Who Benefits from Calorie Labeling? An Analysis of its Effects on Body Mass*

Partha Deb†
Hunter College and the Graduate Center, CUNY
NBER

Carmen Vargas
Hunter College, CUNY

April 2016

Abstract

This paper reports the results of an analysis of calorie labeling laws on body mass index (BMI). We use county-level variation in implementation of calorie labeling laws over time in the US to identify the effects of such laws. A difference-in-difference analysis using the 2003 to 2012 waves of the Behavioral Risk Factor Surveillance System, shows a statistically insignificant change in BMI for women and a small, statistically significant negative average treatment effect for men. We estimate finite mixture models and discover that the average treatment effects mask substantial heterogeneity in the effects across three classes of women and men. For both women and men, the three classes, determined within the model, can be described as a subpopulation with normal weight, a second one that is overweight on average and a third one that is obese on average. Estimates from finite mixture models for women show that the effect is largely concentrated among women with BMI distributions centered on overweight. The effects for men are statistically significant for each of the three classes and substantively large for men in the overweight and obese classes. These results suggest that overweight individuals are especially sensitive to relevant information.

Keywords: Health, nutrition, calorie labeling, heterogeneous treatment effects, difference-in-difference design, entropy balance weights, finite mixture models.

*The authors thank seminar participants at the BU/Harvard/MIT Health Economics Seminar, especially David Cutler and Randy Ellis, and participants at the International Health Economics Association, Milan, Italy 2015, especially Edward Norton, for insights that improved the analysis. All remaining errors are our own.

†Corresponding author: Hunter College, Department of Economics, 695 Park Avenue, West 1524, New York, NY 10065; e-mail: partha.deb@hunter.cuny.edu, Phone: +1 212 732 5435.

1 Introduction

Obesity is a prevalent public health problem. Solutions to the health problem remain elusive. A secular increase in total calorie consumption over time is the more likely explanation for the increase in obesity rather than a decrease in calories expended (Bleich et al., 2008; Cutler et al., 2003). Decreases in prices of calorie-dense foods relative to prices of less-dense, more nutritious foods (Christian and Rashad, 2009) are one plausible explanation for the increase in calorie consumption. Another explanation suggests the cause is consumption of food prepared away from home, which has also increased dramatically over time. The share of daily calories consumed coming from restaurants and fast food establishments has increased from 6% to 20% between 1977-78 and 2007-08 (Lin and Guthrie, 2012). Such food often is typically calorie-dense and has poor nutritional value (Anderson and Matsa, 2011; Currie et al., 2010). Therefore, such consumption is often cited as a cause of increased body weight in the US.

Consumers often do not know the caloric and nutrient value of foods (see Section 6.1 in Cawley (2015) for a review of the literature). Not only do consumers generally underestimate the number of calories, but evidence also suggests that the downward bias is positively related to the actual number of calories in the item (Robert Wood Johnson Foundation, 2009). While the provision of nutrition information on packaged foods has been mandated by the federal government since the passing of the Nutrition Labeling and Education Act of 1990, restaurants have been exempted from this requirement. As a response, some cities, counties and states have passed laws beginning in 2008 that mandate posting of nutrition information (typically calorie counts but also, in some cases, nutrient information) on menus and menu boards. New York City implemented mandatory calorie labeling in July 2008 and other states, counties and cities in the Northeastern and Western regions of the US have followed suit.

In this paper, we examine the effect of calorie labeling on body mass. While the imperfect information arguments made above and in references therein suggest that consumers should respond to calorie information posted on menus in restaurants, whether and by how much consumers actually respond remains an empirical question. We use the 2003 to 2012 waves of the Behavioral Risk Factor Surveillance System (BRFSS) and a difference-in-difference empirical strategy exploiting variation in implementation of calorie labeling laws across counties in the United States over time in our analysis. We estimate models separately for men and women, and for each gender, focus on heterogeneous treatment effects which we elicit using finite mixture models. As Cawley (2015, section 7.7) explains, discovering and explaining heterogeneity of effects is an important research frontier.

Empirical evidence on the effects of calorie labeling is mixed (Robert Wood Johnson Foundation, 2009). A number of studies have examined behavior at the point of purchase. Bassett et al. (2008) examined purchasing behavior at Subway restaurants that voluntarily posted calorie information and found that individuals who noticed the information purchased fewer calories than those who did not. Wisdom and Loewenstein, (2010) used a similar study design and found that disclosing calorie information at a fast food chain restaurant led to consumers ordering fewer sandwiches with fewer calories. Elbel et al. (2009) used a difference-in-difference design comparing treated and untreated chain restaurants before and after implementation of the New York City Law and found that calorie labeling had no impact on consumption. Dumanovksy et al. (2011) employed a similar study design and reached a similar conclusion. Bollinger et al. (2011) used transactions data from Starbucks and found that calorie labeling did result in fewer calories consumed per transaction. They also found that calorie labeling had larger effects for women and a priori high-calorie purchasers. Restrepo (2014) examines effects of mandated calorie labeling in a number of New York counties between 2004 and 2012. Restrepo uses a difference-in-difference design and data from the BRFSS to examine the effects of calorie labeling on the body mass index (BMI). He finds robust evidence of significant decreases in BMI due to calorie labeling. He estimates quantile regressions to show that the effects of calorie labeling are generally larger in the upper quantiles of the BMI distribution.

Our study expands the sample from New York to a substantial portion of the US. While Restrepo (2014) compares a few counties in New York with calorie labeling to those without, we compare a number of counties, cities and states in the US with calorie labeling laws implemented to neighboring geographies without calorie labeling. We conduct a detailed examination of heterogeneous treatment effects using finite mixture models. We expect heterogeneous treatment effects because calorie labeling and posting requirements vary across jurisdictions, because only a fraction of the population may observe calorie labels; and because individuals may substitute calories across restaurant and non-restaurant meals. In fact, Dumanovsky et al. (2011) found heterogeneity across genders and Restrepo (2014) along the distribution of BMI. Our analysis encompasses both those possibilities and allows for more.

We estimate finite mixture models (Mclachlan and Peel, 2000; Deb and Trivedi, 1997) to explore the possibility of such heterogeneity, to estimate heterogeneous effects of migration and to characterize the sources of such unobserved heterogeneity. Finite Mixture Models have received increasing attention in the statistics literature mainly because of the number of areas in which such distributions are encountered (see McLachlan andPeel, 2000; Lindsay, 1995). Econometric applications of finite mixture models include the seminal work of Heckman and Singer (1984) to labor economics, Wedel and DeSarbo (1993) to marketing data, El-Gamal and Grether (1995) to data from experiments in decision-making under uncertainty, Deb and Trivedi (1997) to the economics of health care. More recent applications include Ayyagari et al. (2013) and Deb et al. (2011) in studies of BMI and alcohol consumption, Bruhin et al. (2010) to experimental data and Caudill et al (2009) and Günther and Launov (2012) to issues in economic development.

In difference-in-difference regression analysis, identification of causal effects relies heavily on the assumption of identical pre-program trends. Consequently, it can be useful to weight / reweight the sample such that the covariate distribution of the control group becomes more similar to the covariate distribution in the treatment group (Abadie and Imbens, 2011). In addition, following the principles of synthetic control groups proposed by Abadie and Gardeazabal (2003) and Abadie et al. (2010), one can reweight the sample based on pre-trend values of the outcome itself to have identical (not just parallel) pre-treatment trends in BMI. In this paper, we use a novel technique called entropy balancing (Hainmueller, 2011) to estimate weights to perfectly balance covariates and pre-treatment BMI trends.

