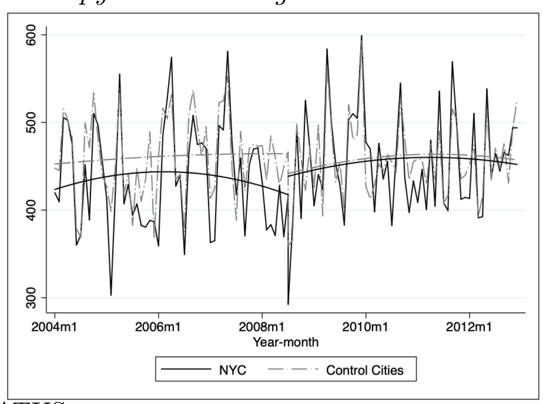
b) Mins. of Sedentary Act.:
Entropy balanced weights



c) Mins. of Light Act.:
Sample weights



d) Mins. of Light Act.:
Entropy balanced weights



Notes: Authors estimations using 2004-2012 ATUS