We find that the average treatment effect is not significantly different from zero for the sample of women but it is negative and statistically significant for the sample of men. These average treatment effects mask substantial heterogeneity in the effects for women and men. The finite mixture models uncover three classes of BMI for women and men: the first is a subpopulation with normal weight, a second subpopulation that is overweight on average and a third one that is morbidly obese on average. There is a statistically significant effect for women in the overweight class while the treatment effects are not significant in the other BMI classes. For men, the effect is statistically significant in each of the three classes but increases substantially in magnitude across normal weight, overweight and obese classes.

The remainder of the paper is organized as follows. Section 2 introduces our empirical strategy and the econometric model and section 3 presents the data. Section 4 discusses the empirical results and provides some interpretations. We wrap up with conclusions in section 5.

2 Empirical strategy and econometric model

2.1 Difference-in-difference specification

Consider a design in which there is a binary treatment indicator G_i (with $G_i = 1$ denoting the treatment group), a binary time indicator T_i (with $T_i = 1$ denoting the intervention period) and X_i denoting a set of control covariates. Then, using the potential outcomes framework, the treatment effect in a difference-in-difference model (Athey and Imbens, 2006) can be written as

$$\tau = E[{y_1}^1 | T_i = 1, G_i = 1, X_i] - E[{y_i}^0 | T_i = 1, G_i = 1, X_i], \tag{1}$$

where ${y_i}^1$ and ${y_i}^0$ denote the potential outcomes with and without treatment respectively. As usual, in the potential outcomes framework, the observed outcome Y is given by

$$y_i = \mathbf{1}[T_i = 1, G_i = 1] \times {y_i}^1 + (1 - \mathbf{1}[T_i = 1, G_i = 1]) \times {y_i}^0 \quad (2)$$

$$= T_i \times G_i \times {y_i}^1 + (1 - T_i \times G_i) \times {y_i}^0 \quad (3)$$

where **1** denotes the indicator function.

In the context of a linear regression model with group and time indicators along with a group-time interaction,

$$y_i = \beta_1 T_i + \beta_2 G_i + \beta_3 T_i G_i + X_i \boldsymbol{\theta} + \varepsilon_i. \quad (4)$$

the two expected potential outcomes can be written as

$$E[{y_i}^0 | T_i = 1, G_i = 1, X_i] = \beta_1 + \beta_2 + X_i \boldsymbol{\theta} \quad (5)$$

and

$$E[{y_i}^1 | T_i, G_i, X_i] = \beta_1 T_i + \beta_2 G_i + \beta_3 T_i G_i + X_i \boldsymbol{\theta} \quad (6)$$

so that the treatment effect $\tau = \beta_3$ is the coefficient on the interaction term in the regression specification. We estimate this basic specification using weighted least squares, where the weights incorporate both survey weights and covariate balance adjustments as described in detail below.

The outcome of interest in our analysis is BMI. Individuals who live in counties with calorie labeling laws implemented are assigned to the treated group while those in counties without implemented calorie labeling laws are assigned to the comparison group. Control covariates include demographic and socioeconomic characteristics and indicators for states and years. Inference is based on standard errors adjusted for clustering at the county level.

2.2 Entropy balance weights

In difference-in-difference regression analysis, identification of causal effects relies heavily on the assumption that pre-program trends of the outcome are identical. In order to achieve balance between treated and control observations, it can be useful to weight / reweight the sample such that the covariate distribution of the control group becomes more similar to the covariate distribution in the treatment group (Abadie and Imbens, 2011). In addition, we include the mean value of the outcome (BMI) for each pre-treatment year of data by county and gender in the set of "covariates" used in the entropy balancing algorithm. Thus, following the principles of synthetic control groups proposed by Abadie and Gardeazabal (2003) and Abadie et al. (2010), not only are the treated and control samples balanced on covariates, they are also balanced on the values of the county-level average outcome in each year of the pre-treatment period.

We use a novel method for generating weights to create balance: entropy balancing. Entropy balancing, a method developed by Hainmueller (2011), produces a set of observation-level weights that directly balances covariate distributions across treated and control groups. Inverse propensity score weighting is the popular method for this purpose (Ho et al., 2007) but the entropy balance method has a number of practical advantages. First, it eliminates the need to back and forth between propensity score specification, estimation and balance checking. Second, propensity score weights can lead to worse balance on some covariate dimensions while improving balance on others (Iacus et al., 2012). Third, while the weights are adjusted as far as is needed to accommodate the balance constraints, at the same time they are kept as close as possible to the base weights to retain information in the reweighted data, so extreme weights are much less likely.

In the propensity score weighting method, every treated unit gets a weight $d_i = 1$ and every control unit gets a weight equal to

$$d_i = \frac{\widehat{p}(x_i)}{1 - \widehat{p}(x_i)}$$

where $\widehat{p}(x_i)$ is the estimated propensity score. In the entropy balancing method, each treated unit gets either a weight $w_i = 1$ or $w_i = s_i$, where s_i is the sampling weight associated with the treated observations, and every control unit gets a weight that satisfies a set of a priori specified balance constraints. Specifically, w_i for the control units are chosen by the solution to

$$\min_{w_i} H(w) = \sum_{i|D=0} w_i \log(w_i/s_i) \tag{7}$$

subject to

$$\sum_{i|D=0} w_i c_{ri}(x_i) = k_r \text{ with } r \in 1, ..., R$$
$$\sum_{i|D=0} w_i = 1$$
$$w_i \geq 0 \text{ for all } i \text{ such that } D = 0$$

where $c_{ri}(Xi) = m_r$ describes a set of R balance constraints imposed on the covariate moments of the reweighted control group. Each balance constraint equates the weighted mean of the covariate in the treated sample to the weighted mean of the covariate in the control sample. In the case of indicator variables, which comprise most of the covariates in the study, equality of means is equivalent to equality of distributions. We conduct entropy balance for each gender and year of data separately, so covariate balance is obtained separately for each gender-year subsample.

We apply the entropy balance algorithm iteratively. In the first application of the algorithm, we set the weights for the treated group to their sampling weights and have one balance constraint for each covariate used in the regression analysis. We use the resulting balance weights to calculate county-level means of BMI by gender in each of the pre-treatment years. The second application of the entropy balance algorithm adds the vector of county-mean BMI in each pre-treatment year (in addition to the regression covariates) into the set of balance constraints. This generates a new set of balance-weights, which we use to recalculate county-level mean BMI for the pre-treatment years. The third application of the entropy balance algorithm uses, once again, balance constraints for covariates and the vector of county-level mean BMI. The revised balance weights have a correlation of 0.99 with the weights from the prior iteration so we consider the process to have converged to a stable set of balance weights.

2.3 Finite mixture model

Most empirical models for estimating treatment effects assume that the effect is constant across the population, or can be interpreted as estimating average effects. Yet there are many reasons for expecting that treatment effects are not constant, or that the average of effects masks substantive and policy relevant heterogeneity. In most large experiments or quasi-experimental designs, there are many opportunities for the actual treatment to be heterogeneous across individual characteristics, household characteristics, sites or geographies, for the intensity of treatment to vary, and for compliance to and effects of treatment to vary by individual or group characteristics. Heterogeneity in each of these dimensions lead to heterogeneity of treatment effects. In this analysis, we expect heterogeneous treatment effects because, for example, calorie labeling and posting requirements vary across jurisdictions, because only a fraction of the population may observe calorie labels; and because individuals may substitute calories across restaurant and non-restaurant meals.

Heterogeneity of treatment effects is typically explored via the use of interaction terms in regression analyses or by stratifying the sample by indicators of the source of heterogeneity. For example, stratified analyses by race or gender are commonplace. However, there are data and statistical limits to the amount of stratification that can be done given a sample, and such analyses increase the risk of false findings. Furthermore, often heterogeneity exists along the distribution of the outcome itself, by complex configurations of observed characteristics, or on unobserved characteristics. Quantile regressions are an appealing technique to explore heterogeneity along the outcome distribution but does not provide insight into the other dimensions of heterogeneity. Finite mixture models can identify heterogeneous treatment effects, if they exist, and characterize that heterogeneity along dimensions of the outcome distribution, observed characteristics and unobserved characteristics.

The density function for a C-component finite mixture (Deb and Trivedi, 1997; McLachlan and Peel, 2000; Deb et al., 2011), is

$$f(y|\mathbf{x};\theta_1,\theta_2,...,\theta_C;\pi_1,\pi_2,...,\pi_C)=\sum_{j=1}^{C}\pi_j f_j(y|\mathbf{x};\theta_j) \qquad (8)$$

where $0 < \pi_j < 1$, and $\sum_{j=1}^{C} \pi_j = 1$ and f_j denotes an appropriate density given the characteristics of the error terms. As we will describe below, normally (Gaussian) distributed components appear to be appropriate in the context of the outcome of interest. We estimate the parameters of this model using maximum likelihood. Inference is based on standard errors adjusted for clustering at the county level.

Examination of the parameter estimates provides information on whether there is substantial heterogeneity across the distribution of the outcome and the treatment effects for each subpopulation. Although the parameter estimates, per se, do not characterize the component distributions, one can use Bayes Theorem in a process to classify individual observations into the components identified by the model estimates. More precisely, post estimation, we use Bayes Theorm to calculate the posterior probability that observation y_i belongs to component c (the prior probability is assumed to be a constant):

$$\Pr[i \in \text{class } c | \mathbf{x}_i, y_{i;} \theta] = \frac{\pi_c f_c(y_i | \mathbf{x}_i, \theta_c)}{\sum_{j=1}^{C} \pi_j f_j(y_i | \mathbf{x}_i, \theta_j)}, \qquad c = 1, 2, ..C. \tag{9}$$

Next, we assign each observation i into class j such that

$$j = \arg\max_c \Pr[y_i \in \text{class } c | \mathbf{x}_i, y_{i;} \theta.]$$

Then we treat the classification as known and stratify the sample by latent class in subsequent descriptive analysis of the relationship between observed covariates and class membership. This provides a characterization of the observations in each class.

3 Data and descriptive statistics

We use data from the 2003–2012 Behavioral Risk Factor Surveillance System (BRFSS) in our analysis.[1] The BRFSS is an annual, nationally representative survey of adults aged 18 and over conducted by the Centers for Disease Control and Prevention in collaboration with state health departments. The survey contains information on

[1]Unfortunately, for 2013 and going forward, the public-use BRFSS files do not contain county identifiers.

self-reported height and weight and extensive demographic and health-related information. Preliminary analyses suggested substantial heterogeneity by gender so we proceed by stratifying our analysis by gender throughout. We also restrict our sample to individuals 21 to 75 years old. We eliminate individuals over 75 years old because, in older adults, decreases in fat-free mass, increases in fat mass and loss of height alter the trajectory of BMI substantially (Villareal et al., 2005). BRFSS reports some implausibly small and large BMI values. To remove these outliers or possibly erroneous values of BMI, we drop observations with the highest and lowest half percent of BMI values separately for women and men. BRFSS also has a small number of extremely low and high sampling weights, which become especially troublesome after entropy balance reweighting, typically resulting in failure to achieve balance. Thus we drop observations with the smallest and largest quarter percent of sampling weights.

We use the Area Resource File (ARF) to classify each county of residence in each year by population and income. We drop all rural counties and small urban areas with populations less than 20,000. The remaining counties are classified as being urban (population greater than 20,000), small metropolitan areas (population less than 250,000), metropolitan areas (population between 250,000 and 1,000,000) and large metropolitan areas with populations greater than 1,000,000. We also control for household median income in each county. In addition, we use the County Business Patterns (CBP) data to measure the size of the limited service restaurant (LSR) sector – the primary target of the calorie labeling laws. To be specific, we measure the size of the limited restaurant sector by the number of employees in such establishments per 1,000 population in each county.

We obtained and cross-referenced county-specific information on legislation and implementation of calorie labeling laws from the National Conference of State Legislatures, the Center for Science in the Public Interest and MenuCalc, a web-based online nutrition analysis platform for the food industry endorsed by the National Restaurant Association. Table 1 shows the counties and states that enforced calorie-labeling laws at some point during the study period. Figures 1 and 2 show that geographies with calorie-labeling laws are concentrated in the Northeastern and Western regions of the US. Therefore, we restrict our analysis to states in those regions for both treatment and comparison observations. In addition to displaying treated counties, figures 1 and 2 also show the comparison states and counties in our study. The key variable in the analysis is an indicator labeled "County has law enforced" that takes a value of one if a county has implemented mandatory calorie labeling in the month and year of interview and zero otherwise. We pool the data from Northeast and Western US counties for the analysis in this paper.

Table 2 displays summary statistics for the samples of women and men separately and stratified according to whether the observation is assigned to a county that has a calorie-labeling law at any point during the sample period or not. Among women, there are 382,581 observations in counties without a law and 197,982 observations in counties with the law. Among men, the corresponding samples sizes are 263,603 and 135,153 respectively.

The summary statistics in table 2 show that the average BMI in treated and comparison counties are not substantially different for samples of women and men. In addition, treated and comparison observations have similar average age and rates of blacks, college graduates, middle income earners and pregnancy among women. On the other hand, treated observations are considerably more likely to be other minority, Hispanic and have low incomes and less likely to be married and have high incomes.

Although our regression analyses take these characteristics into account, it is preferable to minimize reliance on the parametric form of the regression analysis by balancing on covariates and mean BMI in pre-program years. This allows us to interpret the regression estimates as being *doubly robust*. As described above, we adjust the BRFSS sampling weights using entropy minimization to obtain new sampling weights for comparison observations while retaining the original sampling weights for the treated observations. We estimate the entropy-balance weights separately for each gender and each year of data. Sample means using the entropy-balance weights are shown in table 3. Now the samples are "perfectly" balanced. For example, 9.1

percent of Hispanic women lived in comparison counties while 25.9 lived in treated counties in the "unadjusted" analysis shown in table 2. After entropy-balancing, both treated and comparison counties have 25.9 percent Hispanic observations. The mean number of LSR employees is higher in treated counties than control counties when sampling weights are used in the calculations. Once entropy balance weights are used, treated and control counties have identical LSR densities. All other characteristics are similarly balanced.

4 Results

Figure 3 shows the trends in average BMI for treatment and control groups stratified by gender. The graphs in the left panels display means weighted by the BRFSS sampling weights. The graphs in the right panels display means weighted by entropy-balance weights. For the sample of women, the sampling-weighted means are close but show some divergence before calorie-labeling laws went into effect. For men, the sampling-weighted means are appear roughly parallel but somewhat apart in the pre-law period, obscuring possible gains of the calorie labeling laws. The use of entropy-balance weights brings the pre-treatment trends together and enhances the differences in post-treatment outcomes. In both cases, a widening of the BMI gap after 2008 is apparent.[2]

Results of the difference-in-difference entropy-balance weighted regressions for the samples of women and men are shown in Table 4. The effect of the calorie-labeling laws is not statistically significant for the sample of women. However, there is a significant decrease in BMI of about 0.31 among men that can be attributed to the calorie-labeling laws. This translates to about 1.1 percent of the average male BMI. Other coefficients in the regressions are generally significant and have expected signs. For example, BMI increases at a decreasing rate with age, blacks and Hispanics have higher BMI but other minorities have lower BMI. Education and income are

[2]Readers will observe that, for the sample of men, the trends converge sharply in 2012. It remains to be seen if this is a real feature, an artifact of some remaining outliers, or due to changes in values of predictive covariates.

both associated with lower BMI. Among county-level characteristics, while county-level income has the expected negative effect on BMI for both men and women, the density of LSR restaurants is not associated with BMI.

4.1 Finite mixture models

We estimate 2, 3, and 4 component finite mixture models and find that 3-component models have the best fit for the samples of men and women. Thus we proceed by focussing on 3-component models.

Table 5 and figure 4 show some basic characteristics of the three component densities for the full samples of women and men. For the sample of women, the model identifies a distribution centered on normal weight (mean BMI equals 22) with associated probability of 0.32, a distribution centered on overweight (mean BMI equals 27) with a probability of 0.48 and a relatively dispersed distribution centered on obese weight (mean BMI equals 34) with a probability of 0.21. The distributional characteristics of the mixture classes for the sample of men are broadly similar to those of women but have some substantive differences. The model identifies a distribution centered just under the clinical cutoff for overweight (BMI of 25) with associated probability of 0.42, a distribution centered on overweight (mean BMI equals 28) with a probability of 0.43 and a relatively dispersed distribution centered on obese weight (mean BMI equals 34) with a probability of 0.16.

Table 6 reports the difference-in-difference regression coefficients for each latent class. The results show that the reduction in BMI due to the calorie-labeling law is statistically significant only in the second component, which corresponds to a class of women whose BMI is centered on 27. The marginal effect of the law is 0.25, which corresponds to about 0.9 percent of the average BMI for this class of observations. The effects of the calorie-labeling laws are very different for men (see table 6). Calorie-labeling laws have an effect on each of the three classes. Men who belong to the borderline overweight class with a mean BMI of 25 and who are exposed to calorie-labeling laws decrease their BMI by a small, but statistically significant 0.14 units. This corresponds to just under 0.6 percent of the mean BMI for that class. Men in the overweight class, corresponding to the component distribution with a mean of 28 and a occurrence probability of 0.43 decrease their BMI by 0.40 when exposed to calorie-labeling laws. This effect corresponds to 1.4 percent of their average BMI. Men in the obese class decrease their BMI when exposed to calorie-labeling laws by a large and statistically significant 0.70 units – or 2 percent of the average BMI in the class.

For both women and men, there is an effect of calorie-labeling laws on a high weight subpopulation, but men in the morbidly obese subpopulation also experience substantial loss in BMI. These results are broadly consistent with Restrepo (2014) who finds a bigger effect at the higher quantiles of BMI and Bollinger et al. (2011) who find a much higher effect among individuals who already consume above average calories at Starbucks. The estimates of the finite mixture models, however, uncover a subtle yet important difference, which is that calorie-labeling has an affect on a substantial fraction of women who are overweight but appears to have no effect on a smaller but significant fraction of women who are, on average, obese. For men, it appears that the effect of calorie-labeling laws persists through the entire distribution but with the effect being larger (in absolute and percentage terms) in the overweight and obese classes.

4.2 Alternative specifications of the finite mixture models

Self-reported weight and BMI is fraught with measurement error. Cawley (2004) and Burkhauser and Cawley (2008) show that the measurement error is non-classical because, while individuals generally tend to underreport their weight, overweight individuals tend to underreport their weight more. Consequently, in our first check of our main results, we follow Cawley (2004) and predict measured BMI in our BRFSS sample using coefficients from a validation regression equation estimated using data from NHANES. To be precise, we use data from the 2005-2006, 2007-2008, and 2009-2010 rounds of the NHANES to estimate a regression of measured BMI on a quadratic function of reported BMI fully interacted with a quadratic in age and indicators for gender, black and other race and Hispanic ethnicity. We estimate a 3-class finite mixture model using the predicted, measured BMI and report key estimates in the top panel of table 7. The results for both samples of women and men are qualitatively unchanged. The point estimates for overweight and obese classes are somewhat larger, however.

Our second check of the specification of the model concerns the use of entropy balance weights. We have already shown that the use of these weights balances the treated and control samples perfectly in terms of observed covariates. When we estimate the finite mixture model using BRFSS sampling weights instead, the results, shown in the second panel of table 7, are very similar for the sample of women. For the sample of men, the effect for the lowest BMI class (borderline overweight) is no longer statistically significant, the effect for the overweight class is a bit smaller and the effect for the obese class substantially smaller than those obtained in the main specification.

BRFSS is a survey that has been conducted, until recently, exclusively via home (land line) telephone interviews. From 2011 onwards, the BRFSS sampling frame changed to allow for respondents on cellular phones, with just under 11 percent of interviews conducted on cellular phones in 2011 and 17 percent such interviews in 2012. In a check of the estimates to these, potentially different, individuals, we drop them from the sample and reestimate the finite mixture model. The results, shown in the third panel, are qualitatively identical and quantitatively very similar to our main results.

We generalize our difference-in-difference specification to include state-specific time trends. We have estimated finite mixture models with linear and quadratic time-trend specifications and report the results of the quadratic state-specific time trends specification in the fourth panel of table 7. Qualitatively, the results are very similar to the main specification. The point estimates for the overweight and obese latent classes are larger, especially for the sample of men. The results of the model specified with linear trends has the same characteristics.

In our final check of specification, we include the county mean BMI in 2003 as an additional regressor in the model. This is an atypical difference-in-difference specification but has the flavor of a panel-data regression with an initial condition (Wooldridge, 2005). The estimates, shown in the bottom panel, are qualitatively and quantitatively virtually identical to those obtained in our main specification.

4.3 Characterizing the latent classes

Estimates of the parameters of the finite mixture models reveal that the effects of calorie-labeling laws are heterogeneous and substantially affect overweight women and men, and, in addition, obese men. Nevertheless, calorie-labeling laws appear to have no effect on substantial portions of the population of women and only a small effect on a substantial portion of men. In order to determine whether there are observable characteristics (other than BMI) that might distinguish the classes of individuals, we classify the female and male samples separately using parameter estimates from the main specification of the model to estimate the posterior probabilities of class membership (equation 9). We use this estimated posterior classification to summarize the observed characteristics, using the entropy balance weights, of the samples by latent class. Differences in the standardized distributions of the characteristics are displayed in radar plots for the samples of women and men.

Radar plots of means of standardized covariates are shown in figure 5. For the samples of women and men, those in the latent class centered around overweight are likely to be older, high school graduates only, and Hispanic. Note that this is the class of observations with the consistently significant effect across samples and specifications. Women and men in the normal weight class, for whom there are generally statistically insignificant and/or small effects, are likely to be high income and college graduates, live in high income counties and more likely to be married. Finally, while there are no differences in the size of county that women live in on the basis of the BMI classes, men who belong to the obese class are more likely to live in small urban counties while men in the normal weight class are more likely to live in large metropolitan areas. There are virtually no differences in the states of residence of individuals in each of the three latent classes (not shown in the figures).

5 Conclusion

The results in this study indicate that mandatory calorie labeling laws implemented over the past few years in a number of states and counties appear to be having substantial effects in terms of decreased BMI following implementation of such laws. Estimates from linear difference-in-difference models show that calorie labels had no statistically significant effects on women. For men, the effect is small but statistically significant. These results are consistent with much of the literature, finding small effects at best. But, estimates from finite mixture models, designed to elicit potential heterogeneous effects, tell a different story. The estimates show that the effect of calorie labeling laws is substantively large and statistically significant for women and men whose BMI distributions are centered on overweight. For women and men whose BMI is centered on normal weight, the effects are small but statistically significant for men. The effects for women and men whose BMI is centered on obese are also substantively large but the effect is not statistically significant for women.

There are two likely explanations for why the largest effects of the calorie labeling laws are found for overweight women and overweight and obese men. First, if one assumes that the information asymmetry is constant across the population, calorie labeling laws may have largest impacts for the overweight and obese because they are most sensitive to the information. Calorie labeling laws have very small impacts among individuals who are normal weight because they have no reason to change their behavior on the basis of the new information. Second, it is possible that, while individuals in the overweight and obese classes are unaware of the caloric content of restaurant meals, individuals in the normal weight class are more aware of the caloric content. These results run counter to the argument that individuals who consume the most calories (more likely to be overweight or obese) are least likely to change their consumption behavior. In addition, if such high consumption individuals are, in fact, information sensitive, they may well also be price sensitive and sensitive to behavioral prompts.

Our analysis of the observed correlates of class membership shows that women and men in the overweight class are likely to be older, high school graduates only, and Hispanic. We speculate that such individuals might be either most uninformed or most likely to react to new information. Women and men in the normal weight class, for whom there are generally statistically insignificant and/or small effects, are likely to be high income and college graduates, live in high income counties and more likely to be married. Perhaps this justifies the emphasis of most calorie labeling laws on meals in limited service restaurants. Information provision at higher-priced, full service restaurants, frequented by individuals of higher socioeconomic status, may not have substantial new information content for those individuals.

The literature on the effects of calorie labeling at the point of purchase (treated) (e.g., Elbel et al., 2009; Dumanovsky et al., 2011; Bollinger et al., 2011) suggests that the effects of calorie labeling laws are small, at best. In contrast, we find substantial effects at population (intent-to-treat) levels. While it might be tempting to consider these results to be inconsistent with each other, they are, in fact, quite consistent with the view that such labeling is a “nudge” (Thaler and Sunstein, 2008). They may not change behavior a lot at the point of purchase, but do affect behavior substantially in a cumulative fashion and in non-fast food settings by increasing awareness of the calorie content of meals more broadly.

References

[1] Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of Californias Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493505.

[2] Abadie, A., and Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *The American Economic Review*, 93(1), 113132.

[3] Abadie, A., and Imbens, G. (2011). Bias-Corrected Matching Estimators for Average Treatment Effects. *Journal of Business and Economic Statistics*, 29(1).

[4] Anderson, M., and Matsa, D. (2011). Are Restaurants Really Supersizing America. *American Economic Journal: Applied Economics*, 3(1): 152-88.

[5] Area Resource File (ARF) Retrieved in January 2015 from http://ahrf.hrsa.gov

[6] Athey, S., and Imbens, G. (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, 74(2): 431-497.

[7] Ayyagari, P., Deb, P., Fletcher, J., Gallo, W., and Sindelar, J. (2013). Understanding heterogeneity in price elasticities in the demand for alcohol for older individuals, *Health Economics*, 22(1), 89 – 105.

[8] Basset, M., Dumanovsky, T., Huang C., Silver, L., Young, C., Nonas, C., Matte, T., Chideya, S., and Frieden, T. (2008). Purchasing Behavior and Calorie Information at Fast-Food Chains in New York City, 2007. *American Journal of Public Health*, 98(8): 1457-1459.

[9] Bleich, S., Cutler, D., Murray, C., and Adams, A. (2008). Why Is the Developed World Obese? *Annual Review of Public Health*, 29(1), 273295.

[10] Bollinger, B., Leslie, P., and Sorensen, A. (2011). Calorie Posting in Chain Restaurants. *American Economic Journal: Economic Policy*, 3(1): 91-128.

[11] Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica*, 78, 1375 – 1412.

[12] Burkhauser, R., and Cawley, J. (2008). Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics*, 27(2), 519529.

[13] Caudill, S., Gropper, D., and Hartarska, V. (2009). Which Microfinance Institutions Are Becoming More Cost Effective with Time? Evidence from a Mixture Model. *Journal of Money, Credit and Banking*, 41, 651 – 672.

[14] Cawley, J. (2004). The Impact of Obesity on Wages. *Journal of Human Resources*, XXXIX(2), 451474.

[15] Cawley, J. (2015). An economy of scales: A selective review of obesitys economic causes, consequences, and solutions. *Journal of Health Economics*, 43, 244268.

[16] Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Data. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, [2003-2012].

[17] Center for Science in the Public Interest (2015). State and Local Policies for Chain Restaurants. Retrieved in January 2015 from http://www.cspinet.org/menulabeling/state-local-policies.html

[18] Christian, T., and Rashad, I. (2009). Trends in U.S. food prices, 19502007. *Economics and Human Biology*, 7(1), 113120.

[19] County Business Patterns (CBP) Retrieved in January 2015 from http://www.census.gov/econ/cbp/

[20] 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(3): 32-63.

[21] Cutler, D., Glaeser, E., and Shapiro, J. (2003). Why Have Americans Become More Obese? *Journal of Economic Perspectives*, 17(3), 93118.

[22] Deb, P., and Trivedi, P. (1997). Demand for Medical Care by the Elderly: A Finite Mixture Approach, *Journal of Applied Econometrics*, 12, 313-336.

[23] Deb, P., Gallo, W., Ayyagari, P., Fletcher, J., and Sindelar, J. (2011). The Effect of Job Loss on Overweight and Drinking, *Journal of Health Economics*, 30, 317327.

[24] Dumanovsky, T., Huang, C., Nonas, C., Matte, T., Basset, M., and Silver, L. (2011). Changes in energy content of lunchtime purchases from fast food restaurants after introduction of calorie labelling: cross sectional customer surveys. *British Medical Journal*, 343:d4464.

[25] Elbel, B., Kersh, R., Brescoll, V., and Dixon, L. (2009). Calorie Labeling and Food Choices: A First Look at the Effects on Low-Income People in New York City. *Health Affairs*, 28(6): 111021.

[26] El-Gamal, M., and Grether, D. (1995). Are People Bayesian? Uncovering Behavioral Strategies, *Journal of the American Statistical Association*, 90, 1137 – 1145.

[27] Günther, I., and Launov, A. (2012). Informal employment in developing countries: Opportunity or last resort?. *Journal of Development Economics*, 97, 88 – 98.

[28] Hainmueller, J. (2011). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*.

[29] Heckman, J., and Singer, B. (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, 52(2), 271 – 320.

[30] Ho, D., Imai, K., King, G., and Stuart, E. (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, 15(3): 199-236.

[31] Iacus, S., King, G., Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis*, 20(1), 124.

[32] Lindsay, B. (1995). *Mixture Models: Theory, Geometry and Applications*, NSF-CBMS Regional Conference Series in Probability and Statistics, Institute of Mathematical Statistics.

[33] Lin, B., and Guthrie, J. (2012). Nutritional Quality of Food Prepared at Home and Away From Home, 1977-2008. EIB-105, *USDA, Economic Research Service.*

[34] McLachlan, G., and Peel, D. (2000). *Finite mixture models*, New York: John Wiley.

[35] MenuCalc (2015). Menu Labeling Laws. Retrieved in January 2015 from http://www.menucalc.com/menulabeling.aspx

[36] National Conference of State Legislatures (2013). Trans Fat and Menu Labeling Legislation. Retrieved in January 2015 from http://www.ncsl.org/research/health/trans-fat-and-menu-labeling-legislation.aspx

[37] Restrepo, B. (2014). Calorie labeling in chain restaurants and body weight: evidence from New York. *Working Paper. European University Institute.*

[38] Robert Wood Johnson Foundation. (2009). Menu Labeling: Does Providing Nutrition Information at the Point of Purchase Affect Consumer Behavior?

[39] Thaler, R., Sunstein, C. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Yale University Press.

[40] Villareal, D., Apovian, C., Kushner, R., and Klein, S. (2005). Obesity in older adults: technical review and position statement of the American Society for Nutrition and NAASO, The Obesity Society. *The American Journal of Clinical Nutrition*, 82: 923-934.

[41] Wedel, M., and DeSarbo, W. (1993). A Latent Class Binomial Logit Methodology for the Analysis of Paired Comparison Choice Data, *Decision Sciences*, 24, 1157 – 1170.

[42] Wisdom, J., and Loewenstein, G. (2010). Promoting Healthy Choices: Information versus Convenience. *American Economic Journal, Applied Economics*, 2(2): 164-178.

[43] Wooldridge, J. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20(1), 3954.

Table 1: Sample geography by calorie-labeling law status and dates

Date	State	County
8/2008	New York	Queens
8/2008	New York	Kings
8/2008	New York	Richmond
8/2008	New York	Bronx
8/2008	New York	New York City
6/2009	New York	Westchester
11/2009	New York	Ulster
7/2009	California	Statewide
1/2009	Washington	King
8/2009	Oregon	Multnomah
4/2010	New York	Albany
10/2010	New York	Schenectady
11/2010	New York	Suffolk
2/2010	Pennsylvania	Philadelphia
5/2010	Maine	Statewide
11/2010	Massachusetts	Statewide
1/2011	Oregon	Statewide

The following states with no calorie-labeling laws through 2012 are included in our control group: Connecticut, Delaware, New Hampshire, Maryland, New Jersey, Rhode Island, Vermont, Arizona, Idaho, Nevada.

Table 2: Sample means by gender and calorie-labeling law status

	Women		Men	
	No law	Law	No law	Law
Body mass index	26.870	26.735	27.974	27.610
County has law enforced	0.102	0.374	0.105	0.383
Age / 10	4.620	4.544	4.585	4.524
Age (/10) squared	2.432	2.477	2.389	2.380
Black race	0.088	0.087	0.077	0.070
Other minority	0.060	0.104	0.068	0.115
Hispanic ethnicity	0.091	0.259	0.091	0.235
Married	0.613	0.564	0.655	0.615
Less than High School	0.070	0.133	0.076	0.128
High school graduate	0.271	0.213	0.276	0.217
Some college	0.287	0.269	0.252	0.248
Income < $15,000	0.069	0.137	0.051	0.096
Income $15,000 - $25,000	0.123	0.131	0.105	0.122
Income $25,000 - $35,000	0.093	0.091	0.088	0.088
Income $35,000 - $50,000	0.134	0.117	0.135	0.120
Income $50,000 - $75,000	0.161	0.143	0.169	0.148
Income unknown	0.117	0.082	0.089	0.061
Pregnant	0.017	0.019		
Metro county pop. 250K-1M	0.235	0.173	0.240	0.177
Metro county pop. < 250K	0.085	0.045	0.087	0.047
Urban county pop. > 20K	0.067	0.021	0.067	0.022
County median HH income ($10K)	5.706	5.609	5.714	5.679
County LSR employees per 1000 pop.	9.269	9.651	9.271	9.687
Arizona	0.106	0.000	0.106	0.000
California	0.000	0.583	0.000	0.569
Idaho	0.019	0.000	0.021	0.000
Nevada	0.044	0.000	0.048	0.000
Oregon	0.000	0.062	0.000	0.069
Washington	0.086	0.030	0.089	0.037
Connecticut	0.067	0.000	0.068	0.000
Delaware	0.017	0.000	0.017	0.000
Maine	0.000	0.016	0.000	0.017
Maryland	0.109	0.000	0.104	0.000
New Hampshire	0.023	0.000	0.023	0.000
New Jersey	0.163	0.000	0.164	0.000
New York	0.146	0.173	0.137	0.165
Pennsylvania	0.194	0.023	0.197	0.021
Rhode Island	0.021	0.000	0.021	0.000
Vermont	0.006	0.000	0.006	0.000
N	382,581	197,982	263,603	135,153

Means calculated using sampling weights.

Table 3: Balanced sample means by gender and calorie-labeling law status

	Women		Men	
	No law	Law	No law	Law
Body mass index	26.950	26.735	27.770	27.610
County has law enforced	0.091	0.374	0.066	0.383
Age / 10	4.544	4.544	4.524	4.524
Age (/10) squared	2.477	2.477	2.380	2.380
Black race	0.087	0.087	0.070	0.070
Other minority	0.104	0.104	0.115	0.115
Hispanic ethnicity	0.259	0.259	0.235	0.235
Married	0.564	0.564	0.615	0.615
Less than High School	0.133	0.133	0.128	0.128
High school graduate	0.213	0.213	0.217	0.217
Some college	0.269	0.269	0.248	0.248
Income < $15,000	0.137	0.137	0.096	0.096
Income $15,000 - $25,000	0.131	0.131	0.122	0.122
Income $25,000 - $35,000	0.091	0.091	0.088	0.088
Income $35,000 - $50,000	0.117	0.117	0.120	0.120
Income $50,000 - $75,000	0.143	0.143	0.148	0.148
Income unknown	0.082	0.082	0.061	0.061
Pregnant	0.019	0.019		
Metro county pop. 250K-1M	0.173	0.173	0.177	0.177
Metro county pop. < 250K	0.045	0.045	0.047	0.047
Urban county pop. > 20K	0.021	0.021	0.022	0.022
County median HH income ($10K)	5.609	5.609	5.679	5.679
County LSR employees per 1000 pop.	9.651	9.651	9.687	9.687
Arizona	0.184	0.000	0.156	0.000
California	0.000	0.583	0.000	0.569
Idaho	0.019	0.000	0.019	0.000
Nevada	0.079	0.000	0.099	0.000
Oregon	0.000	0.062	0.000	0.069
Washington	0.054	0.030	0.029	0.037
Connecticut	0.060	0.000	0.055	0.000
Delaware	0.022	0.000	0.004	0.000
Maine	0.000	0.016	0.000	0.017
Maryland	0.085	0.000	0.187	0.000
New Hampshire	0.007	0.000	0.008	0.000
New Jersey	0.150	0.000	0.144	0.000
New York	0.094	0.173	0.105	0.165
Pennsylvania	0.206	0.023	0.156	0.021
Rhode Island	0.039	0.000	0.038	0.000
Vermont	0.002	0.000	0.002	0.000
N	382,581	197,982	263,603	135,153

Means calculated using entropy-balance weights

Table 4: Difference-in-difference regressions of BMI by gender

	Women	Men
County has law enforced	-0.129	-0.309***
	(0.079)	(0.064)
County has law	-0.474***	-0.393**
	(0.152)	(0.170)
Age / 10	0.225***	0.047***
	(0.019)	(0.018)
Age (/10) squared	-0.234***	-0.211***
	(0.011)	(0.008)
Black race	2.463***	0.455***
	(0.085)	(0.112)
Other minority	-0.699***	-1.173***
	(0.155)	(0.095)
Hispanic ethnicity	0.831***	0.405***
	(0.105)	(0.091)
Married	-0.388***	0.361***
	(0.056)	(0.054)
Less than High School	2.251***	0.960***
	(0.102)	(0.078)
High school graduate	1.521***	1.093***
	(0.067)	(0.062)
Some college	1.282***	1.086***
	(0.058)	(0.062)
Income < $15,000	1.528***	-0.114
	(0.123)	(0.117)
Income $15,000 - $25,000	1.267***	-0.256**
	(0.080)	(0.102)
Income $25,000 - $35,000	1.032***	-0.157**
	(0.080)	(0.077)
Income $35,000 - $50,000	0.992***	-0.048
	(0.095)	(0.060)
Income $50,000 - $75,000	0.806***	0.136***
	(0.049)	(0.051)
Income unknown	-0.004	-0.354***
	(0.077)	(0.068)
Pregnant	1.169***	
	(0.151)	
Metro county pop. 250K-1M	0.211**	0.218***
	(0.095)	(0.080)
Metro county pop. < 250K	-0.010	0.045
	(0.143)	(0.111)
Urban county pop. > 20K	0.065	-0.293
	(0.150)	(0.181)
County median HH income ($10K)	-0.188***	-0.097***
	(0.031)	(0.032)
County LSR employees per 1000 pop.	-0.003	-0.034
	(0.022)	(0.024)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$
Coefficients on state and year indicators not shown

Table 5: Component characteristics of the finite mixture model

	Women			Men		
	Class-1	Class-2	Class-3	Class-1	Class-2	Class-3
μ	22.316	26.772	33.929	24.742	28.269	33.979
σ	2.115	3.226	5.887	2.429	3.070	5.013
	(0.059)	(0.041)	(0.037)	(0.065)	(0.078)	(0.060)
π	0.320	0.475	0.205	0.417	0.427	0.156
	(0.022)	(0.012)	(0.012)	(0.050)	(0.033)	(0.020)

Table 6: Coefficients from the finite mixture model

	Women			Men		
	Class-1	Class-2	Class-3	Class-1	Class-2	Class-3
County has law enforced	0.017	-0.251***	-0.329	-0.135**	-0.399***	-0.695***
	(0.069)	(0.078)	(0.264)	(0.065)	(0.152)	(0.221)
County has law	-0.172**	-0.495***	-0.630**	-0.082	-0.636**	-0.293
	(0.084)	(0.170)	(0.295)	(0.128)	(0.277)	(0.358)
Age / 10	0.248***	0.335***	-0.131***	0.145***	0.030	-0.177**
	(0.022)	(0.029)	(0.044)	(0.025)	(0.033)	(0.070)
Age (/10) squared	-0.072***	-0.205***	-0.445***	-0.143***	-0.225***	-0.265***
	(0.012)	(0.023)	(0.038)	(0.015)	(0.018)	(0.033)
Black race	1.326***	2.552***	2.981***	0.201**	0.745***	0.702**
	(0.117)	(0.118)	(0.258)	(0.089)	(0.134)	(0.276)
Other minority	-0.426***	-0.493**	-1.486***	-0.807***	-1.290***	-2.072***
	(0.061)	(0.211)	(0.328)	(0.104)	(0.144)	(0.255)
Hispanic ethnicity	0.963***	0.924***	0.113	0.856***	0.371*	-0.880***
	(0.082)	(0.132)	(0.223)	(0.076)	(0.212)	(0.215)
Married	0.155***	-0.281***	-1.543***	0.580***	0.363***	-0.354**
	(0.038)	(0.053)	(0.144)	(0.061)	(0.111)	(0.156)
Less than High School	0.778***	2.593***	2.954***	0.041	1.611***	1.626***
	(0.162)	(0.107)	(0.261)	(0.108)	(0.165)	(0.288)
High school graduate	0.418***	1.768***	2.356***	0.207	1.686***	1.643***
	(0.062)	(0.081)	(0.171)	(0.135)	(0.101)	(0.183)
Some college	0.200***	1.465***	2.155***	0.247**	1.542***	1.849***
	(0.061)	(0.075)	(0.186)	(0.100)	(0.158)	(0.170)
Income < $15,000	-0.062	1.438***	3.239***	-1.252***	0.089	2.053***
	(0.079)	(0.182)	(0.307)	(0.153)	(0.304)	(0.315)
Income $15,000 - $25,000	-0.002	1.294***	2.459***	-0.926***	-0.075	1.054***
	(0.064)	(0.166)	(0.208)	(0.100)	(0.165)	(0.293)
Income $25,000 - $35,000	0.050	1.055***	2.204***	-0.491***	-0.148	0.634***
	(0.081)	(0.152)	(0.276)	(0.110)	(0.115)	(0.243)
Income $35,000 - $50,000	0.135**	1.092***	1.752***	-0.412***	0.027	0.651***
	(0.054)	(0.174)	(0.203)	(0.078)	(0.113)	(0.176)
Income $50,000 - $75,000	0.077	0.896***	1.582***	-0.194***	0.353***	0.405**
	(0.064)	(0.092)	(0.164)	(0.070)	(0.124)	(0.186)
Income unknown	-0.253***	-0.019	0.344	-0.460***	-0.385***	0.099
	(0.058)	(0.099)	(0.238)	(0.078)	(0.125)	(0.267)
Pregnant	0.580***	1.613***	0.781*			
	(0.098)	(0.230)	(0.420)			
Metro county pop. 250K-1M	0.090**	0.340***	0.114	0.028	0.398***	0.227
	(0.044)	(0.122)	(0.257)	(0.063)	(0.130)	(0.211)
Metro county pop. < 250K	-0.066	-0.046	0.170	-0.058	-0.052	0.643*
	(0.077)	(0.166)	(0.300)	(0.089)	(0.136)	(0.334)
Urban county pop. > 20K	-0.093	0.220	0.106	-0.469*	0.005	-0.204
	(0.075)	(0.186)	(0.338)	(0.269)	(0.461)	(0.456)
County median HH income ($10K)	-0.045***	-0.194***	-0.248***	-0.029	-0.127***	-0.186***
	(0.017)	(0.039)	(0.077)	(0.027)	(0.046)	(0.064)
County LSR employees per 1000 pop.	-0.020	-0.026	0.030	-0.027*	-0.038	-0.066
	(0.013)	(0.022)	(0.047)	(0.015)	(0.032)	(0.052)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Coefficients on state and year indicators not shown

Table 7: Coefficients from alternative specifications

	Women			Men		
	Class-1	Class-2	Class-3	Class-1	Class-2	Class-3
	Model of predicted, measured BMI					
County has law enforced	0.019	-0.270***	-0.285	-0.124**	-0.424***	-0.780***
	(0.074)	(0.082)	(0.222)	(0.060)	(0.146)	(0.240)
π	0.296	0.477	0.227	0.447	0.414	0.140
	(0.024)	(0.012)	(0.014)	(0.029)	(0.021)	(0.011)
μ	22.971	27.521	34.466	25.229	28.846	34.834
	Model uses unadjusted BRFSS sampling weights					
County has law enforced	-0.026	-0.185***	-0.132	-0.049	-0.307***	-0.335**
	(0.055)	(0.067)	(0.146)	(0.049)	(0.103)	(0.168)
π	0.337	0.460	0.203	0.415	0.433	0.152
	(0.008)	(0.005)	(0.005)	(0.019)	(0.013)	(0.008)
μ	22.417	26.910	33.939	24.788	28.350	34.219
	Sample drops cell phone interview observations					
County has law enforced	0.021	-0.220***	-0.333	-0.150**	-0.413**	-0.663**
	(0.070)	(0.081)	(0.299)	(0.066)	(0.163)	(0.266)
π	0.320	0.476	0.204	0.415	0.430	0.155
	(0.024)	(0.013)	(0.012)	(0.044)	(0.030)	(0.016)
μ	22.310	26.760	33.925	24.723	28.256	33.966
	Model includes State-specific quadratic time trends					
County has law enforced	-0.041	-0.278***	-0.497	-0.067	-0.629**	-1.198***
	(0.095)	(0.104)	(0.431)	(0.104)	(0.250)	(0.334)
π	0.321	0.475	0.204	0.433	0.416	0.152
	(0.022)	(0.012)	(0.012)	(0.051)	(0.034)	(0.020)
μ	22.322	26.781	33.946	24.800	28.364	34.071
	Model includes county mean BMI in 2003					
County has law enforced	0.021	-0.245***	-0.307	-0.124*	-0.363**	-0.655***
	(0.068)	(0.078)	(0.265)	(0.068)	(0.149)	(0.218)
π	0.321	0.475	0.204	0.406	0.432	0.162
	(0.022)	(0.012)	(0.012)	(0.054)	(0.038)	(0.021)
μ	22.325	26.787	33.934	24.697	28.193	33.830

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Figure 1: Eastern US geography of calorie-labeling laws

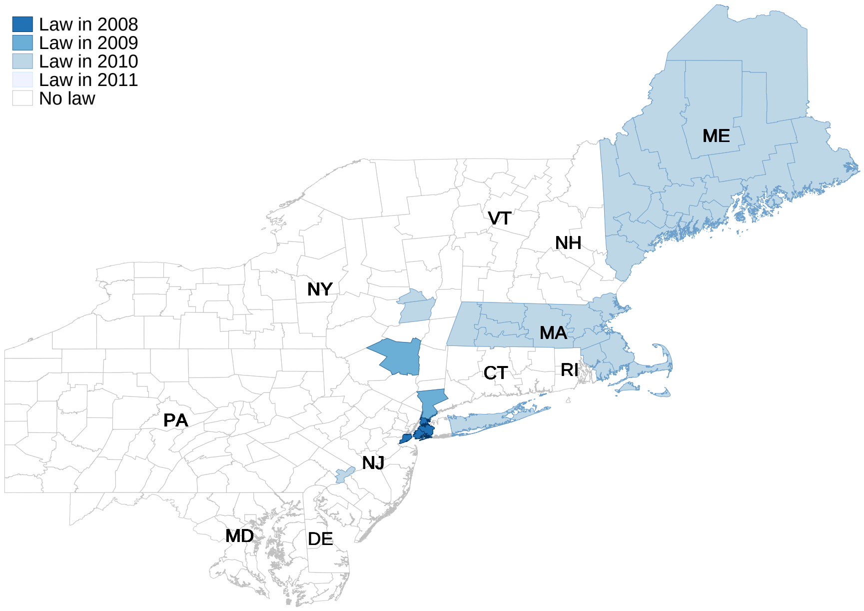


Figure 2: Western US geography of calorie-labeling laws

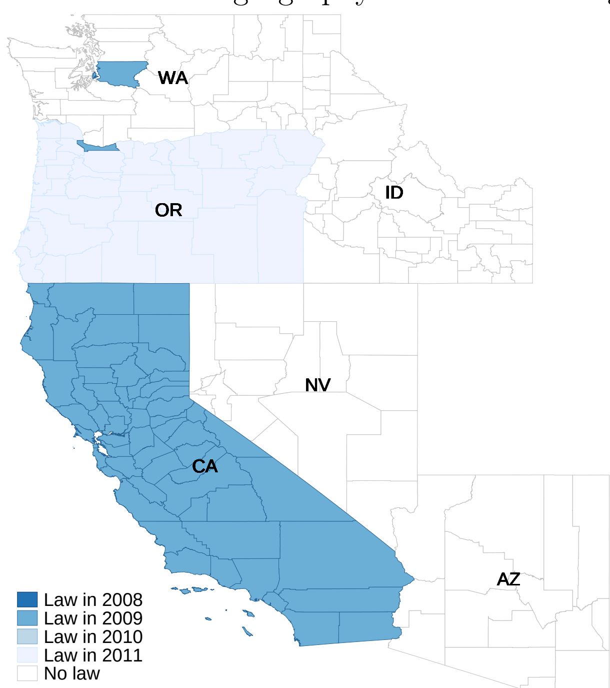


Figure 3: Average BMI trends by gender

Figure 4: Characteristics of the BMI distribution by latent class

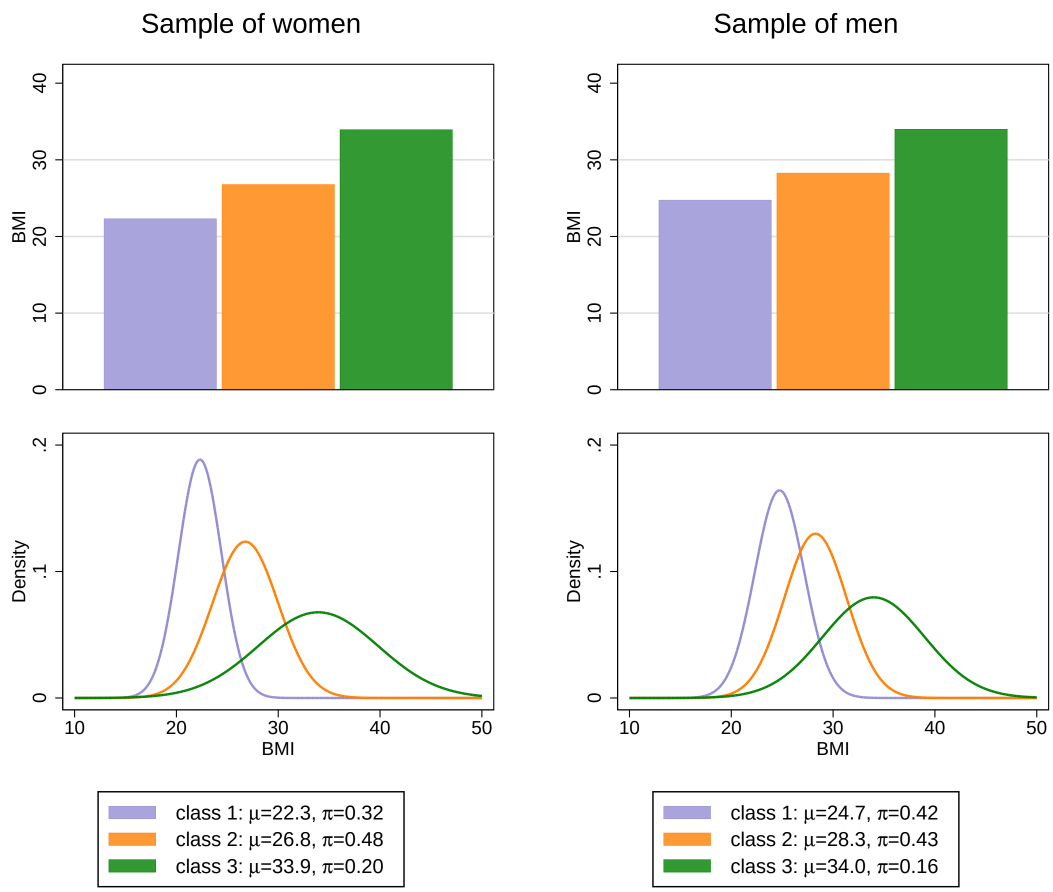
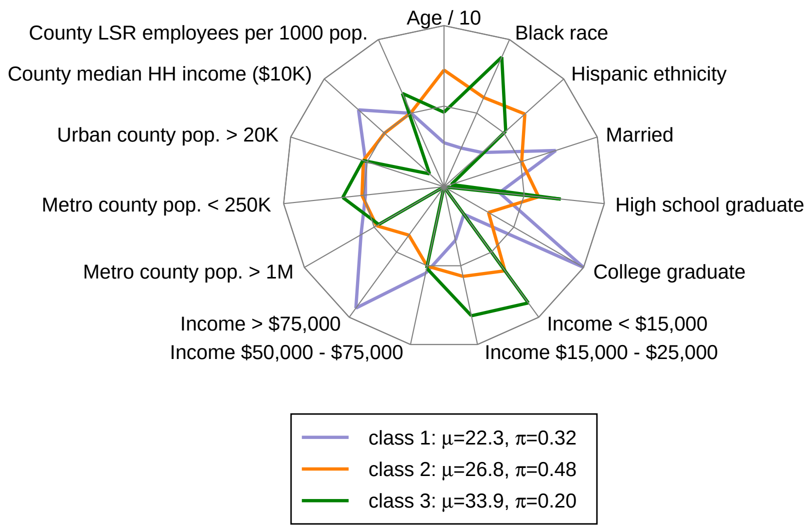


Figure 5: Mean covariate characteristics by latent class

Sample of women



Sample of men